FAKULTÄT
FÜR !NFORMATIK
Faculty of Informatics

Fault-Tolerant Clock Distribution in Grid-Like Networks

DISSERTATION

zur Erlangung des akademischen Grades

Doktor der Technischen Wissenschaften

eingereicht von

Dipl.-Ing. Martin Perner, BSc
Matrikelnummer 00725782

an der Fakultät für Informatik
der Technischen Universität Wien

Betreuung: Univ.-Prof. Dr. Ulrich Schmid

Diese Dissertation haben begutachtet:

Univ.-Prof. Dr. Axel Jantsch

Prof. Dr.-Ing. Miloš Krstić

Wien, 29. August 2019

Martin Perner

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.ac.at

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

FAKULTÄT
FÜR !NFORMATIK
Faculty of Informatics

Fault-Tolerant Clock Distribution in Grid-Like Networks

DISSERTATION

submitted in partial fulfillment of the requirements for the degree of

Doktor der Technischen Wissenschaften

by

Dipl.-Ing. Martin Perner, BSc
Registration Number 00725782

to the Faculty of Informatics
at the TU Wien

Advisor: Univ.-Prof. Dr. Ulrich Schmid

The dissertation has been reviewed by:

Univ.-Prof. Dr. Axel Jantsch	Prof. Dr.-Ing. Miloš Krstić

Vienna, 29th August, 2019

Martin Perner

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Bibliothek
Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

ERKLÄRUNG ZUR VERFASSUNG DER ARBEIT

Dipl.-Ing. Martin Perner, BSc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 29. August 2019

Martin Perner

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

ACKNOWLEDGEMENTS

> If I have seen further it is by standing on the shoulders of Giants.
>
> — Issac Newton

The first shoulder I have to thank is Ulrich Schmid. Being my advisor for many years, he sparked the interest for distributed computing[1] in me. His door was always open for any problem I had, and a solution seemed never out of reach.

Further, my gratitude goes to my (former) colleagues at the Embedded Computing Systems Group for the many hours of shared lunches and enlightening conversations that made this time so much more joyful: Florian Huemer, Jürgen Maier, Roman Kuznets, Robert Najvirt, Varadan Savulimedu Veeravalli, and Martin Zeiner. Especially, I like to thank Matthias Függer, Alexander Kössler, Josef Widder, and Kyrill Winkler for their support in our shared teaching duties.

The institute technicians Karl Malle, Dieter Etz, Heiner Odebrecht, and Thomas Juranek, who kept the environment running, were also a pleasure to work and laugh with. A special thanks goes to Heinz Deinhart for his energetic way of handling problems in the teaching laboratory. And to Traude Sommer, for her pragmatic way at approaching administrative things.

Finally, the first and last shoulders of my life, that provided me support during this time, are my friends and family, especially Julia.

---

[1] This work has been supported by the Austrian Science Fund (FWF) project SIC (P26436) and RiSE (S11405).

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

For Julia
and those who inspired it
and will not read it

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

ABSTRACT

This thesis investigates how to design a system to distribute and use a clock signal in a grid-like network, as employed in, e.g., SoCs and NoCs.

Ideally, the clock is generated by multiple components and has self-stabilization and fault-tolerance properties, i.e., recovers from transient faults even under permanent faults. These clock generation components can be distributed over the whole chip to provide a clock signal for every part of the chip. This, however, comes at a price: The increase in wire length between the components increases the skew between the generating components, and hence the clock skew at the clock boundaries between the components. Furthermore, such clock generation systems are expensive in terms of chip area, and usually require a fully-connected topology. Hence, putting the clock generating components in close proximity is favorable. However, just generating such a clock somewhere in a chip is not enough: After all, the clock signal needs to be distributed to all components that need it.

The challenge for clock distribution is to not diminish or even remove the self-stabilization and fault-tolerance properties of the generated clock during distribution. To this end, we present a clock distribution system, HEX, that allows recovery from transient faults under permanent faults.

Since HEX, which is based on a hexagonal grid, shows a large spread in the node-to-node skew under faults, we explored alternative interconnection topologies. Under the assumption that a suitable clock generation is able to provide a skew bounded by the difference in the wire delay between its components, the TRIX topology has been identified as the best trade-off between performance and implementation efficiency. For TRIX, a transistor cell model of the clock distribution node has been constructed as well.

Given that the distributed clock propagates like a wave through the grid, it provides a synchronization source for the nodes of the grid already. However, the achievable degree of synchrony is not enough to be able to utilize the synchronous design paradigm for communication between the nodes of the grid. We show that high-speed communication is nevertheless feasible in such multi-synchronous GALS architectures, however, by using a FIFO buffer for mitigating the clock skew and thus allowing data transmission in every clock cycle. Using a special buffer management approach, our communication scheme can be guaranteed to be self-stabilizing when the underlying clocking system is. Hence, our communication scheme is fully compatible with, but not limited to, HEX and TRIX.

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

KURZFASSUNG

Diese Arbeit beschäftigt sich mit der Frage, wie man ein Taktsignal in gird-artigen Netzwerken—wie sie in SoCs und NoCs verwendet werden—auf einem Chip verteilen und nutzen kann.

Idealerweise wird dieses Taktsignal von mehreren Komponenten im Verbund erzeugt und besitzt dadurch sowohl selbst-stabilisierende als auch fehlertolerante Eigenschaften. Das bedeutet, dass sich die Komponenten von transienten Fehlern erholen können, selbst wenn einige Komponenten permanent fehlerhaft sind.

Die Komponenten des Verbunds können dabei über den ganzen Chip verteilt sein. Da das erzeugte Taktsignal zu jeder Funktionseinheit des Chips gebracht werden muss, ist dies vorteilhaft. Allerdings hat diese Verteilung auch Nachteile: Durch die langen Leitungen zwischen den Komponenten vergrößert sich der zeitliche Versatz des erzeugten Taktsignals. Dies führt in weiterer Folge zu einem zeitlichen Versatz zwischen benachbarten Funktionseinheiten des Chips, wenn diese von verschiedenen Komponenten ihr Taktsignal beziehen. Da die Komponenten üblicherweise auch vollständig untereinander verbunden sein müssen, benötigten sie viel Fläche am Chip. All das spricht dafür, die Komponenten des Verbundes nahe beieinander zu platzieren. Allerdings müssen die Taktsignale weiterhin allen Funktionseinheiten des Chips zugänglich gemacht werden.

Die Herausforderung bei der Taktverteilung ist es, die selbst-stabilisierenden und fehlertoleranten Eigenschaften des erzeugten Taktes aufrecht zu erhalten. Dazu stellen wir das Taktverteilungssystem HEX vor, welches an den Knoten des Grids Zellen hat, über die nahegelegene Funktionseinheiten mit einem Taktsignal versorgt werden. HEX ist in der Lage, sich von transienten Fehlern selbst unter bestehenden permanenten Fehlern zu erholen.

Da HEX, welches auf einem hexagonalen Raster basiert, jedoch im Fehlerfall einen hohen zeitlichen Versatz zwischen seinen Knotenzellen aufweisen kann, werden alternative Verbindungstopologien untersucht. Unter der Annahme, dass der Takterzeugungsverbund einen Takt mit einem maximalen zeitlichen Versatz in der Größenordnung der Signallaufzeiten zwischen seinen Komponenten erzeugen kann, wurde die TRIX Topologie ausgewählt. TRIX ist ein Optimum zwischen der Qualität der Taktverteilung und des Flächenbedarfs am Chip. Für TRIX wurde auch ein Transistormodell der Knotenzelle entwickelt.

Da sich ein so verteiltes Taktsignal wie eine Welle durch das Grid ausbreitet, stellt es auch eine Synchronisationsquelle für die Knotenzellen dar. Jedoch ist der erreichbare

Grad der Synchronisierung nicht ausreichend, um das synchrone Design-Paradigma für die Kommunikation zwischen den Funktionseinheiten, die von benachbarten Knotenzellen versorgt werden, zu ermöglichen. In dieser Arbeit wird gezeigt, dass dennoch schnelle Kommunikation in derartigen multi-synchronen GALS-Architekturen möglich ist. Dabei wird ein FIFO-Puffer verwendet, um den zeitlichen Versatz zu kompensieren und dadurch Datentransfers mit jeder Taktflanke zu ermöglichen. Durch die Verwendung einer speziellen Pufferverwaltung ist das Verfahren garantiert selbst-stabilisierend, wenn die Takterzeugung und -verteilung dies auch sind. Unser Kommunikationsschema ist daher voll kompatible mit HEX und TRIX, aber natürlich nicht nur darauf beschränkt.

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

CONTENTS

Abstract **xi**

Kurzfassung **xiii**

Contents **xvi**

1 Preface **1**
1.1 Organization 4

2 Clocking a System **9**
2.1 The Role of Time 10
2.2 Metastability 20
2.3 Failure Models 23
2.4 Clock Generation 29
2.5 Clock Distribution 37
2.6 Communication across Clock Domains 41

3 HEX – A Clock Distribution Scheme **47**
3.1 Algorithm & Topology 48
3.2 Skew & Resilience Analysis 50
3.3 Implementation 71
3.4 Simulation Experiments 79
3.5 Fast Clocks 98
3.6 Conclusions 110

4 Alternative Topologies for Clock Distribution Grids **113**
4.1 Alternative Algorithms and Topologies 114
4.2 Methodology 121
4.3 Simulation Results 130
4.4 Conclusion 170

5 TRIX – A Transistor-level Implementation **173**
5.1 Topology & Pulse Forwarding 174
5.2 Firing Logic Implementation 176

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

5.3 Simulations 178
5.4 Conclusion 184

6 Point-to-point Communication in Multi-synchronous Architectures 193
6.1 Clock Model 194
6.2 Description of the Communication Scheme 197
6.3 Analysis 198
6.4 Single-Burst Memory Parameterization 204
6.5 Multi-Burst Constraints – Writer before Reader 209
6.6 Multi-Burst Constraints – Reader before Writer 221
6.7 Evaluation 225
6.8 Conclusions 228

7 Conclusions and Future Work 231
7.1 Future Work 232

A Additional Visualizations of the Simulations of Chapter 4 237
A.1 Plots of the Pulse-based Algorithms in $\mathcal{G}_H$ 237
A.2 Plots of the Pulse-based Algorithms in $\mathcal{G}_A$ 242
A.3 Plots of the Pulse-based Algorithms in $\mathcal{G}_B$ 249
A.4 Plots of the Pulse-Based Algorithms in $\mathcal{G}_C$ 252
A.5 Plots of the Time-based Algorithms with $d = 2$ in $\mathcal{G}_C$ 261
A.6 Plots of the Time-based Algorithms with $d = 1$ in $\mathcal{G}_C$ 266
A.7 Tables of the Time-based Algorithms with $d = 1$ in $\mathcal{G}_C$ 271
A.8 Plots of the Pulse-based Algorithms in $\mathcal{G}_D$ 274
A.9 Plots of the Time-based Algorithms with $d = 0$ and $d = 1$ in $\mathcal{G}_D$ 281
A.10 Plots of the Pulse-based Algorithms in $\mathcal{G}_E$ 290
A.11 Plots of the Pulse-based Algorithms in $\mathcal{G}_F$ 299
A.12 Table for the Pulse-based Algorithms in $\mathcal{G}_F$ 316
A.13 Plots with the 2–local fault Model in $\mathcal{G}_C$ 325
A.14 Tables with the 2–local fault Model in $\mathcal{G}_C$ 342
A.15 Plots with the 2–local fault Model in $\mathcal{G}_D$ 353
A.16 Tables with the 2–local fault Model in $\mathcal{G}_D$ 360

B Additional Plots of Chapter 5 365

Glossary 369

Acronyms 373

Bibliography 377

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

CHAPTER 1

PREFACE

> The amount of logically irrelevant engineering detail inherent in the design and construction of a computer system is great.
>
> — WESLEY A. CLARK, Macromodular computer systems

OUR SOCIETY'S INCREASING dependence on technological systems has reached an unprecedented scale. While we enjoy many benefits to our civilization, the challenges in miniaturizing new technology are mainly due to details which, up until now, were insignificant. In 1965, Moore postulated what became to be known as *Moore's Law* [Moo65], namely, that the number of transistors on an *integrated circuit* (IC) doubles every two years.[1] At that time, the challenges in *very-large-scale integration* (VLSI) design were the lithographic processes needed for the required miniaturization and the conceptual complexity of designing new circuits, caused by the steady growth in complexity and size.

Today, the challenges are mainly due to the reduced size of the structures, which are nowadays in the nm range. The biggest challenges still lie in the lithographic process. The key issues have shifted from achieving miniaturization to handling the issues arising from it, however: With structures that are only a few atom layers thick, we have reached areas where classical mechanics are at their end and quantum mechanics comes into play. For example, quantum tunneling allows an electron, with a certain probability, to pass through a thin insulator. Even with all of these challenges arising, Moore's law still holds true, although this progress is mainly attributable to multi-core processors [Moo06].

---

[1] Initially, the "law" considered a doubling every year, which was later revised to two years, however.

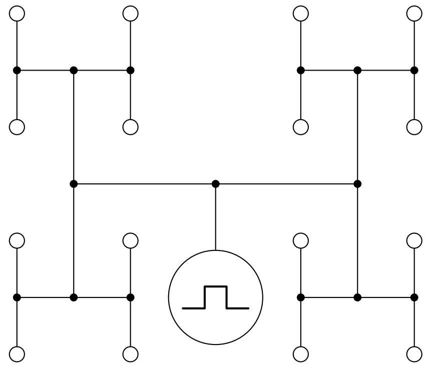


Figure 1.1: A H-tree of depth 4 with a clock source at its root. The white circles represent the leaves of the tree.

One of the challenges is related to the propagation of electrons in the chip. In essence, an IC consists of conducting wires and transistors, which are used as switches. In the traditional and predominant design model for ICs, a common, global, clock signal is required to allow the circuitry to work in unison. In the beginning, ICs were provided with a clock signal from a single oscillator, with a fixed frequency. This single clock source is a simple implementation of a global clock. The predominant circuit design model, the synchronous model, relies heavily on the global clock: It uses the abstraction that, at the beginning of every period of the global clock, all computations in the circuit start instantaneously. Hence, it is assumed that the computations will be completed, with all components in the circuit in a stable state, before the next clock period commences. However, with rising frequencies of the global clock, the limits of the propagation of electrical signals[2] come into play. Additionally, due to the miniaturization, the width of the wires in the circuit decrease, causing an increase in capacitances. Hence, the slew-rate-induced signal delays over the wires become now non-negligible. With all these signal delays between the transistors added to the switching times of the transistors, the convenient abstraction of the synchronous model crumbles: the clock propagates similar to a wave, rather than being at every point of the circuit at the same instant of time. To continue using the synchronous model, it must be ensured that every part of the circuit is in a stable state when the clock wave passes through it.

This is achieved by the usage of H-trees for the distribution of the clock on the chip. A H-tree is, as shown in Figure 1.1, a recursive H-shaped wiring; to be precise it is a fractal. The clock is feed to the H-tree at the root and supplied to the components in the circuit at the leaves at the lowest recursion level. Careful design of the clock tree leads to an equal distance of every leaf to the root of the tree.

However, a H-tree is just an ideal construction. In reality, it does not only comprise

---

[2] The propagation of an electrical wave, such as a clock signal, is limited by the speed of light, due to special relativity [Ein05]. Additionally, the material in which the signal propagates further limits the velocity. Typically, the signal velocity in an electrical circuit is between $c/2$ and $c/3$ [RCN08].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

of simple wires. Every wire has to be modeled via a resistor/capacitor combination, which needs to be periodically charged and discharged. Multiple drivers have to be placed in the tree as a single driver at the root would not be able to charge the overall accumulated capacity with the wanted speed. Furthermore, the clock signal degenerates with increasing distance to the root due to the wire resistance and crosstalk from nearby wires. Adding drivers, however, comes at a price, as they introduce additional offset and variance to the signal propagation delay. This can lead to spatially close leaves, of different branches, in the H-tree to suffer from a high difference in propagation delay, thus further complicating the abstraction of the global clock.

Additionally, a clocking system based on a H-tree, or a similar construct, is not fault-tolerant: A defect of a wire, a driver, or the oscillator will stop (at least a part of) the circuit from receiving a clock signal and thus from working. A mitigation against wire defects are cross-connects between branches of the tree, which also help in reducing the effects of the variance of the signal propagation delays [RHM06]. Nonetheless, cross-connects only help with benign faults: If one of the drivers placed in the tree drives a constant value (stuck-at fault), a cross-connect at a deeper level in the tree may not be able to provide a clock signal to the lower levels.

Tolerating faults becomes increasingly relevant. For a long time, this was only of relevance for safety-critical domains, where even rare failures cannot be ignored, and in aerospace domains, which must cope with the substantially higher radiation levels in the upper regions of the atmosphere and in outer space. The issue with radiation is that ionizing particles can, e.g., cause a closed transistor to involuntary open effectively causing a *single-event transient* (SET) current/voltage pulse (that can cause a memory element to flip its state, which is called a *single-event upset* (SEU)). The reduction of the critical charge required for storing information in a chip, e.g., due to the reduction of the threshold voltage at which a transistor switches, increases the susceptibility of the chip to such radiation effects. Nowadays, even the low radiation levels at sea level induce failures with a substantial failure rate. Compared to permanent hardware defects, however, the effects of ionizing radiation are primarily transient (inducing spurious SET or SEU), see [VPS+13] for more information.

The reduction of the critical charge, however, is not only relevant for the susceptibility to radiation: *Electromagnetic interference* (EMI), and specifically crosstalk, are effects where a signal transmitted on one wire is affected by a signal on another wire. This can range from degradation of the signal's shape, thus increasing the signal's delay, to spuriously inducing a SET.

Furthermore, manufacturing defects, due to impurities in the semiconductor material, have an increasing impact on the chip due to the shrinking size of the structures, called the feature size. Those defects, as well as other process-related effects, may have a transient influence on the chip, e.g., cause a disproportionate increase in signal propagation delay with rising temperature, however, they can also lead to an increase in aging effects and thus permanent defects of the chip. Note that the reduction of the feature size also increases the risk of electromigration and thus permanent wire breakages [GDM+08].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

In recent years, efforts have been made to replace the single oscillator with a group of components that together generated a synchronized clock signal. The goal is to build a clock generation system that can tolerate permanent and transient faults of its components. This not only removes the single point of failure formed by the oscillators, but also increases the *mean time between failure* (MTBF) of the system, which is especially relevant in space systems where repairing a defective component is not an option.

Distributed clock generation comes at a price, though: The resilience of the systems is usually limited to specific numbers and types of faults and, additionally, carry a significant overhead in terms of area required on the chip. This overhead typically increases quadratically with the number of involved components. Hence, the idea of combining such a fault-tolerant distributed clock generation method consisting of a small number of components with a fault-tolerant clock distribution approach emerged. Unlike the H-tree, however, the distribution of such a clock must preserve the improved fault-tolerance properties of the generated clock.

Unfortunately, with a focus on space applications, typical fault assumptions, such as a limit on the number of concurrent faults, may not apply. Indeed, a spaceborne system must be able to handle an overwhelming number of transient component failures, e.g., after a radiation burst like a solar flare. Handling such an event without affecting the operation of the system during said event is usually impossible. However, recovering from such an incident, by eventually resuming correct operation afterwards, is. A system having this property is called *self-stabilizing* [Dij74; Dol00].

Thus, the research question driving this thesis is twofold:

Q1: Can we build a clock distribution that is fault-tolerant, self-stabilizing and implementable in practice?

Q2: Can we utilize the properties of the distributed clock signal to build a communication primitive that is comparable in performance to what is achievable with a global clock?

Our results will allows us to answer both questions with *Yes*.

1.1 Organization

Chapter 2 provides an overview of the challenges and the relevance of providing a clock signal to the components in a chip. In doing so, related work in this field is covered and its relevance to this thesis is worked out.

In Chapter 3 we present a clock distribution approach called HEX, which is both fault-tolerant and self-stabilizing, and works in hexagonal grids. The results of Chapter 3 build the basis for the entire thesis and were, in part, previously published in

- D. Dolev, M. Függer, C. Lenzen, M. Perner, and U. Schmid, „HEX: Scaling Honeycombs is Easier than Scaling Clock Trees“, *Journal of Computer and System Sciences*, vol. 82, no. 5, pp. 929–956, 2016. DOI: `10.1016/j.jcss.2016.03.001`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

- D. Dolev, M. Függer, C. Lenzen, M. Perner, and U. Schmid, „HEX: Scaling Honeycombs is Easier than Scaling Clock Trees“, in *25th ACM Symposium on Parallelism in Algorithms and Architectures*, 2013, pp. 164–175,
- M. Perner, M. Sigl, U. Schmid, and C. Lenzen, „Byzantine Self-Stabilizing Clock Distribution with HEX: Implementation, Simulation, Clock Multiplication“, in *6th Conference on Dependability*, Barcelona, Spain, Aug. 2013, pp. 6–15, ISBN: 978-1-61208-301-8,
- M. Perner, „Self-Stabilizing Byzantine Fault-Tolerant Clock Distribution in Grids“, Master's thesis, Technische Universität Wien, Institut für Technische Informatik, Treitlstr. 3/2/182-2, 1040 Vienna, Austria, 2013,

As usual for collaborative work, it is difficult to identify the individual contributions of the co-authors. It is fair to say, though, that I have been responsible for the planning and setup of the entire experimental evaluation, including the design and implementation of the HEX nodes, and the complete simulation environment. Moreover, I contributed substantially to the identification of the worst-case scenario. Furthermore, I designed the fast-clock mechanism.

Chapter 4 seeks to improve the idea of HEX by exploring alternative constructions of a clock distribution grid. This is motivated by the fact that, while achieving the original goal, there are certain scenarios in which HEX's performance degrades towards its worst-case. Finding an alternative with a smaller range between best-case and worst-case is hence an important goal. While the discussion of possible candidates was for the most part a joint effort, the simulations and evaluations, hence the majority of this chapter, were solely my work.

Using the results of Chapter 4, Chapter 5 presents another clock distribution, TRIX, which is similar to HEX. TRIX delivers far better results in the presence of faults, albeit under the assumption that the provided clock signal is well synchronized. Furthermore, we present a transistor cell implementation of TRIX. The results of Chapter 5 were documented in

- A. Kinali, C. Lenzen, and M. Perner, „Fault-tolerant High-Performance Clock Distribution“, E191 - Institut für Computer Engineering; Technische Universität Wien, Tech. Rep. TUW-278925, 2019. [Online]. Available: `https://publik.tuwien.ac.at/files/publik_278925.pdf`.

As this was the result of joint work with Attila Kinali and Christoph Lenzen, a clear identification of the co-authors individual contributions is difficult. However, the credit for the development and simulation of the transistor-cell model are awarded solely to Attila Kinali. My contribution in this work are primarily the selection of the topology and the basic algorithm as a result of Chapter 4. Furthermore, I conducted the grid-level simulations of TRIX, utilizing the developed transistor-cell model, including the accompanying evaluations.

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

In Chapter 6, we focus on Q2. Specifically, we address communication between components driven by a clock distributed via a system similar to HEX/TRIX. Due to the typically substantial clock skews provided by a fault-tolerant clock distribution system, it is non-trivial to know when the data sent from a neighbor is stable so that it can be safely sampled. In this chapter, we present a self-stabilizing communication primitive designed to work with such clocking systems. The results of this chapter were previously published in

- M. Perner and U. Schmid, „Self-Stabilizing High-Speed Communication in Multi-Synchronous GALS Architectures“, in *24th IEEE International Symposium on On-Line Testing and Robust System Design*, Platja d'Aro, Spain, 2018,
- M. Perner and U. Schmid, „Self-Stabilizing High-Speed Communication in Multi-Synchronous GALS Architectures“, Technische Universität Wien, Tech. Rep. TUW-268547, 2018. [Online]. Available: `http://publik.tuwien.ac.at/files/publik_268547.pdf`,
- M. Perner and U. Schmid, *Self-Stabilizing and Metastability-free Communication in Multi-Synchronous Architectures*, (submitted to MICPRO), 2019.

These results are entirely my own work.

In Chapter 7, we summarize the results and elaborate on further research directions that have arisen during the work done in this thesis.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

CHAPTER 2

CLOCKING A SYSTEM

> Time, space, and causality are only metaphors of knowledge, with which we explain things to ourselves.
>
> — FRIEDRICH NIETZSCHE

COORDINATION IS REQUIRED for a system of independent components to produce meaningful and coherent results. Time is not necessary for coordination, however, ordering of events is. Yet, time is a convenient way for ordering events. In this chapter, we will discuss how a system can be provided with information about time. In this course of action, we will consider three levels of abstraction, with varying detail, during this journey through time.

We rely on abstraction to ignore unnecessary details while focusing on the essential parts, thus allowing a simplified view of a system.

(1) Analog signals and circuits: Our first abstraction is that of analog electronics based on *transistors*. This allows us to consider signals, interpreted as voltage levels, and their behavior over time, instead of electrons and fields.

(2) Digital signals and circuits: While transistor circuits are inherently analog, this continuous value domain is not required for digital logic. Hence, we interpret the voltage of an analog signal as digital “0” and “1”, depending on whether it is below or above a certain threshold, see Figure 2.1. Furthermore, we consider a collection of multiple transistors building a specific function as a gate, through the abstraction of a digital *combinatorial circuit*.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

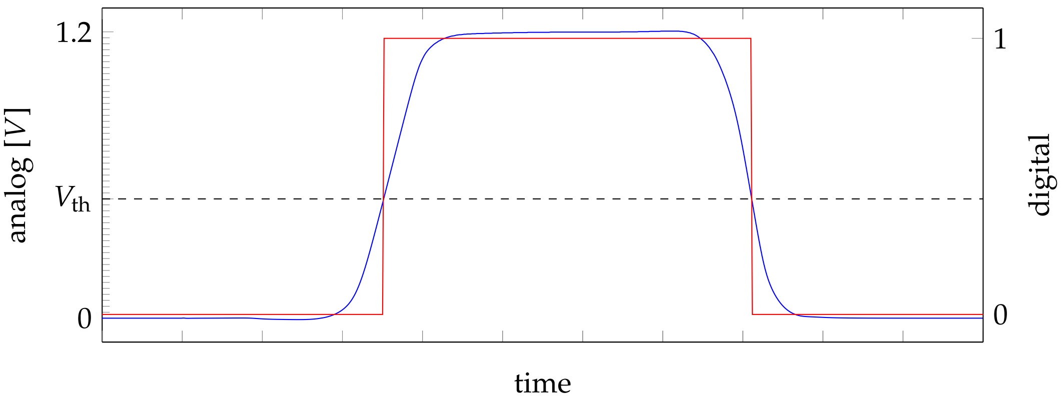


Figure 2.1: This figure visualizes the mapping of an analog signal (—) to its digital abstraction (—). Here, the threshold voltage V_{th} at which an analog signal is considered to change its digital representation is set to 0.5 V. Note that there might be a different threshold for the transition from digital "0" to "1" and from "1" to "0".

(3) Distributed Systems: A system of multiple interacting combinatorial circuits can be seen as a *distributed system*. We abstract from its implementation, i.e., its internal working, by restricting our attention to its externally visible behavior.

2.1 The Role of Time

Time is an integral part of any system description. System implementations may differ in the way time is accessible to its components, however. On the one hand, components may be completely oblivious to the concept of time, and rely on the order of events (causality) only. On the other hand, components may act on certain times only.

2.1.1 Time in a Analog Transistor Circuit

Electronic circuits can be described by means of partial differential equations, namely Maxwell's equations [Max65], which immediately connect the behavior of such systems to time.

The behavior of such a circuit, abstracted as the current and voltage of its signals, can be described by means of ordinary differential equations. However, in most systems relevant to computer science, the digital abstraction shown in Figure 2.1 of such a circuit is sufficient. Yet, the threshold voltages V_{th} may not be uniform for every transistor but may vary by design or due to *process-voltage-temperature* (PVT) variations. Nonetheless, safe thresholds for "0" and "1" can usually be defined, and the gap between these threshold is deemed a forbidden zone. In general, this zone is irrelevant for a digital circuit: As it is assumed that the digital signal has an infinite slope, the signal will not stay inside the zone.

TU Bibliothek WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Instantaneous changes and classical physics, however, do not concur in most cases, and circuits are no exception: Even though we can have instantaneous changes of the logical values, the underlying system still needs time to perform these changes. For the abstraction of a digital circuit to hold, certain conditions with regards to timing must be satisfied.

The circuits considered in this thesis are built as VLSI ICs with *metal-oxide-semiconductor* (MOS) transistors. Especially for *complementary metal-oxide-semiconductor* (CMOS) transistor circuits, a fast transition through the forbidden zone is paramount. Otherwise, both stacks, the one connected to VDD, the P-stack, and the one to GND, the N-stack, of the CMOS circuit may conduct, causing a short circuit.

The MOS technology uses a photolithographic process to imprint structures on a silicon-based substrate. The so called *feature size* relates to the size of these structures,[1] e.g., 22 nm [Iwa09]. The reduction of the feature size is a primary driver for advances in photolithographic process technology. Reducing the feature size allows to put more transistors on the same area, hence reducing manufacturing cost. However, increasing the number of transistors per area also increases the heat that needs to be dissipated. Reducing the supply voltage of the IC, for increasing the switching speed and for decreasing the power required to switch a transistor [Iwa09] is hence inevitable. In turn, however, the shrinking also moves wires closer together. This results in increased crosstalk and mutual capacitances [Vee17]. Recall that crosstalk is a phenomenon where a signal on one wire negatively affects the signal on a neighboring wire. Mutual capacitances lead to an increased wire delay, resulting in the dominance of the wire delays over the switching delays for smaller feature sizes.

2.1.2 Time in Digital Circuits

Digital circuits are based on the digital abstraction and primarily consist of combinatorial circuits called *gates*, which perform Boolean logic operations on the inputs and set their output accordingly. Note that the correspondence between switching circuits and Boolean logic has been established long ago by Shannon [Sha38]. Furthermore, they contain state-holding elements, implemented via feedback loops, whose new output also depends on their old output.

We will demonstrate this abstraction by means of one of the most important components of a digital circuit: The *flip-flop*, a state-holding gate.

A flip-flop consists of two *latches*. They work hand-in-hand to implement the abstraction of a single-bit memory. The inner working is shown in Figure 2.2. As the storage elements within the latches are constructed via inverter loops, transmission gates are required to reduce the load required to force the loop to hold a specific value (“0” or “1”). A transmission gate can be seen as a switch, i.e., it has two operation modes: transparent and high-impedance. When in the latter, the gate disconnects its two ports from each other such that they incur no influence on either part of the circuit.

---

[1] It does not, however, give the size of the smallest structure on a chip.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

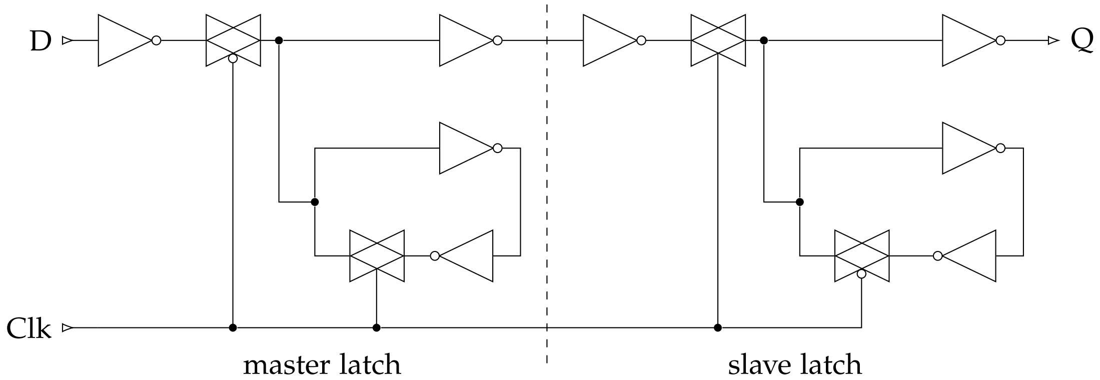


Figure 2.2: A schematic for a Flip-Flop. It comprises of two latches, the master and the slave latch. The master latch samples the input *D* during the low phase of the clock signal *Clk*. On the transition of the *Clk* signal, the transmission gate at the input switches from a transparent mode into a high-impedance mode, allowing the inverter loop to drive its internal value: the state of the input signal on transition of the clock. The slave latch is inverse with regard to the phase of the transmission gates. During the high phase of the clock, the latch transparently transmits the value stored in the master latch to the output *Q*. On the falling clock edge, the inverter loop in the slave latch continues driving this value on the output. This figure has been adapted from [JYG09].

Consider, for example, that the clock signal *Clk* is low. This allows the input *D* of the master latch to overthrow the current value of the storage loop, without the need to overwrite the last inverter in the storage loop. Hence, the loop in the master latch captures the input. We call this the transparent mode of the latch.

During the transition of the *Clk* signal, the transmission gate after *D* disconnects and the one in the storage loop connects, leading to a stable state. This is referred to as the opaque mode. As the slave latch is inversely clocked, it will switch into transparent mode after the *Clk* transition. Thus, it captures the value stored in the storage loop of the master latch and forwarding it to the output *Q*. On the next clock transition, the slave latch will switch into opaque mode, disconnecting itself from the master, and thus continue to provide a stable value to the output, while the master latch re-samples *D* due to it being in the transparent mode.

In principle, this design seems perfect to store the current signal value applied at *D* just before the transition of *Clk*. However, with delays involved, the stability of the input signal is relevant near a transition of *Clk*: It takes some time for the input signal to fully propagate into the master latch and cause a stable value in the storage loop, which is called *setup time*. Additionally, the transmission gate needs some time after the transition of the clock signal to completely switch off. During this time, which is called *hold time*, *D* must also be stable so as not to corrupt the fragile in-between state of the storage loop. If *D* is not stable during these times, which is called a

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

setup/hold violation, the digital abstraction may crumble, as is described in more detail in Section 2.2.

In the early days, digital circuits were built by separate combinatorial gates wired together. This worked, although high effort was required to weed out hazards and race conditions due to the unclocked and asynchronous behavior of these systems [FH90; Huf57]. A motivation for a clocked system was given by Turing in [Tur86]:

> We might say that the clock enables us to introduce a discreteness into time, so that time for some purposes can be regarded as a succession of instants instead of a continuous flow. A digital machine must essentially deal with discrete objects, and in the case of the ACE[2] this is made possible by the use of a clock. All other digital computing machines except for human and other brains that I know of do the same. One can think up ways of avoiding it, but they are very awkward.

To work around this unavoidable volatile behavior of combinatorial circuits, state holding elements are introduced. At a certain point in time, the state holding element considers the current inputs and its current state, and decides on the next state. The determination of this point in time allows to differentiate two principal circuit design methodologies: The synchronous and the asynchronous design model.

2.1.2.1 Synchronous Design Model

In the synchronous model, a global clock signal is responsible for the progress of the system. That is, only the global clock triggers the flip-flops in the circuit, and all combinatorial logic is in the data path [FH90]. Hence, the influence of the clock is limited to the state-holding elements.

This model requires two conditions for its abstraction to work properly: (i) all clocked gates receive the clock signal at the same time, and (ii) the computation of the circuit between the flip-flops is completed before the next clock signal arrives. While (i) can be achieved in a small circuit easily, for larger clock distribution networks, this is a challenging task, as will be shown in Section 2.5. Condition (ii), on the other hand, requires a static timing analysis [BC09]. This analysis must find the slowest path between a flip-flop's output and its successors input in the circuit. Adding to the path delay the output delay of the first flip-flop and the setup-time of the second gives an upper bound on the frequency achievable.

The assumption that all signals in a digital circuit are binary and time is discrete is an abstraction that requires guarantees to be upheld to be applicable [Hau95]. PVT variations can cause a violation of calculated delays, however. Hence, it is possible that the delay of a combinatorial circuit intermittently changes and induces a setup/hold violation [Con02].

The synchronous model is currently the predominate one, hence almost all *electronic design automation* (EDA) tools focus on this model. Tool support starts at the design

---

[2] Automatic Computing Engine

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

phase, where synchronous systems can easily be, and primarily are, described by *finite state machines* (FSMs). A FSM can be modeled by a finite graph where the vertices are the states of the circuit. The edges of the graph are called transitions, and they conditionally are activated by guards consisting of a subset of the input signals and state-related variables. It is a simple task to generated a *hardware description language* (HDL) implementation of a FSM. From the system's HDL implementation, the tools can synthesize a circuit that is ready for manufacturing. Without such tool support, the mapping of the HDL implementation to the gates, as well as performing the placement and routing on the IC, is a challenging task. Not to mention the timing analysis required to determine the maximal achievable operating frequency of the circuit.

2.1.2.2 Asynchronous Design Model

Contrary to the synchronous design model, the asynchronous design model does not employ a global clock signal. This results in a multitude of different design methods, compared to the rather canonical synchronous design model [Hau95]. Depending on the actual methodology, either some sort of completion detection is used or handshakes are required between state holding elements. In any case, asynchronous circuits always run as fast as they can, according to their current operating conditions [Ebe91]. On the other hand, a synchronous circuit is designed for, and runs with, its worst-case timing behavior.

Actually, asynchronous circuits (also called self-timed circuits [Sei80]) are a very natural concept, as awaiting an event is more common than acting at a specific time. Asynchronous computation models are hence common in biological systems: For example, Cardelli, Kwiatkowska, and Whitby describe in [CKW18] how asynchronous circuits can be modeled with chemical reaction networks.

While the asynchronous model does not rely on a global clock, time is still relevant. Due to the finite speed of the transistors, and the signal delay through the connecting wires, circuits that need to relate information propagated via different paths require delay tuning to work as intended.

We can differentiate between three basic restrictions here:

speed-insensitive A speed-insensitive circuit is insensitive to the delay of its gates, but not to the delay of its wires [Mil61].

delay-insensitive The behavior of the circuit is independent of the delay of both the gates and the wires [Cla67; Ebe91].

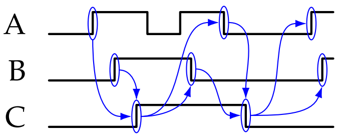
quasi-delay-insensitive Is an extension of delay-insensitivity, where it is assumed that, for specific *isochronic forks*, delay bounds apply [Mar90; MM95]. A simplistic version would be the assumption that the delay of the wires after a fork are identical.

These restrictions also have an effect on the principal design of the completion detection logic of an asynchronous circuit. For example, the *Null Convention Logic* (NCL)

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

A	B	C
0	0	0
0	1	C_{-1}
1	0	C_{-1}
1	1	1

(a) Truth table of a Muller C-Gate. The inputs A and B define the output C. C_{-1} denotes the previous value of the output C.



(b) Timing diagram of a Muller C-Gate. As can be seen in the truth table on the left, the transitions of the output signal C are synchronized with the transitions of both inputs A and B. The blue markings show the dependence of the signal transitions on each other.

Figure 2.3: Truth table and timing diagram illustrating the behavior of a Muller C-Gate with the inputs A and B, and the output C. Copied from [Per13].

provides a delay-insensitive way to build a circuit [FB96]: The 2-valued Boolean logic is extended by an intermediate value (NULL), which has to be reached between two data values. Thereby, a clear differentiation between an old and a new value can be made.

2.1.2.3 Signal Transition Graphs

In an asynchronous computation, progress does not a priori depend on the passage of time, but on the transition of signals. In fact, for synchronous computations, the externally relevant event for the progress of the system is a transition of the clock signal.

To capture this, a *signal transition graph* (STG) can be used to model the behaviour of a system. A STG is a Petri Net, where the transitions are labeled with signal names and a transition polarity (+ for rising, − for falling) [Hau95]. Petri Nets are a graphical formalism to model concurrent systems [CKK+97]. They consists of places and transitions, which are connected by arcs. We say that a STG fires when a token passes over a transition. This requires that the transition is enabled, i.e., the input signal corresponding to the transition performs the labeled change in polarity resp. the output signal is able to make this change.

Assuming that the firing of a transition in the STG causes the corresponding signal to make the specified transition, it can be used to model asynchronous circuits. By enforcing some restrictions on the STGs, it is even possible to automatically generate a circuit from a STG [CKK+97]. Thus, like state machines for the synchronous model, STGs are the primary tool for describing asynchronous systems.

We show an example STG for a Muller C-gate [MB56; Mul55], which is basically an AND-gate for signal transitions. That is, if the last transition at both inputs is a rising one, the output performs a rising transition, as shown in Figure 2.3b, and vice

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

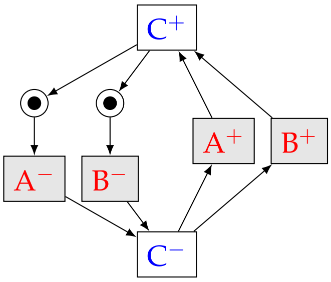


Figure 2.4: An example for a STG, which could be synthesized to a Muller C-Gate, with the inputs A and B and the output C. The nodes with *signal*$^+$ mark a rising transition of the signal, whereas *signal*$^-$ marks a falling edge of the respective signal. This is a correct STG, but does not describe the general behavior of a Muller C-Gate, as there are restrictions on the transitions of the input signals as well. Copied from [Per13].

versa for falling transitions. The truth table of the Muller C-gate is given in Figure 2.3a, with the corresponding STG in Figure 2.4. It is immediately apparent from Figure 2.3a that the Muller C-gate is a state-holding element [SEE96]. Note that the STG shown in Figure 2.4 represents the ideal flow of transitions; it does not also capture possible intermediate transitions of A and B.

2.1.3 Distributed Algorithms

In distributed algorithms, one considers a system of distributed processes that compute concurrently and communicate via message-passing with each other. The definition of what constitutes as a distributed system can be given as follows [Gho14]:

Multiple processes. The system consists of more than one process that can compute independently of each other. They can be separate physical processors, different processes on the same processor, or a mixture of both.

Interprocess communication. Processes communicate in this system via messages transmitted from a sender to one or more receivers.

Disjoint states. The only way for the process to interact is via messages. This requires that, even for processes on the same processor, there is no shared memory.

Collective goal. The processes must interact in order to achieve the common goal.

With this general description of a distributed system, there is enough room for further restrictions. This has resulted in a multitude of models of computations, each of which with a different set of restrictions. However, here we focus on the two most extreme models, as they cover the whole spectrum of system discussed in the following chapters. We will use one of the basic problems in distributed computing to outline the differences between the two general models.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Furthermore, distributed system are typically not considered to be perfect. That is, it is usually assumed that some part of the system can stop working towards the collective goal. These parts are called faults. Their effect on the system can vary between benign faults that prohibit working towards the common goal, to malicious faults that actively try to prevent the system from achieve the common goal.

2.1.3.1 Fault-Tolerant Consensus

Consensus describes the ability of a system to reach a common decision based on local inputs. That is, some voting is commenced upon the specific input values of all processes which, e.g., could be measurements of a temperature sensor. We borrow the definition of the consensus problem from [AW04]:

> Consider a system in which each processor p_i has special state components x_i, the *input*, and y_i, the *output*, also called the *decision*. Initially, x_i holds a value from some well-ordered set of possible inputs and y_i is undefined. Any assignment to y_i is irreversible. A solution to the *consensus* problem must guarantee the following:
>
> **Termination** In every admissible execution, y_i, is eventually assigned a value, for every non-faulty processor p_i.
>
> **Agreement** In every execution, if y_i and y_j are assigned, then $y_i = y_j$, for all non-faulty processors p_i and p_j. That is, non-faulty processors do not decide on conflicting values.
>
> **Validity** In every execution, if, for some value v, $x_i = v$ for all processors p_i, and if y_i is assigned for some non-faulty processor p_i, then $y_i = v$. That is, if all the processors have the same input, then any value decided upon must be that common input.

Note that validity is required to exclude trivial solutions, e.g., all processes always decide on the value 42 in every execution. Furthermore, by adapting the conditions, closely related problem can be described. For example, by weakening agreement and validity to allow values within the interval spanned by the initial values, approximate agreement can be defined.

Applications for consensus are widespread, as, e.g., clock synchronization (cf. Section 2.4.3) can also be seen as a consensus problem.

2.1.3.2 Synchronous Model

The synchronous round model assumes a lock-step computation. The progress of the system happens in so called rounds, which consist of 3 phases each:

1. All processes send messages to their peers (which actually happens at the end of the previous round).

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

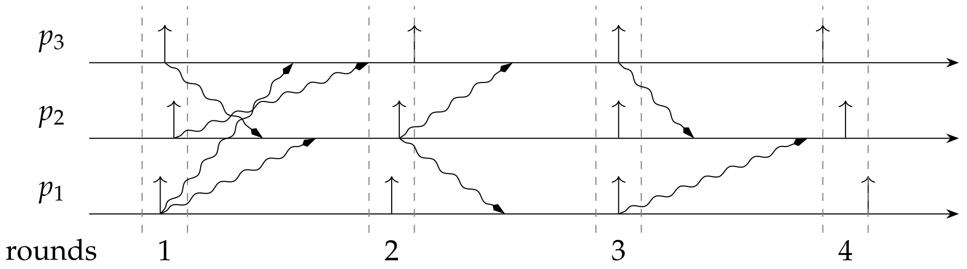


Figure 2.5: An example for a synchronous lock-step execution. The dashed, gray, lines mark the time slots in which a process has to make its computation step of the specific round, to fulfill the lock-step synchronous abstraction. The wiggly lines represent the communication between the processes.

2. All processes wait for a defined amount of time, while they receive messages sent in this round.

3. Every processes performs its local computation in zero time (and sends messages to its peers for the next round).

Phase 2 brings time into the system. It is an abstraction to cover the fact that the delivery of messages (as well as the computation) is not instantaneous.

It should be noted that perfect synchronization of the rounds is not required. As can be seen in Figure 2.5, as long as the processes agree on the common time within a given clock skew, the lock-step model can be simulated.

Algorithm 2.1: Consensus Algorithm in the presence of crash failures [AW04]: code for process $p_i, 0 \leq i \leq n-1$

```
Initially $V = \{x_i\}$ ;                      /* $V$ contains $p_i$'s input.  */
round $k$, $1 \leq k \leq f+1$
    send $\{v \in V : p_i$ has not already sent $v\}$ to all processes;
    receive $S_j$ from $p_j, 0 \leq j \leq n-1, j \neq i$;
    $V := V \cup \bigcup_{j=0}^{n-1} S_j$;
    if $k = f+1$ then
        $y := \min(V)$ ;                                /* decide */
```

Solving consensus can be trivial in this model. Assume that at most f processes can crash, with the effect that the crashing process will send its message only to a subset of the other processes and will be silent forever after that. Under such a failure model, we can achieve consensus within $f+1$ rounds by the protocol, shown in Algorithm 2.1: In every round, every process sends its input values, which were not sent already, to all of its neighbors. Every process will hence receive these messages in the same round from every correct process. If a process p_j has crashed during this round, only a subset

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

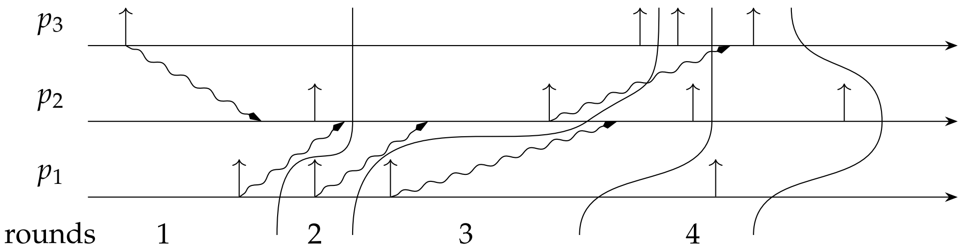


Figure 2.6: An example for an execution in the asynchronous model. Note that, compared to the synchronous model, there is no clean separation between the round. This can be seen as the event sent by p_1's first and second computation, is received by p_2 before its second computation.

of the processes may have heard from p_j and hence know x_j. Hence, another round is required for all other processes to become aware of this input value. The number of rounds where a process can crashed is limited by the maximum number of faults f. Therefore, after $f+1$ rounds, every process has the same set V: An unanimous decision can be made.

2.1.3.3 Asynchronous Model

The asynchronous model is the most general applicable timing model, as it has no restrictions on the progress that an individual process has to make. This freedom of choice is, in turn, too weak to allow solving a vast number of distributed computing problems. For example, it is impossible to decide if a process is faulty or "just" late. This has a major impact, if the decision of the system has to potentially integrate knowledge of every alive process. In particular, it is impossible to solve consensus in the asynchronous model if just a single process can crash [FLP85]. In [DDS87], the authors precisely characterized what must be added to the asynchronous model in order to make consensus solvable.

An example of an asynchronous computation is given in Figure 2.6. Compared to the synchronous case, it is not necessarily possible anymore to cleanly separate computation and corresponding communication in distinct rounds. For example, the message send by p_2 in round 2 is received by p_3 in round 4, while p_1's round 2 message is already available to p_2 in round 2. However, we can see that every event spans a causal cone, which comprises those actions of the other processes that might be causally affected by the event [Mat89].

2.1.3.4 Causality

The reader might have noticed that we did not consider time per se when considering distributed algorithms so far. This is true in some sense, as timestamps are rarely

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

required. What distributed algorithms actually need, however, is causality and hence some ordering of events. Clearly, order does not a priori need time.[3]

Lamport's seminal work [Lam78] on the happened-before relation captures causality in distributed systems in a formal way. Two events are in a happened-before relation, if one of the following holds:

1. They are events of the same process
2. They are corresponding send/receive events
3. There is an event with which both events are in a happened-before relation.

For processes to observe this happened-before relation, they can be equipped with an internal counter that is increased after every step the process makes. The value of this counter is attached to every message sent by the process. Additionally, if a message is received, the next value for the internal counter has to be larger than the current internal counter value and the counter value received. These *logical clocks* allows to order events according to the happened-before relation: When for two events the happened-before relation holds, then the logical clock values show the temporal order of them. The converse, however, is not true: As every process increases its counter after every step, the counter values of two processes are independently increased if they do not communicate with each other. This is a consequence of the happened-before relation being a partial order, while the counter values form a total order [AW04].

This restriction, however, can be lifted by using *vector clocks*, as introduced by [Fid91; Mat89]. Instead of a single counter value, every process keeps a vector with size n (= number of processes in the system). The process i increments after every step the value at the index i in the vector. The vector is attached to every transmitted message. On reception of a message, every index of the vector is set to the maximum between the received datum and the local value. With these vector clocks, it is possible to order events. Furthermore, this allows to determine if two events happened concurrently, namely, when the two vector clocks are incomparable.

2.1.3.5 Communication

For communication in distributed systems, we only consider message passing between processes. A message is sent from one party to the other via point-to-point links. While in transit, the message, depending on the fault model, may be altered or lost. Furthermore, some models may not enforce message ordering, i.e., the order in which the messages are received may not reflect the order in which they were sent.

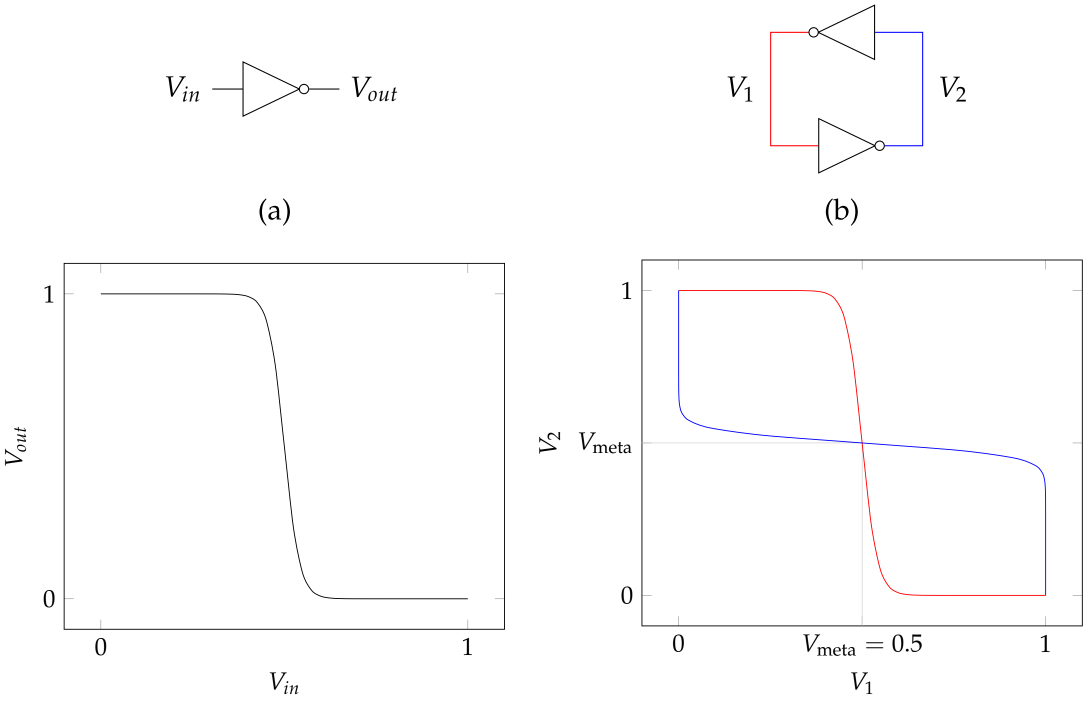
2.2 Metastability

As electrical circuits are per-se analog, i.e., work with continuous voltage levels, voltages close to the threshold voltage V_{th} are interpreted differently: $V_{\text{th}} - 1\,\mu\text{V}$ leads to "0",

---

[3] Of course, time provides order.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) (b)

(c) (d)

Figure 2.7: (c) shows the transfer function of a single inverter (a). The inverter loop (b) is similar to the one used in the flip-flop in Section 2.1.2. Superimposing the transition function of the two inverters, as both share input/output, results in the transition function of the inverter loop (d). Besides the stable equilibrium points at $(0, 1)$ and $(1, 0)$ we see there a third semi-stable or meta-stable equilibrium point at $(0.5, 0.5)$.

$V_{\text{th}} + 1\,\mu\text{V}$ to "1".[4] Consequently, minor disturbances like signal noise or PVT variations of V_{th} have a major effect on the digital abstraction if the analog signal remains close to V_{th} for a non-negligible time.

Of course, the question arises whether such a dangerous near-threshold signal can happen. Unfortunately, the answer is yes: a metastable upset can cause such a signal. In this subsection, we will briefly explain what happens in such a case.

Consider the boundary between two systems, each with its own notion of time, over which a digital signal must be transmitted. Assuming that the receiving system is synchronously clocked, a natural idea is to use a flip-flop as described in Section 2.1.2 for sampling the transmitted signal. Clearly, the sampling points of the received signal are determined by the receiver's clock, whereas the transitions of the transmitted signal are dictated by the sender. As there is no relation between the transitions of the signal

---

[4] Note that $1\,\mu\text{V}$ is an arbitrarily chosen value here.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Wien Bibliothek Your knowledge hub

and the transitions of the receiver's clock, it can happen that there is a signal transition arbitrarily close to a clock transition, which constitutes a setup/hold violation.

To answer the question as to what happens in this case, consider the following simple example: As in [Kin08], we consider an inverter loop, as shown in Figure 2.7, similar to the one used by the flip-flop of Section 2.1.2. The two inverters normally output the inverse of their input. However, Figure 2.7d reveals another stable state: if one of the inputs is forced to the intermediate voltage level V_{meta} the circuit remains in this semi-stable or metastable equilibrium point. In reality, at some point, noise and other effects will cause the circuit to leave this metastable state and end up in a stable one. Note carefully, however, that the time when this will happen cannot be bounded in general.

While this example may look contrived, to showcase this effect, we need to stress that (i) this type of storage loop is a common design element, and (ii) that every bistable circuit also has a metastable state [Ebe91; Mar81].

The effect arising from this metastable state is termed metastability, the event leading to metastability is called a metastable upset. The first mentioning of metastability in computer science dates back to [Cat66; LC66]. Albeit, it has been known under the term *Buridan's ass* [Lam12] already by the French philosopher Jean Buridan in the 14th century. It states that

> An ass placed equidistant between two bales of hay must starve to death because it has no reason to choose one bale over the other.

The main issue with metastability is that it cannot be avoided on time domain crossings [KC87; Mar81]. Furthermore, metastability cannot be contained, i.e., can spread arbitrarily far into a circuit, leading to a system failure.

The reader may be aware that, even though our daily computer systems crash from time to time, they do not suffer significantly from metastability: While metastability cannot be avoided, the probability of metastable upsets can be made very small. A simple way to achieve this is by serially chaining multiple flip-flops, so called stages, for building a *synchronizer*. With each stage of the synchronizer, the probability that metastability is resolved, in a previous stage, increases significantly [CW75; KW76]. Synchronizers are hence commonly used to prevent a metastable signal from entering a circuit, and there is a huge literature on synchronizer design [DB99; KBY02; Kin08; PM95]. However, synchronizers cannot completely remove metastability from a system [Gin03]. Moreover, the increase in safety has to be paid by an increase in delay, though, as every stage requires another clock cycle for the signal to pass.

Even though metastability cannot be avoided and contained, recent research has shown that it is possible to prevent it from spreading arbitrarily: In [FFL18], Friedrichs, Függer, and Lenzen show that some restricted logical masking is possible for certain classes of circuits. However, this does not contradict the impossibility of resolving metastability, since only a "safe" output of the logical gates is guaranteed, not a correct one.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

2.3 Failure Models

In his seminal work [Von56], Von Neumann introduced the concept of reliability of logic gates. Failures in circuits were introduced by analogy with biological systems, neurons to be precise. Via probabilistic analysis, it was shown that redundancy can alleviate failures in the system, provided enough overhead is added.

Today, the need for reliable systems increases due to autonomous driving and other instances of cyber-physical systems where correct operation is crucial. As failures can have catastrophic consequences here, the system must be designed to prevent failures in the first place, and transparently handle the ones that cannot be prevented.

2.3.1 Definitions

Following [Kop97], a *fault* is the primitive event that can cause a component to deviate from its designed behavior at the current time. An *error* is the manifestation of a fault in the state of a component. A *failure* is a deviation from the intended behavior of a component. Due to this, the failure of a component can cause a fault of another component.

2.3.2 Taxonomy of Failures

We consider a classification of failures according to their duration and severity [ALRL04; Kop97], which are sufficient for our overview. For a refined taxonomy of errors, failures and faults, we refer to [ALRL04].

A failure's duration, or persistence, can be classified into the following three categories:

A *transient failure* is caused by an external event, typically originating in the environment. This type of failure is hard to control, albeit, supply voltage [SS04] and feature size can have an impact [PMAS04] in some settings. For example, a smaller feature size increases the probability of SEUs.

Intermittent failures can be caused by specific data patterns causing, e.g., crosstalk. Furthermore, temporary changes of the environment (PVT variations) can cause parts of the circuit to be slightly out-of-specification.

The difference to a transient failures is that this type of failure depends on the occurrence of specific operating conditions, possibly for a longer time. By contrast, a transient failure is caused by a single event.

A *permanent failure* has manifested itself in a circuit and cannot be resolved by the system on its own, i.e., external repair is required. Causes for permanent failures can be the aging of the semiconductor material, e.g., due to electromigration, or due to cumulated dose defects caused by long-term, respectively serve, radiation [Wat97].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

The severity of a failure, i.e., the behavior of a faulty component towards the other components of the system, can be classified into three basic failure-types [AW04]:

The *Fail-Silent or Crash Failure* is the simplest behavior a faulty component can show. In a distributed system with message passing, for example, the name speaks for itself: From a certain time on, messages are no longer sent. At the digital logic level, a reasonable correspondence would be the absence of a transition of the signal: stuck-at failures can be seen as crash failures.

In severity above crash failures are *omission failures*. In the context of distributed systems, one distinguishes send omissions, where (a subset of) messages may not be sent, and receive omissions, where (a subset of) messages may not be received.

A *Byzantine Failure* [LSP82] is an unrestricted type of failure. A Byzantine faulty component can behave inconsistently towards it neighbors, can create artificial messages, and can violate timing assumptions.

Note that when considering the digital abstraction, a Byzantine fault is limited to "0" and "1" outputs. However, as we have discussed earlier, a metastable signal is not bound to this abstraction and thus is a stronger failure type—in the sense that it cannot be contained and handled on a logical level [FFS09].

For example, consider a single fork that connects a faulty sender, having a out-of-specification output voltage, to two inputs with slightly different threshold voltages (V_{th}). The inputs may be able to resolve the signal in time to a valid digital signal. If the inputs resolve differently, the sender is perceived as a Byzantine fault. However, if the inputs become metastable, the power of a Byzantine fault is exceeded.

2.3.3 Fault Models

Systems are composed of components, which themselves are made up of components. For example, distributed systems are often considered to consist of $n \geq 2$ processors interconnected by point-to-point communication links or some network. Every processor consists of many components (registers, ALU, ...) that may fail in some way. A very important question for designing fault-tolerant systems is when and in which way components may fail. There are two fundamentally different approaches to this problem:

Probabilistic models: Every components fails probabilistically, typically with some given failure rate. Often, component failures are assumed to be independent of each other [KK07].

Adverserial models: Failures are under the control of an adversary, who is restricted by some failure model. For example, in a system of n processors, at most f may fail (in some way) during the systems lifetime. Many different failure models,

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

including hybrid ones [BSW11; LR93; LSP82; SCY98; SWR02; WS07], have been studied in the literature.

The increase in reliability requirements in current applications, especially aerospace, is an import factor during the system design phase. Thus, the definition of a suitable fault hypothesis is key, as a violation of the fault hypothesis will typically lead to a fatal system error. Given the fault hypothesis, a fitting fault model that, e.g., specifies how many components may fail, can be defined. However, as soon as the fault model allows sufficiently severe failures, the complexity of the fault-tolerance mechanisms required to handle them increases. Hence, the fault model should be as restrictive as possible, while still covering the fault hypothesis.

To define a suitable fault hypothesis for a system, it is necessary to be aware of the possible sources of faults in the system. A primary source of faults is the IC itself. Due to the reduction in feature size, the wires on an IC cannot be considered in isolation anymore [PMAS04]. Nearby wires can interfere with each other as, e.g., the parasitic capacitances between the wires can lead to crosstalk. In turn, this can cause, e.g., out-of-specification delays. These out-of-specification delays are excessive delays that may lead to a signal transition happening too close to a clock transition, leading to metastability [PMAS04]. However, crosstalk is not the only source for out-of-specification behavior. For example, a droop in the supply voltage of the IC [FKLW18] decreases the switching speed of the transistors, or an increased in temperature (PVT variations) can also lead to a slow-down of the circuit's operation. Fortunately, the occurrence of this kind of faults is transient in nature. That is, they affect the circuit only during a limited time span. Nonetheless, the dramatic reduction in feature sizes, and operation voltages, reduce existing "safety margins" for tolerating transient faults, making this type of faults the primary source of failures in current VLSI ICs [DW11; VPSS12].

Aging effects can also lead to an increase in wire delays and transistor switching time [McP07; MSF14], resulting in permanent timing violations, among others. Similar to transient faults, the likelihood of aging effects also increases with the reduction of the feature size. Hence, circuits created with newer semiconductor technology have an increased susceptibility to transient and permanent faults [GEB+06; SKK+02].

Another cause for both transient and permanent faults is ionizing radiation. Radiation effects occur when an ionizing particle hits the IC and causes a temporal disturbance of the charges within the IC [Bau05]. While there are multiple effects that can be caused by such *single-event effects* (SEEs), the general differentiation is between "soft" SEEs, which cause transient effects, and "hard" SEEs, leading to a permanent defect of the device. In terrestrial applications, soft SEE are the norm, while the relevance of hard SEEs is limited to harsh environment like space applications. However, the importance of SEEs in terrestrial applications rises with the continuing reduction of the feature size, and they are expected to become the major cause for failures in the future [CKT+11; WHG+09].

A SET is a SEE that affects a circuit element due to an induced voltage spike. This happens when an ionizing particle hits the junction of a switched-off transistor. The electron-hole pairs induced by this hit are separated by the electric field and thereby

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

make a current pulse at the output. However, the hit transistor may not need to be completely switched-on: If the induced pulse is sufficient to charge a successor transistor above its threshold voltage, the SEE causes a SET. Note that if the hit transistor has multiple successors with differing threshold voltages, the SEE may not cause a SET for every successor. If a SET is latched by some subsequent storage element, it may lead to a SEU. Unfortunately, with decreasing feature size, the probability that the hit of a single ionizing particle affects more than one component of the circuit increases [BBR+08; DW11]. However, the substrate change in newer VLSI technologies, used for lower features sizes, led to a higher resilience to SEE, due to a reduced volume of material susceptible to SEEs [RAGM13].

2.3.4 Fault-Tolerance Techniques

Basically, there are three main paradigms for handling faults: Fault prevention, fault detection and recovery, and fault masking.

A popular *fault prevention* technique is radiation hardening, which allows to reduce the effects of SEEs on a circuit. This can be achieved by special layout considerations [BLSC10; Lac08] in form of physical arrangement and design of the relevant parts of a transistor.

Fault detection and fault masking strive for containment of faults. We define a *fault-containment region* (FCR) as a collection of components that together share one or more common resources, which can be affected by a single fault [KK07; Kop97]. FCRs have to be designed such that the failing of two FCR is uncorrelated. As failures can propagate, some sort of detection and containment, or masking is required to ensure this. Recall that metastability cannot be contained by a FCR [FFS09], as its power exceeds even Byzantine failures.

The ability to detect a fault depends on the applicable system model and the knowledge about the behavior of the components available. Usually, *fault detection* is implemented by enhancing the data generation of a component with redundant information, via specific encoding, that allows to detect the occurrence of an error [KK07]. Hence, the recipient is able to verify the integrity of the received data. Simple codes, e.g., parity, allow only the detection of an error. More advanced codecs, however, support the recovery of the original information, given that only a limited number of bits are affect, e.g., *error-correcting codes* (ECCs) [Ham50; Nic05; PW72]. Fault detection is typically used in the context of *failure detection, isolation and recovery* (FDIR) approaches, which use dedicated circuitry (the FDIR unit) to detect a failure, diagnose the fault causing it, and to initiate recovery actions. Hence, FDIR architectures employ methods to isolate faulty components from the remainder of the system. This is not limited to the removal of a component from a quorum, but can also include putting the system in a safe mode. Furthermore, this type of architecture allows the re-integration of a previously faulty components after its recovery, e.g., via a reset. Obviously, the affected component cannot provide its usual service during the latter. FDIR is thus suitable for

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

high-reliability applications, e.g., spacecrafts [SD13] where maintenance is infeasible, but not for high-availability ones.

For high-availability applications, the preferable fault-tolerance technique is *fault masking*. This techniques generates sufficiently redundant information to transparently mask a failure, and thus prevents an error from being forwarded. For example, a redundant design on the digital gate level [GM00] allows to detect an induced soft error that is then masked by a majority voting. Such a voting is also used by *triple modular redundancy* (TMR) [DKM05; LV62; Von56]. TMR triplicates every component and performs a majority voting upon the results of the components. The cost for implementing TMR-like approaches are mainly paid in area and power. If time is not an issue and faults are benign, duplicate execution with the same component is also a possibility [CNV+00]. However, duplicate execution is susceptible to intermittent faults of the component.

In any case, the voting is a critical part of approaches using information redundancy. The system's design must ensure that the voter cannot become a single point of failure, which would effectively negate the purpose of the added redundancy. Furthermore, a simple binary comparison by the voter may not suffice when the outputs of the components can produce non-identical results in non-faulty executions. If this cannot be avoided, some sort of *k-plurality voting* is required [KK07], where the voter decides on an outcome based on k of its inputs being within some ε-environment, according to a defined measurement.

From the fault masking techniques presented above, it is apparent that *redundancy* is the key. Note that the requirement for redundancy is not limited to the computation alone. Recall the fault-tolerant consensus problem: Under the assumption that the system experiences at most f fail-silent components, at total number of $f+1$ components in the system is sufficient for safe operation. However, for f Byzantine components $3f+1$ components are required [PSL80]. Furthermore, due to the inconsistent behavior a Byzantine failure can exhibit, simple detection and masking methods cannot be used [DHP+04]. Hence, Byzantine fault-tolerant architectures [LSP82] also require a minimum connectivity between the components of the system [Dol82]: The faults must not be able to modify/forge a significant part of the communication in the system.

2.3.5 Self-Stabilization

Self-stabilizing systems [Dij74; Dol00; Sch93] are designed in such a way that they recover to a correct state after being forced into an illegitimate state due to, e.g., an overwhelming number of transient failures caused by ionizing particles [VPS+13]. An essential part of this definition is that the system will (i) reach a known good state in bounded time (*convergence*), from every illegitimate state, and (ii) once in a good state, the system will not leave it without external intervention (*closure*). This holds, however, only if no further faults occur during stabilization and thereafter. The period between the time when no new transient failures occur and reaching a good state is

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

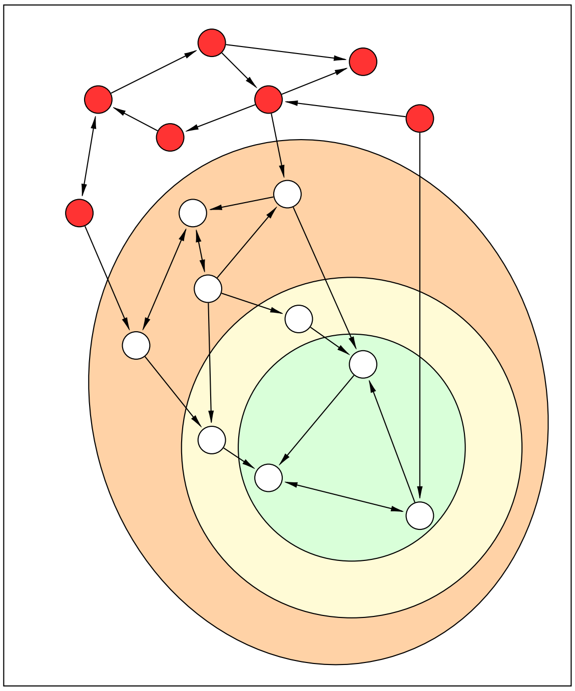


Figure 2.8: This figure shows the complete state space, which is a superset of multiple partial sets. The red nodes symbolize states which are arbitrary. The nodes in the orange area are pseudo-legal states, the nodes in the yellow area are safe states and the green area is the set of legal states. Copied from [Per13].

called *stabilization time*. Following the visualization from Figure 2.8, we can divide the state space into 4 types:

Legal States The legal states are the good states of the system, and colored green in the figure. They are the intended state space in which the system should operate. Hence, without a failure, the system will not leave this state space.

Safe States This state space, yellow, includes the legal states and has the property that all transitions from states in this space go to the legal state space.

Pseudo-Legal Sates This orange set of states distinguishes itself from the safe state space in that there is a sequence of transitions leading to the legal states space, albeit it may also reach erroneous states before doing so. That is, while the system behaves within its specification towards its environment, internally it has still a partially corrupted state.

Erroneous States This set covers the remaining states.

On an algorithmic level, many algorithms have been found that achieve self-stabilization. Generally speaking, the ability to self-stabilize is the result of a limited

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

effect of the (faulty) late evolution on the current state of the algorithm. That is, after some time the erroneous history ceases to affect the future, and the system converges to a legal state. Compared to FDIR, this removes the need for an explicit reintegration of a previously faulty component. An emerging property is that self-stabilizing systems do not need to be initialized after the system's boot to reach a well defined state after bounded time.

Consider that fault-tolerance allows a system to operate correctly in the presence of the designed number of failures f, by means of masking them. Unfortunately, however, if the system experiences more than f failures, the behavior of the system is unknown and correct operation at some future point cannot be guaranteed. Combining fault-tolerance with self-stabilization allows correct operation under f failures, and recovery to correct operation after a temporary violation of the bound on concurrently active failures f. However, this requires an extension of classic self-stabilization to fault-tolerant self-stabilization, where a bounded number (f) of persistently faulty components are allowed also during stabilization. Fault-tolerant self-stabilization can hence be seen as a *never give up strategy* (NGU) [Kop97], which is a last resort to bring the system back into a well defined state.

However, self-stabilization is a non-trivial property for a system to achieve and may not be achievable in every problem domain. For example, the composition of self-stabilizing circuits may not be self-stabilizing if a feedback loop is involved, unlike purely forwarding circuits [DFL+14]. Moreover, a layered architecture cannot be made self-stabilizing by just providing a self-stabilizing layer, without the layer below being self-stabilizing as well. Furthermore, self-stabilization comes at a price: Besides the need for a periodic deadlock prevention event and (small) bounded-sized variables, self-stabilizing solutions usually do not provide the application with a notification that the system has stabilized.

2.4 Clock Generation

In this section, we will introduce methods for distributed clock generation. However, before we talk about generating a clock, we need to define what we understand by a *clock*. A clock is a device that measures the progress of time by generating so called *ticks* with a defined temporal distance, called the period. Distributed clock generation is the problem of implementing a clock C_i at process p_i in a distributed system consiting of processes $p_1, \ldots, p_n$, $n \geq 2$.

Definition 2.1 (Clock):
A clock $C_i : \mathbb{R} \to \mathbb{N}$ maps (continuous) real-time $t \geq 0$ to (discrete) clock time $T = C_i(t)$. Correct clocks generate (right-continuous) step functions, i.e., assume all values $0, 1, 2, \ldots$ in strict succession, where t_i^k with $t_i^0 = 0$ denotes the smallest real-time with $C_i\left(t_i^k\right) = k$. We say that C_i generates tick k at time t_i^k. □

In terms of clock synchronization, correct clocks will be characterized by (i) their *clock skew*, a time bound within which they produce the same tick k, (ii) their *tick*

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

separation time between any two consecutive ticks, and (iii) their *precision* bound, which is the maximum difference in the clock readings between any two correct clocks at the same real-time. Formally, these parameters are defined as follows:

Definition 2.2 (Clock parameters):
Given a fixed pair of correct clock i, j, they are characterized by:

1. *Skew bound* $\boldsymbol{\sigma}'_{i,j}$, defined as $\boldsymbol{\sigma}'_{i,j} = \left[\sigma'^{-}_{i,j}, \sigma'^{+}_{i,j}\right]$ with $\sigma'^{-}_{i,j} \leq t^k_j - t^k_i \leq \sigma'^{+}_{i,j}$ for all $k \geq 0$.
2. *Tick separation time lower bound* Γ'_i, defined as $\Gamma'_i \leq t^{k+1}_i - t^k_i$ for all $k \geq 0$.
3. *Signed precision bound* $\boldsymbol{\pi}_{i,j} = \left[\pi^{-}_{i,j}, \pi^{+}_{i,j}\right]$, which is formally defined via $\pi^{-}_{i,j} \leq C^k_j(t) - C^k_i(t) \leq \pi^{+}_{i,j}$ for all $t \geq 0$. □

2.4.1 Traditional Clock Generation

The commonly used mechanism for generating a clock signal for a digital circuit is a quartz crystal oscillator, which is easy to implement and cheap. The behavior of quartz oscillators is well understood, however, they need compensation for temperature and voltage drift, especially when used to generate high frequencies, as to keep the resulting frequency drift low. Furthermore, the crystal is susceptible to mechanical shock, which is problematic for space applications: During the launch phase, it is common that high g-load shocks occur.

In the last years, new developments made it feasible to use *micro-electro-mechanical system* (MEMS) as oscillators [NYH+15]. Experimental data shows exceptional temperature stability, resilience to gravitational shock, and aging properties [SiT14; SiT15]. However, in sharp contrast to the mostly analog circuitry for controlling the oscillation frequency of a quartz crystal, the control circuitry of a MEMS is implemented using digital VLSI technology. Hence, ionizing radiation affecting the control circuitry can affect the frequency of the generated clock [TKP+15]. Nonetheless, special adaption for space application seems reasonable as the MEMSs themself show high resilience towards radiation effects [BBD+14].

2.4.2 Clock Generation via Transistor Circuits

Tick generating components can also be implemented with transistor-level circuits, such as astable multivibrators [PS57]. However, we will focus on fault-tolerant clock generation, i.e., implementations that do not have a single point-of-failure.

Considering only combinatorial gates, the simplest method for generating a periodic signal is a ring oscillator: A ring oscillator consists of an odd number of inverters in a feedback loop. It is easy to implement but has several deficiencies: It is not fault-tolerant and the frequency of the oscillation is defined by the delays of the wires and the gates in the feedback loop. This is also the source of a large frequency instability: PVT

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

variations cause fluctuations of the delays, leading to jitter and phase noise [HLL99]. Furthermore, SETs can insert stray pulses into the feedback loop, effectively causing a multiplication of the frequency, if additional hardware for removing these pulses is not implemented [CLM+14].

An interesting approach has been proposed by Fairbanks and Moore: In [Fai06; FM05] an accurate distributed oscillator is presented that consists of simple transistor circuits arranged in a grid. Albeit, its behavior under faults has not been studied, initial observations (cf. Section 7.1) suggest that the circuitry is indeed fault-tolerant to some extent.

The system is constructed by two types of circuit cells that are alternately arranged in a two-dimensional grid. The cells are called pull-up and pull-down cells. This division leads to a setup where the four neighbors of a pull-up cell are pull-down cells, and vice versa. The name of the cells describes their basic function: Every cell can only drive the wires shared with their neighbors into one state, i.e., the pull-up cells can only drive the wires to high. The transition itself is feed back into the cell and weakens the strength of the drivers, allowing the internal state of the cell to be overwritten by its neighbors over the shared wires.

The system thus oscillates by half of the cells inverting their output, and hence the wires to their neighboring cells. This forces the other half of the cells, in turn, to also invert their output, leading to the desired oscillating behavior. The frequency of the oscillation is controlled by careful sizing of the transistors in the circuit.

A simpler approach at distributed clock generation is presented in [AMB07; MA03; MA04]. Inverters are connected in regular structures here, such that loops with an odd number of inverters are formed. Simply put, the smallest sub-structure is a triangle with 3 inverters. The proposed solution spreads this structure over a larger area, thus combining clock generation with a clock distribution, by allowing components to connect to the structure at the closest node to get their clock signal. Except at the border of the IC, every inverter has at least two other inverters driving its input. Hence, some limited fault-tolerance is in principle possible, albeit this clock generation has not been explored systematically yet.

Similarly, [GC00; KCS+10; SGAZ14] present approaches of a sparse rectangular grid of *phase locked loops* (PLLs), with [SGAZ14] using *all digital PLLs* (ADPLLs). As will be described in more detail in Section 2.4.4, every PLL has its own oscillator and clocks a small section, called tile, of the IC. At the boundary to the adjacent tiles, phase detectors, which measure the skew between its inputs, are placed to compare the local clock to the clocks of the neighboring tiles. The skew difference is used as a parameter for adjusting the frequency of the local tile's frequency. The proposed approaches primarily differ in the method used to implement the phase detectors. Depending on the relation between the phase difference and the output generated by the phase detector, these solutions may be prone to be caught into a so called mode-lock [PN95]: While their is a phase difference to every neighbor, the overall error function may be

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

such that they cancel each other out. Thus, a local correction is prevented and the synchronization error with respect to each neighbor is never corrected.

All papers thus employ methods to prevent a mode-lock. [GC00] uses phase detectors that produce a non-linear error function that focuses on high-frequency changes in the phase difference. [KCS+10] uses a bang-bang style error function that produces a binary signal. This is a commonly used error function for a phase detector, except that the error signal of each neighbor is weighted differently. Hence, all neighbors contribute to the synchronization when the local clock is, e.g., ahead of each neighbor. However, if the clock is between neighbors, it will synchronize primarily with the neighbor(s) with the highest weight.

The method used in [SGAZ14] not only considers the sign of the phase error, but also its value. Additionally, it uses a weighting similar to [KCS+10] as well. However, the topologies proposed are different: They propose a rectangular structure, where every inner node considers the phase information from 4 neighbors. The irregularity at the borders allows to integrate an external reference clock as a fourth neighbor at some point. Whereas in the regular structures every node considers its complete neighborhood, in the "swimming-pool" topology [SAGZ14] the nodes at the border do not: Similar to a pool edge, the inner nodes (the water) have no influence on the border. That is, while the nodes at the border only consider their neighbors which are also in the border for synchronization, the other nodes do not limit their neighborhood in any a way.

To the best of our knowledge, none of these approaches has been analyzed for their fault-tolerance properties, not to speak of self-stabilization. We conjecture that the mode-lock prevention logic is a favorable point of attack for Byzantine faults, however. Furthermore, [SGAZ14] uses *time-to-digital converters* (TDCs) to determine the value of the phase difference, which can be prone to metastability. However, it is not stated whether the phase detectors have been implemented in a somewhat metastability-containing way [FFL18].

2.4.3 Distributed Algorithms for Clock Synchronization and Clock Generation

Clock synchronization, which is the problem of implementing synchronized clocks on top of free-running hardware clocks, has been studied in distributed computing since decades [AW04; Sch86; SWL90]. In a way, it can be viewed as repeatedly reaching consensus on certain clock values.

We follow the definition of clock synchronization from [AW04], and denote with $HC_i(t)$ the hardware clock of process i at time t. As the hardware clock may be inaccurate, its value can differ from real time, i.e., the clock *drifts*. Hence, repeatedly, a correction has to be performed such that the adjusted clocks $AC_i(t) = HC_i(t) + adj_i(t)$, with adj_i denoting the adjustment variable, remain synchronized. For achieving clock synchronization, any hardware clock employed must have *bounded drift*, defined as

$$(1+\rho)^{-1}(t_2 - t_1) \leq HC_i(t_2) - HC_i(t_1) \leq (1+\rho)(t_2 - t_1). \tag{2.1}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Given such a bounded-drift hardware clock, we define the following properties that a clock synchronization has to fulfill based on [AW04]:

Clock Agreement There exists a constant precision π such that in every execution, for all times t and all non-faulty processors p_i and p_j,

$$\left|AC_i(t) - AC_j(t)\right| \leq \pi. \tag{2.2}$$

Clock Validity There exists a positive constant γ such that in every execution, for all times t and every non-faulty processor p_i,

$$\begin{aligned}(1+\gamma)^{-1}\left(HC_i(t) - HC_i(0)\right) \leq AC_i(t) - AC_i(0) \leq \\ \leq (1+\gamma)\left(HC_i(t) - HC_i(0)\right).\end{aligned} \tag{2.3}$$

The goal of a clock synchronization algorithm is to fulfill the above properties with small π and γ.

Lower bounds on the achievable quality of synchronization differ from setting to setting. For example, Biaz and Welch show in [BW01] that it is impossible for a deterministic clock synchronization algorithm to guarantee a worst-case precision between *all* pairs of processes that is better than $^{D\varepsilon}/_{2}$, where D is the diameter of the underlying communication graph and ε the end-to-end delay uncertainty of message transmission. In the setting of gradient clock synchronization, where the goal is to closely synchronize only direct neighbors in a sparse topology, Lenzen, Locher, and Wattenhofer have shown that the skew is lower bounded by $\Omega(\varepsilon \log D)$ [LLW10].

Before advancing into related work on this topic, we note that the optimal resilience for clock synchronization in the presence of f Byzantine faulty processes/clocks is bounded by $n \geq 3f+1$ [PSL80].

Srikanth and Toueg proposed in [ST87] a Byzantine fault-tolerant distributed algorithm for clock synchronization in fully connected networks. The principle is simple: Every process broadcasts a message when it has reached the k^{th} resynchronization point, which can be viewed as the beginning of synchronization round k. Upon reception of $2f+1$ round k messages, a process *accepts* that a majority of processes have reached round k. Furthermore, if a process receives $f+1$ messages for some round r, it assumes that round r is the current round, i.e., the process *catches up*, and broadcasts a round r message. In all cases, it advances its adjusted clock accordingly. The worst-case precision between correct processes achievable by this algorithm is proportional to the maximum communication delay between them.

The algorithm is fault-tolerant, due to the thresholds based on f required for deciding. Furthermore, it allows some sort of reintegration of faulty processes due to the catch-up rule with threshold $f+1$. However, the algorithm is not self-stabilizing if all clocks are desynchronized.

Lundelius-Welch and Lynch [LL88] invented the well-known fault-tolerant average/midpoint clock synchronization algorithm, which works as follows: Every process

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

broadcast a message upon reaching the beginning of the next synchronization round, and sets a local timer that timeouts when the process should have received messages from all its neighbor. Every received message from some neighbor is time-stamped on arrival with the process's local clock value. On timeout of the timer, the f smallest and f largest differences of sending and receiving times are dropped, and the average/midpoint[5] of the remaining values is computed. Based on the average/midpoint, a correction value for the local clock is computed. This algorithm achieves a skew proportional to the uncertainty of the end-to-end delays involved. As [ST87], the algorithm requires a fully connected topology.

Implementations of this type of algorithm are being used in industry, e.g., in [BBB+03; KB03]. These implementations are usually software/hardware-hybrids, which allow to reduce the message delay uncertainty.

A pure hardware implementation of this approach is given by Kinali, Huemer, and Lenzen in [KHL16]. It works for $n \geq 3f + 1$, and is able to recover from transient faults if at least 2/3 of the processes remain synchronized.

In [KL16] Khanchandani and Lenzen show how to merge a system based on the clock synchronization concept of [LL88] with FATAL [DFLS14], hence enabling self-stabilization and, at the same time, achieving better precision bounds.

A problem related to distributed clock synchronization is distributed clock generation. The challenge here is to generate a clock, i.e., a sequence of clock ticks at the processes, which are approximately synchronized, but without having free-running hardware clocks underneath.

The DARTS approach [FDS10; FFSK06; FSFK06] implements a distributed tick generation based on [WS09], a variant of the consistent broadcasting primitive used in [ST87]. The basic idea is to use pipelines like Sutherland's micropipelines [Sut89] to store the incoming ticks from the neighboring processes. However, due to its internal workings, a reintegration of a faulty process is not always possible. Furthermore, transient faults can inject stray ticks into the pipelines, which become persistent and thus severely degrade the system's resilience.

In [DFLS14] Dolev *et al.* presented a self-stabilizing Byzantine fault-tolerant clock synchronization termed FATAL, which can also be implemented using asynchronous hardware [DFL+14]. On top of FATAL, a high-speed tick generation termed FATAL$^+$ can be executed, which allows to improve both the precision and the generated frequency. However, FATAL is only probabilistically self-stabilizing. That is, the stabilization time is bounded only with a certain probability. However, FATAL can stabilize even if a third of the processes behave Byzantine during the stabilization.

2.4.4 Frequency multiplication

The frequency of the clock ticks provided by C_i might not be sufficient for specific application. For example, the synchronized clocks generated by approaches like

---

[5] The midpoint is the average of the lowest and the highest element in a set.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

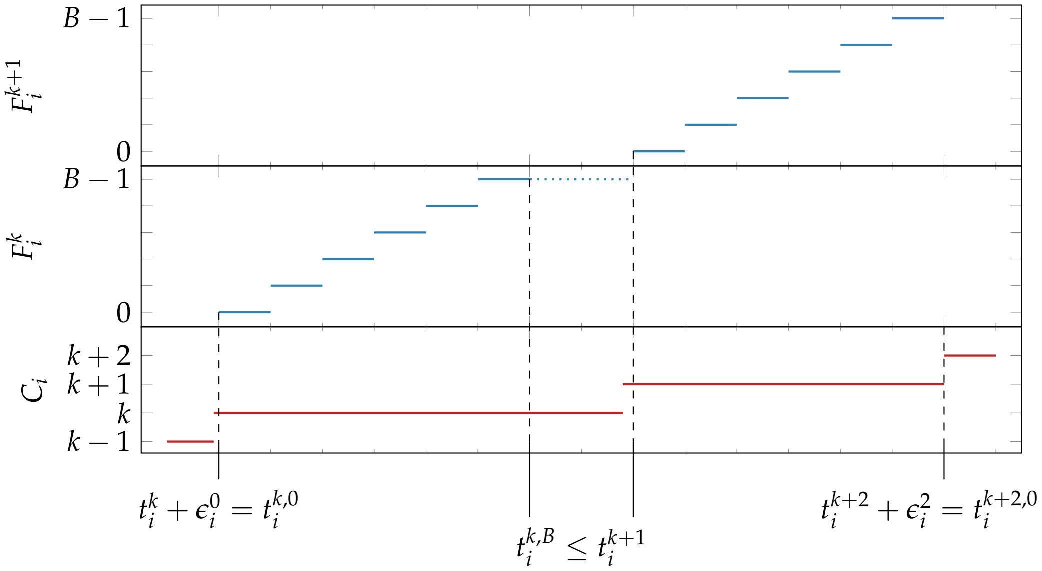


Figure 2.9: Visualization of the slow clock C_i and its corresponding fast clocks F_i^k and F_i^{k+1} with $B = 6$. Note that while ϵ_i^0 and ϵ_i^2 can differ, they both are in $\left[\epsilon_i^-, \epsilon_i^+\right]$.

[DFL+14; DFL+16] have typically a fairly low frequency. Hence, generating additional clock ticks between the ticks of the slow clock C_i is required.

We therefore define the *fast clock* F_i^k as follows: $F_i^k : \mathbb{R} \to \{0, \ldots, B-1\}$ maps (continuous) real-time $t \geq 0$ to (discrete) clock time $T = F_i^k(t)$. We will refer to C_i as the corresponding *slow clock* to remove ambiguity. We assume that every tick k of the slow clock C_i starts, after a possibly varying slack time within $\left[\epsilon_i^-, \epsilon_i^+\right]$, a dedicated fast clock F_i^k. This fast clock generates exactly B ticks $0, \ldots, B-1$ in strict succession, a so called *burst*, plus a final *virtual* tick B. This virtual tick is only used conceptually for cleanly separating consecutive fast clocks F_i^k and F_i^{k+1}. Analogously to the slow clocks, we say that a fast clock F_i^k generates tick b at time $t_i^{k,b}$. Incorporating slack between the slow clock tick and the first tick of the fast clock allows to cover circuit delays and other timing uncertainties.

Formally, we require from a fast clock the following properties:

Definition 2.3 (Fast clock parameters):
Let $F_i^k(t)$ be the fast clock started by the correct slow clock C_i's tick k, such that $t_i^k + \epsilon_i^- \leq t_i^{k,0} \leq t_i^k + \epsilon_i^+$ for $0 \leq \epsilon_i^- \leq \epsilon_i^+$, cf. Figure 2.9. It is characterized by

1. the frequency range of the fast clock: $\forall k \geq 0, \forall b \in \{0, \ldots, B-1\} : \frac{1}{f_{i,\max}} \leq t_i^{k,b+1} - t_i^{k,b} \leq \frac{1}{f_{i,\min}}$,
2. the number of fast clock ticks B, which must satisfy $\frac{B}{f_{i,\max}} \leq t_i^{k,B} - t_i^{k,0} \leq \frac{B}{f_{i,\min}} \leq \Gamma_i$.

□

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

The following Lemma 2.4 is an easy consequence of Definition 2.2 and Definition 2.3, which takes into account the (variable) but bounded slack for starting the fast clocks:

Lemma 2.4 (Fast clock skew parameters): *Correct fast clocks* F_i^k, F_j^k *started by correct clocks* C_i, C_j *satisfy* $\sigma_{i,j}^- \leq t_j^{k,0} - t_i^{k,0} \leq \sigma_{i,j}^+$ *and* $\Gamma_\ell \leq t_\ell^{k+1,0} - t_\ell^{k,0}$ *for all* $k \geq 0$ *and* $\ell \in \{i, j\}$, *where* $\sigma_{i,j} = \left[\sigma'^-_{i,j} - \epsilon_i^+ + \epsilon_j^-, \sigma'^+_{i,j} - \epsilon_i^- + \epsilon_j^+\right]$ *and* $\Gamma_i = \Gamma'_i - \epsilon_i^+ + \epsilon_i^-$. □

Theoretically, the concept of a fast clock is simple. In practice, however, attention has to be paid to implementing it correctly. Using a free-running clock underneath, e.g., a quartz oscillator, requires the masking of the clock pulses at the end of the burst. This masking, however, must be done with care to ensure a proper timing of every clock pulse. To achieve this, upon deciding to unmask the output, the masking system has to wait for the end of the current pulse. Otherwise, the unmasking may happen at an arbitrary point during the first pulse, thus causing the output of a degenerated pulse. From this follows that such a system has a slack in the range of a pulse period.

Furthermore, this type of designs are prone to metastability, as has been shown in [NS14]. Related approaches, e.g., [NS15], can remove the metastability issues but not the loss of a fast clock tick.

Hence, techniques that directly multiply the slow clock's ticks are of interest here. The standard most common component to achieve this is a *phase locked loop* (PLL). A PLL is a circuit that tries to match the phase of the generated clock signal with the input clock signal. By means of a feedback loop, the signal generated by an internal oscillator is compared to the input. Adding a prescaler into the feedback loop allows for an output signal that has a higher frequency than the input signal. When the scaled signals match in frequency and phase, the error generated by the phase detector will go to zero, and the PLL has *locked on*: an output signal is generated with a stable frequency in this case.

Unfortunately, PLLs are sensitive to jitter and noise of the input signal, as this causes an unstable error function produced by the phase detector. This prevents the PLL from keeping the lock on the input signal, hence leading to an unstable output signal.

From a design point of few, a PLL can be made rather insensitive to jitter of the input, at the cost of increased sensitivity to the jitter of the oscillator. Nonetheless, the key challenge is the time until a lock has been established: Without a frequency lock, the strict relation of B fast clock ticks per slow clock tick cannot be guaranteed, and Definition 2.3 and Lemma 2.4 will be violated.

Unlike a PLL, an *all digital PLL (ADPLL)* consists of digital circuitry only. It is typically less sensitive to noise, jitter, and other environmental factors. An implementation of an ADPLL is described in [ON03]. The proposed design uses a free running oscillator, based on a ring oscillator, that drives two counters. One measures the current period of the input signal, the other one generates the output clock signal. The intention of the design is similar to our fast clocks: The reference counter provides enough information

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

to the output counter to evenly space the cycles of the generated signal during a period of the input signal. However, the proposed design is susceptible to jitter on the input signal, in the sense that ticks of the burst may be lost after a rapid change in frequency. Additionally, the required synchronizers can only reduce the probability of a metastable upset.

The digital PLL implementation in [WY03] uses a *time-to-digital converter* (TDC) to determine the phase and frequency difference, which allows faster locking. However, it suffers from metastability issues due to the usage of a TDC. Furthermore, the design also tries to lock on the input signal like every other PLL, which may lead to the same issues with large input jitter.

Thus, as with PLLs, ADPLLs cannot guarantee the properties needed for a fast clock.

Digitally controlled clock multipliers are similar to PLLs in the sense that their goal is also to increase the frequency of their input signal. However, instead of continuously generating an output signal, they generate only a fixed amount of clock cycles corresponding to their multiplication factor [NT96]. To ensure that the multiplication factor is kept, there must be a gap, i.e., a margin, between the end of the last tick of the output clock, and the next input clock tick. In [ONM+00], a method is described how this margin can be reduced, by gradually changing the frequency of the internally oscillator according to the frequency of the input signal. The principle used to control the frequency is the same as in the successor work [ON03] reviewed above.

With enough margin, even a high input jitter does not prohibit the successful operation of a digitally controlled clocked multiplier. Thus, using an appropriate parameterization, a digitally controlled clock multiplier can be used to produce a fast clock signal. The multiplication factor is equal to the burst length, which is hence easily configurable.

2.5 Clock Distribution

As all the state-holding components of a synchronous circuit require a clock signal, the problem of distributing a synchronized clock signal to a large number of spatially distributed components on the chip arise. It is of course infeasible to add an oscillator to every component and try to keep those oscillators synchronized. Hence, it is common to have a single oscillator and distribute its signal via a *clock distribution network* (CDN), e.g., via a clock tree. In a clock tree, as the one shown in Figure 1.1, a clock signal is provided to all components attached to the leaves of the tree. In large clock trees, the clock signals need to be refreshed by inserting clock buffers,[6] which keep the signal rise times within the desired bounds. This is important as a slow rise time may cause the transition to reach the threshold voltage of different components at vastly different times and thus increases the skew.

---

[6] Clock buffers are typically constructed by having two inverters in series.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

A promising approach is to split the design into multiple clock domains and drive each domain via a smaller clock tree. These clock domains can either be independent, which is the case in a *globally asynchronous, locally synchronous* (GALS) architecture, perform some sort of clock synchronization, or are derived from a common base clock that is distributed via a low-degree clock distribution grid. Nonetheless, a fairly small clock tree is required at some point to keep the clock deviations due to delay differences and variations low.

2.5.1 Clock Trees

We consider the classical H-tree first, which is a common method for distributing a clock signal in a circuit. Recall the structure of a H-tree (Figure 1.1), which is a planar graph that would allow for a routing of the tree in one layer of the IC. A H-tree guarantees an equal wire length from its root to all leaves, which ensures that the clock signal arrives at all leaves (almost) simultaneously. The regular structure results in a height that is logarithmic in the number of leaves, while the number of wires grows linear with the leaves.

There are two main measures for defining the quality of a CDN:

1. The difference in the arrival time of the same clock tick at two leafs is called *skew*. The skew can be differentiated in the local skew, between neighboring components and the global skew between arbitrary components.

2. The variance in the arrival of successive clock ticks at the same leaf, causing a fluctuation in the frequency, is called the *jitter*.

To achieve a skew/jitter in the ps range, however, the wire geometry and the buffers need to be carefully engineered [LKM10] to ensure equal signal propagation delays. This is further complicated by possibly different fan-ins, wire resistance, and capacitances of the attached leafs [JH01] that need to be charged at every clock transition. The delays thus depend heavily on the technology and the feature size [Fri01], and suffer from PVT variations as well [BRd+06; CCL+08]. Another subtle variance comes from the fact that a component may not switch completely at every clock cycle, depending on the data applied, which causes a difference in the capacitance that needs to be charged by the clock signal.

In the following, we discuss techniques that focus on minimizing the skew. Fortunately, most designs do not require zero skew between arbitrary components. The global skew from one corner of the IC to the other is hence usually irrelevant, as one usually avoids signals traveling over the whole IC [LM11]. However, due to the structure of a tree, also spatially close components may be clocked by very different branches of the tree. As these branches may already have diverged at the root of the tree, this can result in a fairly large local skew. The main technique for optimizing the skew is to introduce some artificial delays in the clock tree. As this decreases the skew, the maximal achievable operational frequency can be increased accordingly [Fri01].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

There is a limit on the size of a clock tree, however, as a larger clock tree requires larger buffers and longer wires, due to the logarithmic height of the clock tree. Studies have shown that the clock tree alone can consume 30–50% of the IC's power [Fri01; LKM10; RRA15]. This is not surprising, as the CDN switches with every clock cycle and the dynamic power of a IC is calculated by $P = CV^2f$, with C being the capacitance, V the voltage (VDD), and f the frequency. Furthermore, the buffers in the clock tree introduce a delay that can add up to several clock cycles [CCL+08], and also increase the delay variance. Consequentially, buffers are nowadays the principal source of the remaining skew in a CDN [Fri01].

As PVT variations are, in general, long-term effects that gradually change, they can be compensated. It should be noted, however, that PVT variations primarily influence the conductance of transistors, so the buffers are typically affected more than the wire resistances [FP86]. Designs for compensating PVT variations can be found in [GD98; KBD+01; SS01], for example. In a nutshell, the phase difference of the clock signals at certain points in the tree are measured and used to configure variable delay elements in the CDN, thus reducing the skew by affecting the signal propagation delay. Similarly to the variable delay elements, solutions based on so-called deskew buffers have been proposed in [ORM04; ORM05], which allow to synchronize a clock signal to a reference clock signal. The advantage of deskew buffers is that they affect the current clock cycle, compared to the delay elements that can only affect future clock cycles.

Short-term effects on the delays, however, require structural changes. Adding cross-links between branches of the clock tree have been shown to reduce to the skew between neighboring branches [MK11; RHM06]. Advancing the concept of cross-links, Lee and Markov [LM11] proposed multi-level trees: A multi-level tree has additional root-to-sink paths, where a section of the path is shared with other such paths. These multi-level trees have also shown to counteract PVT variations.

2.5.2 Clock Mesh

Moving the idea of multi-level trees even further results in clock meshes. A clock mesh is based on a rectangular grid structure that spans the whole IC, to which a component can be connected, called *tapped*, at any point. The advantages of clock meshes are that they are rather insensitive to changes in position and size of their load [SS01], and that spatially close components also acquire their clock signal from taps of the clock mesh that are nearby [RWM06; YWC+06]. The independence of the load also reduces the design effort required for a clock mesh. Moreover, due to the spatially closeness, meshes achieve very low local skew. However, they cause a significant overhead in terms of area and power requirement, due to the need to charge higher capacitances [RMW+01]. Furthermore, meshes of certain size still require an overlaying clock tree, over which the clock signal is applied to the mesh, via buffers. Thus, intermediate architectures, intermixing clock mesh(es) and the clock tree(s), have also been considered [YWC+06].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

2.5.3 Faults in Clock Distribution Networks

Having a clock tree as the last element in the CDN before the actual components reveals a fatal flaw: A single wire defect, e.g., due to some manufacturing defect or aging, may cause the underlying branch of the clock tree to black-out. This type of fault in the CDN has been proven to be hard to find [MDMR01], as such faults may only locally affect operation by degrading the shape of the clock signal. Approaches like adding cross-links may alleviate this type of problem, albeit, their main goal is still the reduction of the skew [MK11; RHM06]. The relevance of defects in the CDN has been shown in a study by Metra, Di Francescantonio, and Mak [MDM04], revealing that the probability for faults in the CDN is two order of magnitude higher than for any other type of fault in the IC.

Given the susceptibility of clock trees to faults it is worth noting that the usage of clock meshes at the lowest level in the CDN would not solve the problem: Clock meshes are also affected by stuck-at faults of the buffers, and are equally influenced by SEEs. The effects of SEEs on the CDN have been studied extensively in recent years [ACB+15; CCK+12; MGC+16; WHG+09], as the soft error rate in a CDN cannot be neglected [SSP+05] any more. The soft-error rate caused by a clock mesh itself is less than the one of a clock tree, due to the higher capacitance of the clock mesh compared to the wire between two buffers in a clock tree. However, if a SEE affects a buffer supplying the mesh, the whole mesh is affected. Hence, the effective error rates of clock meshes are higher than those of clock trees [CK14; WMC+16].

To reduce the effect of SEEs on CDNs, hardened standard cells designs [GK08; ZS05] and clock buffer designs [DGKC09; MZ08; SSP+05] have been proposed. However, the analysis of these designs focuses on the specific elements alone, without considering the effects of on the whole CDN. A holistic view of the CDN [CCH12] is, however, required as not every element needs to be hardened: The most vulnerable elements are buffers that are small or have a low connectivity in the CDN [CK14; CKTR12; KCOW08; SSP+05], i.e., drive a lower capacitance. Hence, radiation-hard-by-design approaches [CCH12] or duplication of the clock tree [ACB+15; GCRC15] have been considered in literature. As one can expect, these approaches are effective in reducing the error rate, however, they require additional area and power making them infeasible in some applications.

For most of the considered faults, their probability increases proportional to the size of the CDN. An approach reducing the effective size of the CDN is a GALS architecture [Cha84]. In a GALS architecture, a small set of components forms a synchronous island, supplied by a common CDN. With asynchronous communication between these islands, the complete design has multiple independent CDNs and thus lacks a single point of failure that could affect the whole design. However, if a common base clock is provided to the synchronous islands for implementing multi-synchronous GALS, this does not hold anymore. In multi-synchronous GALS, mesochronous clocking [Mes90; SG03; TGL07] is used to also guarantee some synchrony between different islands. See Section 2.6 for further information.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

For the sake of completeness, we present a clock distribution method applicable to the common base clock GALS approach [SS01], which is similar to the clock generation mechanisms describe in [GC00; KCS+10]. There, a *master PLL* generates a base clock that is distributed, via variable delay elements, to secondary PLLs. The latter supply a clock signal to a local clock distribution grid for every island. To compensate the skew between the grids, phase differences are measured and the error is considered in the parameterization of the delay elements. Furthermore, the clock provided to one grid is fed-back into the master PLL. This prevents mode-locking by construction, as only delay variations are compensated while avoiding any feedback between the local grids.

The single point of failure issues is not limited to a single IC, however. Clock distribution in 3D chip design technology, where multiple chips are stacked onto each other, has problems similar to on-chip CDNs. The connections between these chips are primarily established via *through silicon vias* (TSVs), which are vertical conductors with pads at the bottom/top of the silicon, allowing to bond two chips together. Note that the TSVs are not only used for data transmission, but can also be used for clock distribution. A precise bonding of the chips is paramount to achieve high quality connections over the TSVs: Misalignment of the pads can lead to an increased resistance and hence delay variations [LML+08], or to even no connection at all. Furthermore, classical breakdowns, i.e., wire defects by electromigration, of the TSVs can also occur over time [LML+08]. An approach to reduce the effect of the misalignment problem is to place redundant TSVs [LSH+11; PK13; PK15]. However, they require complex calculations for their positioning, i.e., assessing whether and where an additional TSV needs to be placed. Moreover, as this redundancy should only be used when needed, additional area is required for routing and transmission gates. Note that the transmission gates are also required so that a redundant TSV can be considered as a backup for more than just one TSV. However, multiplexing of a TSV is not possible with the transmission gates.

2.6 Communication across Clock Domains

Since the synchronous design model for the entire chip is a thing of the past [TGL07], modern systems comprise of multiple clock domains, sometimes with different frequencies. As discussed in Section 2.1.2, implementing data communication between these islands is tricky. Without stringent guarantees on the synchronization of these islands, synchronizers are required to reduce the probability of metastable upsets (cf. Section 2.2). Besides the lingering metastability issues, there are other pitfalls that can lead, e.g., to inconsistent data [Gin03].

By contrast, if certain precision bounds and guarantees on the relation of the frequencies of sender and receiver can be guaranteed, metastability-free communication can be implemented. We distinguish the following guaranties, based on the definitions from [Mes90; TGL07]:

- A *mesochronous* pair of clock signals has the same frequency but a constant-value phase difference. This can, e.g., be the result of a common clock signal where the phase differences originate in the delays of the CDNs connecting the islands.

- Clock signals in a *multi-synchronous* relation also share the same frequency, but with a bounded phase difference that can drift over time.
- Clock signals in a *plesiochronous* relation have the same nominal frequency, but may suffer from a slight mismatch in frequency that leads to an unbounded phase difference.
- If none of the above applies, a *heterosynchronous* frequency relation exists, i.e., unrelated frequencies of independent oscillators. However, as [SS01] outlined in Section 2.5.3 reveals, it is not uncommon for some islands of a GALS to be clocked by the same base clock. Such approaches implement *ratiochronous* relations, where the frequencies are multiples of a common base clock and thus lead to a predictable periodic phase relation.

In general, however, communication in systems with such frequency relations is typically considered in the context of GALS.

2.6.1 Communication in GALS

Globally asynchronous, locally synchronous (GALS) architectures define systems consisting of synchronous islands that communicate asynchronously. GALS are quite common in *System-on-Chips* (SoCs), which are called *Network-on-Chips* (NoCs) if based on an underlying communication subsystem. Typically, instead of one global CDN, this approach utilizes multiple smaller CDNs [FKG+12] for each island. While, in general, no relation between the synchronous islands can be assumed, typically relations like the ones introduced before exist, hence allowing a more synchronous approach for communication.

A well-known method for asynchronous communication in general has been introduced by Sutherland under the term *micropipeline* [Sut89]. A micropipeline consists of Muller C-gates that build the control elements for the pipeline stages, similar to a FIFO buffer, where data is stored in a handshake-based fashion. Inevitably, however, at the interface between a synchronous island and the asynchronous communication channel, metastable upsets caused by signal transitions concurring during a clock transitions can occur. As usual, adding synchronizers only reduces the chance of metastability, at the cost of performance [DGS04].

A more complex system based on asynchronous communication is presented in [SFGP09]. The scope of this approach are NoCs, with the additional possibility for usage in off-chip communication, e.g., via TSVs in 3D chips. Even though the effects of faults have been analyzed, with a focus on glitches, the system can be deadlocked by faults.

In [KGGV07; TGL07], three methods for ensuring safe communication in GALS are presented. Besides the aforementioned usage of synchronizes, pausible clocks can be

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

used, which have already been considered in [Cha84]. The idea of pausible clocks is to pause the generation of the local clock if the risk of a metastable event arises. This, however, requires a local clock generation that can be influenced, i.e., paused, by the circuit itself. Communication primitives utilizing pausible clocks can pause their clock also for flow control reasons, i.e., to prevent a buffer overflow [MTMR02; SM00].

The last method listed in [TGL07] is, of course, using a priori synchronization guarantees between the islands as introduced above. A major advantage of such GALS systems is that they allow to build a global notion of time throughout an IC. Not only does this greatly simplify the design of the system at higher levels of abstractions, it also allows to avoid metastability completely. For example, the NoC architecture Aelite [HSG09] supports time-triggered transmission scheduling/routing, hence avoiding flow control to prevent buffer overruns.

Given a mesochronous, or multi-synchronous, frequency relation, which guarantees bounded precision, a high bandwidth communication can be implemented using FIFO buffers. In [PHS09], Polzer, Handl, and Steininger introduced a model that allows to size a buffer such that metastability-free communication is ensured, as long as the clocks adhere to their mesochronous properties. While this approach is not fault-tolerant on its own, it can nonetheless be utilized in Byzantine fault-tolerant architectures, in the sense that it can guarantee correct communication between non-faulty processes. If a fault affects the communication system by, e.g., corruption the read/writer pointer, or the underlying clock, some state recovery is required to resume correct operation.

In [SG03], synchronizers are introduced that can handle issues arising in mesochronous and multi-synchronous systems. It is also shown experimentally that noise from neighboring clocks due to crosstalk negatively influences the synchronizers, thus decreasing their mean time between upsets drastically.

There are multiple other approaches to achieve communication over FIFO buffers. They commonly use special methods to prevent over- respectively underruns of the buffer via control signals. However, these control signals depend on data modified by both clock domains to calculate the fill level of the FIFO. Obviously, this inter-clock domain calculation can lead to metastable upsets. Depending on the approach, either synchronizers are used to reduce the probability of metastability [BV06], or assumptions on the clocking system are employed to prevent metastability all together [CD03; CG03; CH09; CN04; DGS04]. In [MG07], both methods for metastability reduction have been considered.

2.6.2 Multi-Hop Communication

Communication between pairs of distinct components can be used to implement multi-hop communication between arbitrary components.

In large scale systems, like NoCs, dedicated communication networks are implemented that allow the communication between every connected component. NoC communication networks typically consist of routers, to which a local component may be connected, and a grid network connecting these routers. There is a rich literature on the specific details of implementing the routers [BM06]. Special focus is given

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

to congestion handling and reducing the number of wires between the routers, thus reducing the area requirements of the networks. Moreover, fault-tolerant communication is paramount for NoCs [RFZJ13], due to the criticality of communication for the functionality of the system. Fault tolerance in NoCs can be achieved on many levels: From hardening the routers, using error-correction codes for the messages, over sending duplicate messages to routing the same message over independent paths. As we will not consider multi-hop communication in this thesis, we will not dive further into these issues.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

CHAPTER 3

HEX – A CLOCK DISTRIBUTION SCHEME

> Simple solutions seldom are. It takes a very unusual mind to undertake analysis of the obvious.
>
> — ALFRED NORTH WHITEHEAD

WE TACKLE the problem of fault-tolerant clock distribution by proposing an alternative way for distributing a synchronized clock signal throughout an IC. As has been examined in Section 2.5, fault-tolerance on a level sufficient to shield Byzantine faults on such a low abstraction level has not been considered yet. Hence, we present an approach, termed HEX, that is based on a sufficiently connected wiring topology, namely, a *hexagonal grid*. At each grid point, we place an (intermediate) node that controls when the clock pulses are forwarded to adjacent nodes and supplies the clock to nearby functional units, typically using a *small* local clock tree. It will turn out that HEX compares favorably to classic clock trees in most aspects. As the intermediate nodes do not simply forward pulses, an injection of a stray pulse by a SET is blocked. However, this in turn increases the complexity of the system's skew analysis: Determining the skew at a certain point in the grid requires not only an analysis of the signal propagation delays between the nodes, but also an analytic handling of the behavior of the nodes.

Following [DFL+16], we start by introducing the HEX topology and algorithm as well as the system model in Section 3.1. In Section 3.2, we provide a detailed analysis of the worst-case neighbor skew of HEX, discuss faulty nodes and their effect on the skew, and analyze the self-stabilizing properties of HEX. We present in Section 3.3 an implementation of the HEX architecture in *very high speed integrated circuit hardware description language* (VHDL) used for the simulations conducted in Section 3.4. These

TU WIEN Bibliothek Your knowledge hub

simulations are used to validate the plausibility of the analytical results. Additionally, the extensive simulation experiments conducted give a glimpse on the average-case behavior of the system. Furthermore, our custom ModelSim-based simulation and analysis framework is presented. As HEX can only forward pulses with a fairly low frequency, we present in Section 3.5 an approach to generated a higher frequency fast clock at a HEX node. In Section 3.6 we conclude this chapter.

3.1 Algorithm & Topology

We consider a set of nodes executing a pulse generation and forwarding algorithm, which communicate by message passing over a communication network whose underlying undirected communication graph is a cylindrical hexagonal grid.

Formally, the directed communication graph (V, E) of our HEX grid is defined as follows (see Figure 3.1 and Section 2.5): Letting $L \in \mathbb{N}$ denote its length and $W \in \mathbb{N}$ its width, the set of nodes V is the set of tuples $\{(\ell, i) \in [L+1] \times [W]\}$. Here, $[L+1] := \{0, \ldots, L\}$ denotes the row index set, referred to as *layers*, and $[W] := \{0, \ldots, W-1\}$ the column index set of the nodes in the grid.

For each node $(\ell, i) \in V$, $\ell \in [L+1] \setminus \{0\}$, $i \in [W]$, the edges in E define the directed links to other nodes in the grid. We call incoming and outgoing links to such neighboring nodes of the same layer, e.g., from (ℓ, i) to $(\ell, i-1 \mod W)$, intra-layer links. Whereas, links to/from nodes in different layers are called inter-layer links. In general, inter-layer links will be directed from a lower layer to a higher layer.[1]

For HEX the following links are in E: Incoming and outgoing links to neighboring nodes of the same layer, namely from (ℓ, i) to $(\ell, i-1 \mod W)$, called the left neighbor of (ℓ, i), and to $(\ell, i+1 \mod W)$, called the right neighbor (and vice versa from the left and the right neighbor to (ℓ, i)). Every node (ℓ, i) also has incoming links from $(\ell-1, i)$, called its lower left neighbor, and $(\ell-1, i+1 \mod W)$, called its lower right neighbor. Hence, if (ℓ, i) is in a layer $\ell \in [L]$, then it has outgoing links to $(\ell+1, i-1 \mod W)$, its upper left neighbor, and $(\ell+1, i)$, its upper right neighbor. Figure 3.1 depicts the structure of the resulting HEX grid and shows a node's communication channels within the grid. The neighboring nodes of node (ℓ, i) form a hexagon, hence the name HEX grid. Due to the fact that column coordinates are modulo W, the HEX grid has a cylindrical shape.

Each node of the grid runs an algorithm that can broadcast *trigger messages* (representing clock pulses) over its outgoing links, as well as receive trigger messages over its incoming links. Each fault-free link guarantees a communication delay, i.e., the time between sending and receiving a trigger message, within $[d^-, d^+] \subset (0, \infty)$, where $\varepsilon := d^+ - d^-$. Having ε without any constraint, however, could render a few worst-case constructions (based on Definition 3.2) overly conservative. This can be avoided by the additional constraint $\varepsilon \leq d^+/2$, which guarantees a property similar to the triangle

---

[1] Otherwise, the logic of the nodes would have to ensure that no stray/duplicate pulses are distributed in the grid.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

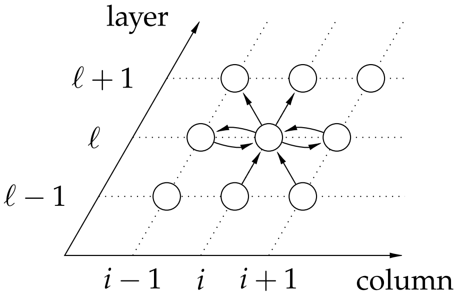


Figure 3.1: Node (ℓ, i) and its links in the cylindrical hexagonal grid topology. Column coordinates are modulo W and layer coordinates (rows) are between 0 and L.

Algorithm 3.1: HEX pulse forwarding algorithm for nodes in layer $\ell > 0$.

```
upon receiving trigger message from neighbor do
    memorize message for $\tau \in [T^-_{\mathbf{link}}, T^+_{\mathbf{link}}]$ time;
upon having memorized trigger messages from (left and lower left) or
        (lower left and lower right) or (lower right and right) neighbors do
    broadcast trigger message; // produce pulse
    sleep for $\tau \in [T^-_{\mathbf{sleep}}, T^+_{\mathbf{sleep}}]$ time;
    forget previously received trigger messages;
```

inequality. Each node further has access to a (possibly inaccurate) clock to measure timeouts.

Nodes at layer 0 are special as they act as primary clock sources, i.e., they execute a pulse generation algorithms like [DFLS14; FS12] that generates synchronized and well-separated consecutive initial trigger messages. For each pulse number $k \in \mathbb{N}$, the time between any (non-faulty) node in layer 0 generating its k^{th} trigger message and another node in layer 0 generating its $(k+1)^{\text{th}}$ trigger message is sufficiently large. The precise meaning of "sufficiently large" depends on the desired fault-tolerance properties; we will elaborate on this in Section 3.2.4. Note that it is desirable to keep the maximal time between pulses small in order to guarantee a high operating frequency.

Nodes at layers larger than 0 run the HEX pulse forwarding algorithm specified in Algorithm 3.1. Basically, nodes forward pulse k once they received trigger messages for pulse k from two adjacent neighbors. Since clock pulses and trigger messages carry no information beside their occurrence, care must be taken in order not to generate multiple trigger messages for a single pulse. The simple solution used here relies on a sufficiently large separation between the k^{th} and $(k+1)^{\text{th}}$ pulse (for each k), removing the need to locally keep track of the pulse count. Further, for each link, a node memorizes a received trigger message only for some time between $T^-_{\mathbf{link}}$ and $T^+_{\mathbf{link}}$, $T^+_{\mathbf{link}} \geq T^-_{\mathbf{link}}$, where the slack $T^+_{\mathbf{link}} - T^-_{\mathbf{link}}$ accounts for inaccurate local timers. After the timer expired, the node forgets the reception of the message by clearing the *memory*

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

flag associated with the link. Following the local generation resp. forwarding of a pulse, a node goes to sleep, i.e., will not locally trigger further pulses, for some time between $T^-_{\mathbf{sleep}}$ and $T^+_{\mathbf{sleep}} \geq T^-_{\mathbf{sleep}}$. Upon waking up, it clears all its memory flags. Note that there would be no need for the individual link timeout mechanism $\left([T^-_{\mathbf{link}}, T^+_{\mathbf{link}}]\right)$ described above if the algorithm would always start from a properly initialized state. It is required, however, for also guaranteeing self-stabilization from arbitrary states in the presence of persistent Byzantine faults.

The precise conditions for $T^-_{\mathbf{link}}$ and $T^-_{\mathbf{sleep}}$ follow from the analysis and are discussed in Section 3.2.4. Due to its simplicity, Algorithm 3.1 can easily be implemented by means of an asynchronous state machine, as described in Section 3.3.

3.2 Skew & Resilience Analysis

In this section, we analyze skew and fault-tolerance properties of the HEX algorithm in the topology presented in the previous section. Recall that nodes in layer 0 generate synchronized pulses, which the nodes in higher layers just propagate upwards; this results in a "pulse wave" as depicted in Figure 3.22. By $t^k_{\ell,i}$, we denote the *triggering time* of node (ℓ, i), i.e., the time when it forwards the k^{th} pulse of the grid. Generally, we will use superscript $^{\text{k}}$ to denote variables associated with the k^{th} pulse.

3.2.1 The Fault-free Case

We will now analyze the propagation of a single pulse wave,[2] assuming that the constraints (C1) and (C2) below are satisfied and no nodes are faulty. In a nutshell, our constraints ensure that, initially, all nodes have cleared their memory flags and are waiting for the next pulse generated by the nodes in layer 0.

(C1) $T^-_{\mathbf{link}}$ is sufficiently large so that no trigger message from a neighbor is "forgotten" before the corresponding message from another neighbor arrives. Thus we can be sure that a node which is not sleeping will be triggered by a wave.

(C2) $T^-_{\mathbf{sleep}}$, $T^+_{\mathbf{sleep}}$, and the time between pulses (controlled by the layer 0 nodes) are large enough so that (i) every node will be triggered at most once per wave and (ii) no node sleeps when the next wave arrives.

Specific values for the parameters that ensure (C1) and (C2) will be given in Section 3.2.4.

We first introduce the concept of "left zig-zag paths", which will play an essential role in bounding the worst-case triggering times of adjacent nodes in a single pulse wave.

Definition 3.1 (Causal Links and Paths):
A node is *left-triggered* / *centrally triggered* / *right-triggered* in a given execution, if the satisfied guard from Algorithm 3.1 causing the node to trigger is having received trigger

---

[2] To keep the notation simple, we will drop the superscript ^{k} indicating the pulse number during this section.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

messages from the *left and lower left* / *lower left and lower right* / *lower right and right neighbors*, respectively. In each case both of the respective links are *causal*. A *causal path* consists of causal links only. □

Note that a link being causal implies that its endpoint is triggered at least d^- time after its origin. For instance, if (ℓ, i) is left-triggered, the links $((\ell, i-1), (\ell, i))$ and $((\ell-1, i), (\ell, i))$ are causal, while $((\ell, i+1), (\ell, i))$, $((\ell-1, i+1), (\ell, i))$ are not.

The following definition backtraces a sequence of causal links from a given destination node, either to the node in layer 0 starting the causal chain or to some specific column of interest. Note that this can be done for any destination node in a given execution.

Definition 3.2 (Left Zig-Zag Paths):
Given are a layer $\ell \in [L+1] \setminus \{0\}$ and column indices $i, i' \in [W]$, $i < i'$.[3] The causal *left zig-zag path* $lzp^{i' \to (\ell,i)}$ is composed of rightward links $((\ell', j-1), (\ell', j))$ and up-left links $((\ell'-1, j+1), (\ell', j))$. It is inductively defined as follows. We start with the 0-length path $((\ell, i))$. Suppose that in some step of the construction the current path originates at node (ℓ', j) with $\ell' > 0$. If (ℓ', j) is left-triggered, we extend the path by adding the rightward link $((\ell', j-1), (\ell', j))$ as first link (and $(\ell', j-1)$ as its origin). Otherwise, the up-left link $((\ell'-1, j+1), (\ell', j))$ is causal and can be added as prefix to the path (and $(\ell'-1, j+1)$ as its origin). In the case of adding an up-left link the construction terminates if either (i) $j+1 = i'$ and the path now contains more up-left than rightward links (we will call $lzp^{i' \to (\ell,i)}$ a *triangular path* in this case) or (ii) $\ell' - 1 = 0$ and $j+1$ arbitrary (a non-triangular path). □

The following simple facts about left zig-zag paths follow almost immediately from their definition.

Lemma 3.3: *Every left zig-zag path* $lzp^{i' \to (\ell,i)}$ *constructed according to Definition 3.2 is finite. If* $lzp^{i' \to (\ell,i)}$ *is a triangular path and starts at* (ℓ', i')*, for some* $0 \leq \ell' < \ell$*, then each of its prefixes* α *is also a triangular path.* □

Proof: Since causal paths are acyclic, there must be fewer than W left links before the construction goes down one layer; the finiteness of ℓ hence implies the finiteness of $lzp^{i' \to (\ell,i)}$. Now assume that some prefix α of a triangular path $lzp^{i' \to (\ell,i)}$ starting at (ℓ', i') is not a triangular path, i.e., has at least as many rightward links as up-left ones. Then, the suffix of α must start in (ℓ'', i'') with $i'' \geq i'$, and must have more up-left links than rightward ones. Since the suffix must hence cross column i' from right to left, the construction of Definition 3.2 would already have terminated here. ■

---

[3] Recall that column indices are implicitly taken $\bmod W$, so in principle we would have to account for this in the definition. However, in our proofs we will exploit the translation and mirror symmetry of the grid w.r.t. column indices so that we can always assume that $i < i'$, avoiding more involved notations. For example, if $i = W-1$, then some node in the column to the right (which has column index 0) is considered as having index $i' = W$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

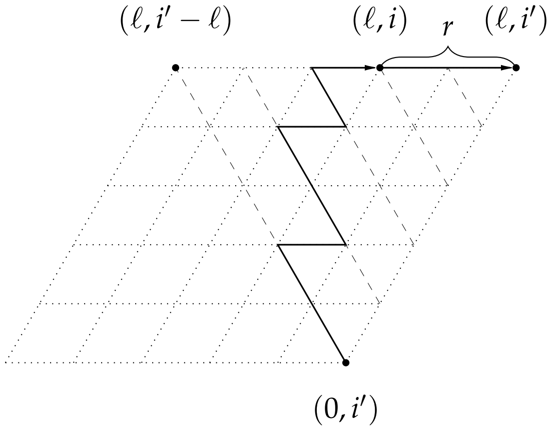


Figure 3.2: Illustration of the situation in Lemma 3.4.

We now provide a very important technical lemma, which reveals a connection between the triggering times of two nodes (ℓ, i) and $(\ell, i+1)$ at the same layer ℓ: If the left node is the end of a left zig-zag triangular path starting at node (ℓ', i') and has a distance of $r > 0$ columns to the right node, the latter cannot trigger later than the left node plus a time offset of at most $rd^- + (\ell - \ell')\varepsilon$. Note that we assume here that $i' - \ell \geq 0$, i.e., that W is large enough such that no wrap-around occurs. The bound provided by Lemma 3.4 also holds in the general case, but then it may not be tight.

Lemma 3.4: *Suppose that path α is a prefix of some left zig-zag triangular path $lzp^{i' \to (\ell'', i'')}$, and that α starts at node (ℓ', i') and ends at node (ℓ, i) with $\ell > 0$. Let $r > 0$ be the number of up-left links minus the number of rightward links along α. Then $t_{\ell,i'} \leq t_{\ell,i} + rd^- + (\ell - \ell')\varepsilon$.* □

Proof: By Lemma 3.3, α is a triangular path and hence indeed $r > 0$. For simplicity of our arguments, we set $\ell' = 0$, i.e., we shift all layer indices by ℓ' and the new value of ℓ now represents $\ell - \ell'$, and assume that $i' - \ell \geq 0$, i.e., that W is large enough such that no wrap-around occurs within the triangle.[4] Consequently, we only need to consider the set S of nodes in the triangle with corners $(0, i')$, $(\ell, i' - \ell)$, and (ℓ, i') shown in Figure 3.2.

Observe that $lzp^{i' \to (\ell'', i'')}$ starts at the lower corner of the triangle and the prefix α never leaves it. By induction on the k^{th} left-diagonal $(k, i'), \dots, (\ell, i' - (\ell - k))$ (for $k \in [\ell + 1]$) of the triangle, we will prove that each node p that is both on the diagonal k and either on α or to the right of α is triggered at the latest at time

$$t_p \leq t_{\ell,i} - (\ell - r)d^- + kd^+. \tag{3.1}$$

---

[4] Our proof also holds for the general case, though, provided (i) one just neglects the fact that some of the index pairs may actually refer to the same node (which does not affect our argument) and that (ii) certain left zig-zag paths in our construction cannot occur (which may lead to an overly conservative bound).

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Since (ℓ, i') is on diagonal ℓ, this implies $t_{\ell,i'} \leq t_{\ell,i} + rd^- + \ell\varepsilon$. Undoing the initial index shift, i.e., replacing ℓ by $\ell - \ell'$, the claim of the lemma follows.

First, we show directly that Equation (3.1) holds for each node p on α: Observe that node (ℓ, i) is on diagonal $(\ell - r)$. Hence, a node p that is h hops from (ℓ, i) on α must be on a diagonal $k \geq (\ell - r) - h$. Since $lzp^{i' \to (\ell'', i'')}$ is causal, it follows that

$$t_p \leq t_{\ell,i} - hd^- \leq t_{\ell,i} - (\ell - r)d^- + kd^- \leq t_{\ell,i} - (\ell - r)d^- + kd^+, \tag{3.2}$$

showing the statement for nodes on α.

Note that all nodes on diagonal 0 are either on or to the left of α, hence we already covered the induction anchor at $k = 0$. For the induction step from k to $k + 1$, observe that any node will be left-triggered within at most d^+ time once both its left and lower-left neighbors are triggered. For any node p on the $(k+1)^{\text{th}}$ diagonal that is strictly to the right of α, its left and lower-left neighbor are on the diagonal k of S and either on α or to the right of α. The statement for diagonal k thus implies $t_p \leq t_{\ell,i} - (\ell - r)d^- + kd^+ + d^+$ as required. On the other hand, nodes lying on α are covered by Equation (3.2), which completes the induction step. ■

In the following definition, we will introduce the different notions related to the skew between nodes. Besides the maximum (unsigned) intra-layer skew of neighboring nodes at the same layer and the (signed) maximum inter-layer skew w.r.t. the layer below, we also define the *skew potential* of layer ℓ. Informally, the latter provides a measure for the adversary's ability to exploit the existing skew of the nodes in layer ℓ to increase the skew of neighboring nodes in layer $\ell + 1$. By this, we mean that the adversary can, in the worst case, force a node at layer $\ell + 1$ to left-trigger strictly before it is centrally triggered by its layer-ℓ neighbors. It is not too difficult to prove[5] that this is only possible if the skew between neighbors at layer ℓ is strictly larger than d^-. Hence, we define the skew potential below in a way that results in a positive value only in the latter case.

Definition 3.5 (Distance, Skew, and Skew Potential):
For $i, j \in \mathbb{Z}$, let $d := i - j \mod W$ and define the cyclic distance as $|i - j|_W := \min\{d, W - d\}$. For $\ell \in [L + 1]$, we define

(i) the *intra-layer skew* of layer ℓ as $\sigma_\ell := \max_{i \in [W]} \{|t_{\ell,i} - t_{\ell,i+1}|\}$,

(ii) the *skew potential* on layer ℓ as $\Delta_\ell := \max_{i,j \in [W]} \{t_{\ell,i} - t_{\ell,j} - |i - j|_W d^-\}$.

For $\ell \in [L + 1] \setminus \{0\}$, we define

(iii) the *inter-layer skew* of layer ℓ as $\hat{\sigma}_\ell := \max_{i \in [W]} \{t_{\ell,i} - t_{\ell-1,i}, t_{\ell,i} - t_{\ell-1,i+1}\}$. □

---

[5] In fact, this proof is embedded in the proof of Lemma 3.7. In a nutshell, it shows that such an early left-triggering would only be possible if the left neighbor itself was early left-triggered as well. By continuing this argument inductively over the entire layer, a left neighbor will eventually be reached that cannot be left-triggered, which provides the required contradiction.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

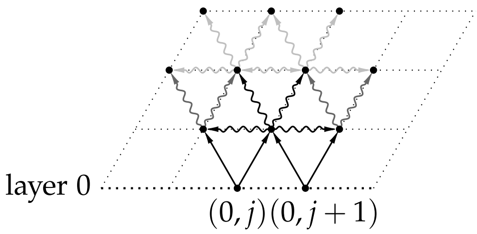


Figure 3.3: Illustration of the induction in the proof of Case 2 of Lemma 3.6. The solid lines mark the trigger messages sent by time t_0, the wiggly black lines symbolize the trigger messages that will be sent by time $t_0 + d^+$. The gray wiggly lines resp. the light gray ones represent (some) trigger messages that will be sent by time $t_0 + 2d^+$ resp. $t_0 + 3d^+$.

Note that every pair of nodes i, j occurs twice (as i, j and j, i) in the max-term of the skew potential in (ii) above, which implies that only a non-negative time difference can determine Δ_ℓ. Moreover, as $j = i$ is not excluded, we always have $\Delta_\ell \geq 0$.

We start by proving a weak bound on the maximal skew at the upper layers that holds *independently* of the initial skew potential Δ_0. Note that this result implies tolerance of HEX against arbitrary layer 0 skews, at the expense of “losing” layers $\ell \in [W-2]$. This behavior is also clearly visible in the simulation results shown in Figures 3.23 and 3.28.

Lemma 3.6: *For $W > 2$ and all $\ell \in \{W-2, \dots, L\}$, $\Delta_\ell \leq 2(W-2)\varepsilon$.* □

Proof: Consider any fixed $i, i' \in [W]$, $i < i'$ (wrap-around cases are symmetrical) and assume that $\ell = W - 2$; we will argue later on why the proof below also covers $\ell > W - 2$. We distinguish two cases.

Case 1: $lzp^{i' \to (\ell,i)}$ starts at node (ℓ', i') for some $\ell' \in \{1, \dots, \ell - 1\}$
Then, by Lemma 3.4,

$$t_{\ell,i'} \leq t_{\ell,i} + (i'-i)d^- + (\ell-\ell')\varepsilon < t_{\ell,i} + (i'-i)d^- + \ell\varepsilon. \leq t_{\ell,i} + (i'-i)d^- + (W-2)\varepsilon. \tag{3.3}$$

Case 2: $lzp^{i' \to (\ell,i)}$ starts at node $(0, j)$, $j \in [W]$
Then the path has a length of at least $2\ell - (i'-i)$, since at least ℓ up-left and $\ell - (i'-i)$ right links are required for the path to originate at layer 0. Denote by t_0 the earliest time when a pair of two adjacent nodes in layer 0 are both triggered. Clearly, the k^{th} node on $lzp^{i' \to (\ell,i)}$, $k \geq 2$, cannot be triggered before time $t_0 + (k-1)d^-$ because it is connected by a causal path of length $k-1$ to a layer-0 node that is triggered at or after time t_0. Hence, $t_{\ell,i} \geq t_0 + (2\ell - (i'-i))d^-$ and thus

$$t_{\ell,i} \geq t_0 + (2(W-2) - (i'-i))d^-. \tag{3.4}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Denote by $(0, j)$ a node with $\max\{t_{0,j}, t_{0,j+1}\} = t_0$; by the definition of t_0, such a node exists. We claim that all nodes in layer $W-2$ are triggered no later than time $t_0 + 2(W-2)d^+$. This follows by induction on the layers $\lambda \in [W-1]$, where the hypothesis is that all nodes $(\lambda, j-\lambda), (\lambda, j-\lambda+1), \ldots, (\lambda, j+1)$ are triggered until time $t_0 + 2\lambda d^+$; an illustration of the first layers is shown in Figure 3.3. Since in layer λ these are $2+\lambda$ nodes, i.e., all W nodes in layer $W-2$, this will prove the claim of our lemma.

By the definition of t_0, the induction hypothesis holds for $\lambda = 0$. To perform the step from λ to $\lambda+1$, observe that all nodes $(\lambda+1, j-\lambda), (\lambda+1, j-\lambda+1), \ldots, (\lambda+1, j)$ are triggered no later than time $t_0 + (2\lambda+1)d^+$, since by the hypothesis their lower left and lower right neighbors are triggered at least d^+ before that time. Until time $t_0 + 2(\lambda+1)d^+$, nodes $(\lambda+1, j-(\lambda+1))$ resp. $(\lambda+1, j+1)$ must also follow since they are right- resp. left-triggered (if not triggered differently before), which completes the induction.

The result of our induction proof implies $t_{\ell,i'} \le t_0 + 2(W-2)d^+$ and hence, by using Equation (3.4),

$$t_{\ell,i'} - t_{\ell,i} \le (i'-i)d^- + 2(W-2)\varepsilon. \tag{3.5}$$

Overall, since i and $i' > i$ were arbitrary, from the two cases and the symmetry properties of the grid, we conclude that $\Delta_\ell = \max_{i,i' \in [W]} \left\{ t_{\ell,i'} - t_{\ell,i} - |i'-i|_W d^- \right\} \le 2(W-2)\varepsilon$, as claimed.

Finally, since the above proof did not require any specific property to be respected by layer 0 nodes, it also applies literally if we replace layer $W-2$ by ℓ and layer 0 by layer $\ell - W + 2$. This concludes the proof of Lemma 3.6. ■

Next, we derive more refined bounds on the intra-layer skew between two neighboring nodes at the same layer $\ell > 0$: In contrast to Lemma 3.6, we now take the maximal skew in previous layers into account. Unfortunately, its proof is complicated by the fact that we need to distinguish three different cases that might lead to the worst-cast skew $\vec{\sigma}_{\ell,i}$ between nodes (ℓ, i) and $(\ell, i+1)$. They depend on whether the skew $\vec{\sigma}_{\lambda,i}$ of the corresponding nodes (λ, i) and $(\lambda, i+1)$ at some (suitably chosen) layer λ is $\vec{\sigma}_{\lambda,i} \le d^+$ (Case 1) or else $\vec{\sigma}_{\lambda,i} > d^+$ (Cases 2 and 3).

Informally, Case 1 is characterized by a V-shaped growth of the worst-case skew: Our detailed proof will show that $\vec{\sigma}_{\lambda,i}$ increases a most by ε with every layer. By contrast, in the other cases, the nodes (λ, i) and $(\lambda, i+1)$ at layer λ are already “torn apart”. The worst-case skew in this case is determined by a left zig-zag path $lzp^{i+1\to(\ell,i)}$ that causes $t_{\ell,i}$ to be as small as possible on the one hand, and a “slow” causal path ending at $(\lambda, i+1)$ that makes $t_{\ell,i+1}$ as large as possible on the other hand. Cases 2 and 3 are distinguished according to the two possibilities where $lzp^{i+1\to(\ell,i)}$ can start here: Case 2 applies when it originates at some node $(0, j_0)$ with $j_0 \ne i+1$, whereas Case 3 is characterized by a triangular path starting at $(\ell', i+1)$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Lemma 3.7: *For all* $\ell_0 \in [L]$ *and* $\ell \in \{\ell_0+1,\dots,L\}$*, it holds for each* $i \in [W]$ *that*

$$|t_{\ell,i} - t_{\ell,i+1}| \leq d^+ + \left\lceil \frac{(\ell-\ell_0)\varepsilon}{d^+} \right\rceil \varepsilon + \Delta_{\ell_0}. \qquad \square$$

Proof: Fix some value of $\ell \geq 1$ and assume, without loss of generality, that $\ell_0 = 0$. To simplify our arguments, we also assume $t_{\ell,i} < t_{\ell,i+1}$ (as the other cases are symmetric, this is sufficient).

Define $\lambda_0 := \lfloor \ell d^- / d^+ \rfloor$, which maximizes λ_0 under the constraint that $\lambda_0 d^+ \leq \ell d^-$. Thus, a "slow" chain of trigger messages will complete λ_0 hops within the time a "fast" chain requires for ℓ hops. Recalling that $-\lfloor x \rfloor = \lceil -x \rceil$, we obtain

$$\ell - \lambda_0 = \ell - \left\lfloor \frac{\ell d^-}{d^+} \right\rfloor = \left\lceil \frac{\ell \varepsilon}{d^+} \right\rceil. \tag{3.6}$$

We distinguish three cases:

Case 1:
V-shaped skews (Figure 3.4a): $t_{\lambda,i+1} \leq t_{\lambda,i} + d^+$ for some $\lambda \geq \lambda_0$.

We choose λ maximal with this property, so that $t_{\lambda',i+1} > t_{\lambda',i} + d^+$ for all $\lambda' \in \{\lambda+1,\dots,\ell\}$. Notice that this implies that, for all such λ', node (λ', i) cannot be right-triggered, as the links $((\lambda', i+1), (\lambda', i))$ cannot be causal. Hence, all links $((\lambda'-1, i), (\lambda', i))$ must be causal in this case. By induction on λ', we can thus infer $t_{\ell,i} \geq t_{\lambda,i} + (\ell-\lambda)d^-$.

Furthermore, $t_{\lambda',i+1} > t_{\lambda',i} + d^+$ ensures that the trigger message from (λ', i) to $(\lambda', i+1)$ arrives well before time $t_{\lambda',i+1}$. Thus, node $(\lambda', i+1)$ will be triggered at the latest when the trigger message from its lower left neighbor $(\lambda'-1, i+1)$ arrives. Again by induction on λ', we infer that $t_{\ell,i+1} \leq t_{\lambda,i+1} + (\ell-\lambda)d^+$, and hence

$$t_{\ell,i+1} \leq t_{\lambda,i} + (\ell-\lambda+1)d^+. \tag{3.7}$$

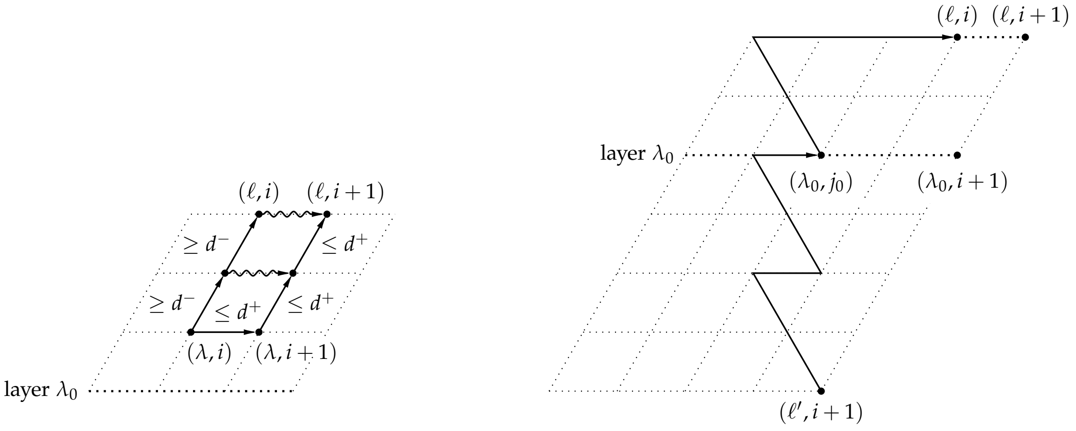
Combining these bounds and applying Equation (3.6), we obtain

$$t_{\ell,i+1} - t_{\ell,i} \leq (\ell-\lambda)\varepsilon + d^+ \leq d^+ + \left\lceil \frac{\ell\varepsilon}{d^+} \right\rceil \varepsilon. \tag{3.8}$$

Case 2:
Non-V-shaped skews, non-triangular (Figure 3.4c): Case 1 does not apply and the $lzp^{i+1\to(\ell,i)}$ starts at some node $(0, j_0)$, for $j_0 \neq i+1$.

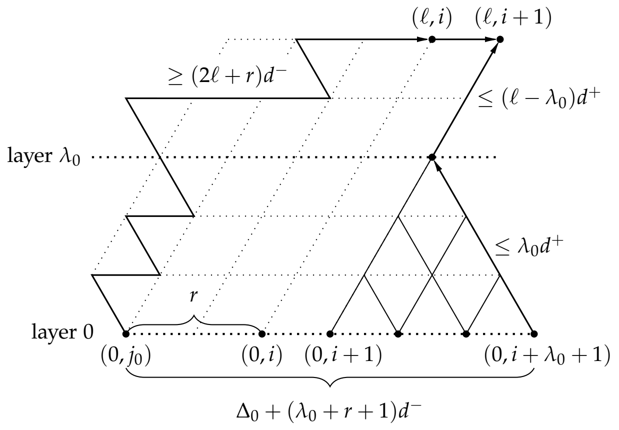
If $lzp^{i+1\to(\ell,i)}$ contained more left-up links than rightward links, it would contain a subpath originating at a node in column $i+1$ that also would have more left-up than rightward links. This is not possible, since then the construction would have terminated at this node, either resulting in the path originating at a layer $\ell' > 0$ or at node $(0, i+1)$. Hence $lzp^{i+1\to(\ell,i)}$ is of length $2\ell + r$ for some $r \geq 0$ and $j_0 = i - r \mod W$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) Case 1: V-shaped skews

(b) Case 3: Non-V-shaped skews (triangular)



(c) Case 2: Non-V-shaped skews

Figure 3.4: Illustrations of the different cases in the proof of Lemma 3.7.

For all indices $j \in \{i+1, i+2, \ldots, i+1+\lambda_0\}$, we have that $|j - j_0|_W \leq j - i + r$, and hence, by definition of the skew potential, also

$$\begin{aligned} t_{0,j} - t_{0,j_0} = t_{0,j} - t_{0,j_0} - |j - j_0|_W d^- + |j - j_0|_W d^- \leq \Delta_0 + |j - j_0|_W d^- \leq \\ \leq \Delta_0 + (\lambda_0 + r + 1)d^-. \end{aligned} \tag{3.9}$$

Recalling the length of $lzp^{i+1\rightarrow(\ell,i)}$ established above, we obtain that

$$\begin{aligned} t_{\ell,i} &\geq t_{0,j_0} + (2\ell + r)d^- \\ &\geq t_{0,j} - \Delta_0 - (\lambda_0 + r + 1)d^- + (2\ell + r)d^- \\ &= t_{0,j} - \Delta_0 + (2\ell - \lambda_0 - 1)d^-. \end{aligned} \tag{3.10}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Moreover, by induction on $\lambda \in \{0, \dots, \lambda_0\}$, it follows that all nodes $(\lambda, j') \in \{(\lambda, i+1), \dots, (\lambda, (i+1+\lambda_0-\lambda) \mod W)\}$ are triggered at time $t_{\lambda,j'} \le \max_{i<j\le i+\lambda_0+1}\{t_{0,j}\} + \lambda d^+$. Plugging in Proof 4 implies

$$t_{\lambda,j'} \le t_{\ell,i} + \Delta_0 - (2\ell - \lambda_0 - 1)d^- + \lambda d^+ \tag{3.11}$$

and hence

$$t_{\lambda_0,i+1} \le t_{\ell,i} + \Delta_0 - (2\ell - \lambda_0 - 1)d^- + \lambda_0 d^+ \le t_{\ell,i} + \Delta_0 - (\ell - \lambda_0 - 1)d^-; \tag{3.12}$$

the second inequality holds by the definition of λ_0, which implies $\lambda_0 d^+ \le \ell d^-$.

Since Case 1 does not apply, we have $t_{\lambda,i+1} > t_{\lambda,i} + d^+$ for all $\lambda_0 \le \lambda \le \ell$. We can hence use the same argument as used for deriving Equation (3.7) to show that $t_{\ell,i+1} \le t_{\lambda_0,i+1} + (\ell - \lambda_0)d^+$. It follows that

$$t_{\ell,i+1} \le t_{\lambda_0,i+1} + (\ell - \lambda_0)d^+ \le t_{\ell,i} + d^- + (\ell - \lambda_0)\varepsilon + \Delta_0 = t_{\ell,i} + d^- + \left\lceil \frac{\ell\varepsilon}{d^+} \right\rceil \varepsilon + \Delta_0, \tag{3.13}$$

where the last equality follows from Equation (3.6).

Case 3:
Non-V-shaped skews, triangular (Figure 3.4b): Neither Case 1 nor Case 2 apply.

In this case, $t_{\lambda,i+1} > t_{\lambda,i} + d^+$ for all $\lambda \in \{\lambda_0, \dots, \ell\}$, and $lzp^{i+1\to(\ell,i)}$ is a triangular path starting at node $(\ell', i+1)$ for some $\ell' < \lambda_0 - 1$: By construction, the first (causal) link of $lzp^{i+1\to(\ell,i)}$ is $((\ell', i+1), (\ell'+1, i))$, implying that node $(\ell'+1, i+1)$ is triggered no later than time $t_{\ell'+1,i} + d^+ = \max\{t_{\ell',i+1} + d^+, t_{\ell'+1,i} + d^+\}$. Hence, since Case 1 does not apply, we must indeed have $\ell' + 1 < \lambda_0$.

Let (λ_0, j_0) be the last node on the causal path $lzp^{i+1\to(\ell,i)}$ that is still in layer λ_0. Observe that $j_0 + r - u = i$, where r (resp. u) is the number of rightward (resp. up-left) hops of $lzp^{i+1\to(\ell,i)}$ after (λ_0, j_0). We apply Lemma 3.4 to the prefix α of $lzp^{i+1\to(\ell,i)}$ ending at (λ_0, j_0), i.e., set $i := j_0$, $i' = i+1$ and $r := i+1-j_0$ in this lemma, which yields

$$t_{\lambda_0,i+1} \le t_{\lambda_0,j_0} + (i+1-j_0)d^- + \left(\lambda_0 - \ell'\right)\varepsilon. \tag{3.14}$$

Since Case 1 does not apply, we can use the same induction as used before Equation (3.7) to prove that $t_{\ell,i+1} \le t_{\lambda_0,i+1} + (\ell - \lambda_0)d^+$. We thus obtain

$$\begin{aligned} t_{\ell,i+1} &\le t_{\lambda_0,j_0} + (i+1-j_0)d^- + \left(\lambda_0 - \ell'\right)\varepsilon + (\ell - \lambda_0)d^+ \\ &= t_{\lambda_0,j_0} + (\ell - \lambda_0 + i + 1 - j_0)d^- + \left(\ell - \ell'\right)\varepsilon. \end{aligned} \tag{3.15}$$

By construction, $lzp^{i+1\to(\ell,i)}$ is of length $2(\ell - \ell') - 1$ and its prefix ending at node (λ_0, j_0) is of length $2(\lambda_0 - \ell') - (i+1-j_0)$. Therefore, the length of the suffix of $lzp^{i+1\to(\ell,i)}$ starting at (λ_0, j_0) is $2(\ell - \lambda_0) + (i - j_0)$. As this suffix is a causal path, we have

$$t_{\ell,i} \ge t_{\lambda_0,j_0} + (2(\ell - \lambda_0) + (i - j_0))d^-. \tag{3.16}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

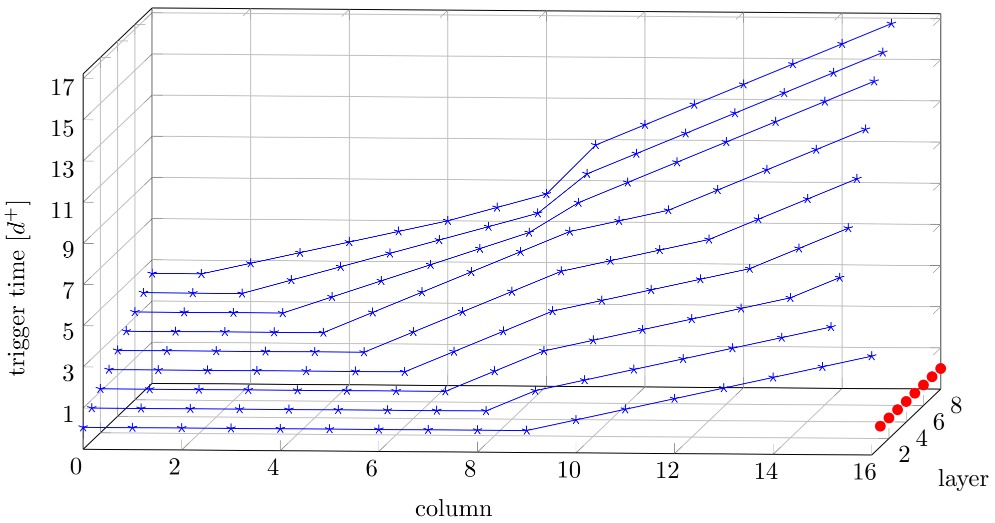


Figure 3.5: Visualization of a worst-case pulse propagation wave, maximizing the skew between the top-layer nodes in columns 8 and 9. To focus on the essential central part of the grid, we introduced a barrier of “dead” nodes in column 16. Nodes in and left of column 8 are left-triggered (except for the “flat” region) with minimal delays of d^-. Nodes in and right of column 9 are slow due to large delays of d^+ and large initial skews in parts of layer 0.

Altogether, we arrive at

$$\begin{aligned} t_{\ell,i+1} - t_{\ell,i} &\leq (\ell - \ell')\varepsilon - (\ell - \lambda_0 - 1)d^- \\ &\leq \ell\varepsilon - \left(\frac{\ell\varepsilon}{d^+} - 1\right) d^- \\ &= d^- + \frac{\ell\varepsilon^2}{d^+} \\ &\leq d^+ + \left\lceil \frac{\ell\varepsilon}{d^+} \right\rceil \varepsilon \end{aligned} \tag{3.17}$$

according to Equation (3.6).

Since the claimed bound holds in each of the (exhaustive) cases considered, the proof of Lemma 3.7 is completed. ■

We remark that it is possible to construct, by deterministically choosing appropriate link delays, worst-case executions that almost match the bounds established in Lemma 3.7; an example is shown in Figure 3.5.

In the proof of Lemma 3.7, in particular, in Case 2 (Figure 3.4c), we silently assumed that the starting node $(0, j_0)$ of the left zig-zag path $lzp^{i+1 \to (\ell,i)}$ on the left side does not “collide” (due to a wrap-around) with one of the slow nodes $(0, i+1), \ldots, (0, i+\lambda_0+1)$ on the right side. Whereas this is reasonable for wide grids, this is not

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

realistic if W is small. Considering such a collision prohibits some of the worst-case scenarios considered, and hence possibly makes the worst-case skew result provided by Lemma 3.7 overly conservative. We therefore provide the following corollary, which takes this width constraint into account.

Corollary 3.8: *Set $\delta := d^-/2 - \varepsilon$. For each layer $\ell \in \{W, \dots, L\}$ and all $i \in [W]$, it holds that*

$$|t_{\ell,i} - t_{\ell,i+1}| \leq \max\left\{d^+ + \left\lceil \frac{W\varepsilon}{d^+} \right\rceil \varepsilon, \Delta_{\ell-W} + d^+ - W\delta\right\}. \qquad \square$$

Proof: The proof is mostly analogous to the one of Lemma 3.7, with ℓ_0 resp. Δ_0 replaced by $\ell - W$ resp. $\Delta_{\ell-W}$. Case 2 needs a slightly different treatment, though, by assuming w.l.o.g. $\ell = W$ and recomputing the bound Proof 4 for all indices $j \in \{i+1, i+2, \dots, i+\lambda_0+1\}$ as

$$t_{\ell,i} \geq t_{0,j} - \Delta_0 - |j - j_0|_W d^- + (2\ell + r)d^- \geq t_{0,j} - \Delta_0 + \frac{3\ell d^-}{2}, \tag{3.18}$$

where we conservatively set $r = 0$ and exploit that $|j - j_0|_W \leq W/2 = \ell/2$ in the second step. The analogon of Equation (3.11) in the proof of Lemma 3.7, for $\lambda \in \{0, \dots, \lambda_0\}$ and $(\lambda, j') \in \{(\lambda, i+1), \dots, (\lambda, (i+1+\lambda_0-\lambda) \mod W)\}$, hence reads

$$t_{\lambda,j'} \leq t_{\ell,i} + \Delta_0 - \frac{3\ell d^-}{2} + \lambda d^+ \tag{3.19}$$

and thus leads to

$$t_{\lambda_0,i+1} \leq t_{\ell,i} + \Delta_0 - \frac{3\ell d^-}{2} + \lambda_0 d^+ \leq t_{\ell,i} + \Delta_0 - \frac{\ell d^-}{2}, \tag{3.20}$$

where we used that $\lambda_0 d^+ \leq \ell d^-$ by the definition of λ_0. Finally, re-using the result $t_{\ell,i+1} \leq t_{\lambda_0,i+1} + (\ell - \lambda_0)d^+$ from the proof of Lemma 3.7, we arrive at

$$\begin{aligned} t_{\ell,i+1} &\leq t_{\ell,i} + \Delta_0 + (\ell - \lambda_0)d^+ - \frac{\ell d^-}{2} \\ &\leq t_{\ell,i} + \Delta_0 + \ell\varepsilon + d^+ - \frac{\ell d^-}{2} \\ &= t_{\ell,i} + \Delta_0 + d^+ - W\delta, \end{aligned} \tag{3.21}$$

where we used Equation (3.6) to derive the second inequality.

Checking the bounds from Case 1 and Case 3 in Lemma 3.7, we see that they are smaller or equal to the left term in the maximum on the right hand side of the claimed bound. The bound for the differently treated Case 2 matches the right term in the maximum. ■

We are now ready to derive our main result, namely, bounds on the worst-case skews between neighbors.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Wien Bibliothek Your knowledge hub

Theorem 3.9 (Skew Bounds—Fault-free Case): *Suppose that* $\varepsilon \leq d^+/7$. *If* $\Delta_0 = 0$, *then the intra-layer skew* σ_ℓ *(recall Definition 3.5) is uniformly bounded by* $d^+ + \lceil W\varepsilon/d^+ \rceil \varepsilon$ *for any* $\ell \in [L+1]$. *In the general case,*

$$\begin{aligned}\forall \ell \in \{1,\dots,2W-3\}: &\quad \sigma_\ell \leq d^+ + 2W\varepsilon^2/d^+ + \Delta_0.\\ \forall \ell \in \{2W-2,\dots,L\}: &\quad \sigma_\ell \leq d^+ + \lceil W\varepsilon/d^+ \rceil \varepsilon.\end{aligned}$$

The inter-layer skew of layer $\ell \in [L+1] \setminus \{0\}$, *for all* $i \in [W]$, *is determined by*

$$\begin{aligned} t_{\ell-1,i} - \sigma_{\ell-1} + d^- &\leq t_{\ell,i} \leq t_{\ell-1,i} + \sigma_{\ell-1} + d^+ \text{ and}\\ t_{\ell-1,i+1} - \sigma_{\ell-1} + d^- &\leq t_{\ell,i} \leq t_{\ell-1,i+1} + \sigma_{\ell-1} + d^+.\end{aligned}$$

□

Proof: Assume first that $\Delta_0 = 0$. For the sake of the argument, imagine that the HEX grid would start at layer $-(W-1)$, where for all $i \in [W]$ and all $\ell \in \{-(W-1),\dots,0\}$ we would have that $t_{\ell,i} = \ell d^+$. Clearly, starting from any execution on the actual grid, this would result in a feasible execution on the extended grid if we choose all link delays on the imagined links to be d^+. It follows that $\Delta_\ell = 0$ for all $\ell \in \{-(W-1),\dots,0\}$. From Lemma 3.6, we obtain that $\Delta_\ell \leq 2(W-2)\varepsilon$ for all $\ell \in \{1,\dots,L\}$ (since we have negative layer indices until $-(W-1)$, the lemma also applies to layers $1,\dots,W-3$). Now we apply Corollary 3.8 to all layers $\ell \in \{1,\dots,L\}$, yielding that

$$\sigma_\ell \leq \max\left\{d^+ + \left\lceil \frac{W\varepsilon}{d^+} \right\rceil \varepsilon, W(2\varepsilon - \delta) + d^+\right\}. \tag{3.22}$$

Since $\varepsilon \leq d^+/7$, we have $d^- \geq 6d^+/7$ and $\delta \geq 2d^+/7$ and thus $2\varepsilon - \delta \leq 0$; the maximum in Equation (3.22) is hence dominated by the first term. This proves the first statement.

Now consider the case where Δ_0 is arbitrary. The bound on σ_ℓ for $\ell \in \{1,\dots,2W-3\}$ follows from Lemma 3.7. For $\ell \geq 2W-2$, observe first that we can apply Lemma 3.6 to all layers $\ell \in \{W-2,\dots,L\}$. Hence the same bound as in the previous case holds due to Corollary 3.8 applied to layers $\ell \in \{2W-2,\dots,L\}$.

The third inequality of the theorem holds since

$$\begin{aligned} t_{\ell-1,i} - \sigma_{\ell-1} + d^- &\leq \min\{t_{\ell-1,i}, t_{\ell-1,i+1}\} + d^- \qquad (3.23)\\ &\leq t_{\ell,i} \leq \max\{t_{\ell-1,i}, t_{\ell-1,i+1}\} + d^+ \leq t_{\ell-1,i} + \sigma_{\ell-1} + d^+; \qquad (3.24)\end{aligned}$$

the last inequality is proved analogously. ■

3.2.2 Fault Model

As usual network properties are not provided by HEX, like (almost) fully connectivity or bidirectional communication, we cannot hope to match classic resilience results, like tolerance to f Byzantine node failures provided $n \geq 3f+1$ [Dol82; PSL80].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

A promising alternative are f–local failure models [BV05; Koo04; PP05], which assume that every node has at most f faulty nodes among its n direct neighbors. It is particularly suitable for regular topologies like HEX, and has the advantage that the network-wide number of faulty nodes scales with the network size.

To define suitable failure semantics, we first consider the possible effect of faults on the HEX grid. Hence, we elaborated the possibilities the two most basic fault types have:

Fail-Silent A fail-silent or crash faulty node can only stop sending trigger messages at some point, and never resume sending from this point on. It thus can only slow-down the firing of a neighbor, by extending the causal path to that respective neighbor.

Byzantine As a Byzantine faulty node actually has no limitations on its behavior, the implementation presented in Section 3.3, has a problem. The state-holding element, especially the memory flags inside the receiver, cannot be ensured to be invulnerable to metastability upset [Mar81] . However, as laid out in more detail in [DFL+14], the probability that a faulty node produces a signal transition within the picosecond-range window of vulnerability is extremely small and can be further decreased by means of, e.g., synchronizers [Kin08]. Thus, we must restrict Byzantine faults from inducing metastability.

It follows that a Byzantine fault is limited to

- moderately speeding-up its neighboring nodes, as two causal links are necessary for a node to trigger.
- arbitrary slow-down of its neighboring nodes. However, the trigger message could be arbitrarily delayed for pulse k, which may in turn result in an arbitrary speed-up for pulse $k+1$! Whereas, a fail-silent fault can be modeled by omission of a trigger message.

As Algorithm 3.1 is designed to withstand one Byzantine faulty neighbor, this property can be used as a first base of our fault model:

Assumption 3.10: *Every node in the HEX grid has at most one faulty neighbor.* □

Under this restriction, a faulty nodes requires that the 18 nodes with a distance of ≤ 2 hops are non-faulty, cf. Figure 3.6. This number may be smaller for narrow grids or nodes near the top or the bottom. Requiring such a large neighborhood of correct nodes is rather restrictive, as the birthday paradox[6] [AM70] causes the probability of

---

[6] In [AM70] a solution is presented for the question who likely it is that f people have their birthdays at least $k/2$ days apart. As these $k/2$ days span in both directions, they represent the size of the neighborhood of the fault in which no other fault must be placed. Unfortunately, this 1D problem does not fully match the 2D situation in our grid: While the mapping, by itself, from these $k/2$ day interval from the 1D timeline to the 2D grid would not produce problems, the 1D solution allows no overlapping of the neighborhoods of two persons which is possible in our case.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

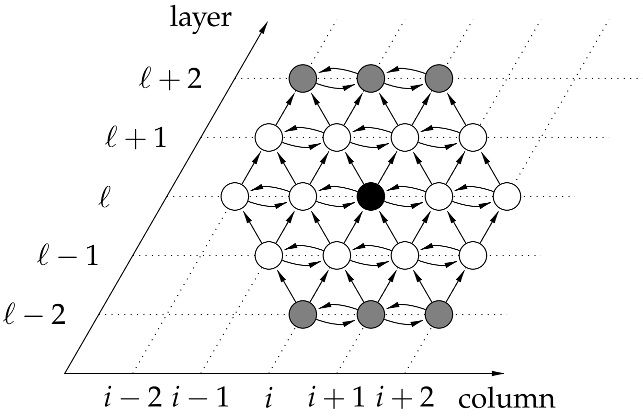


Figure 3.6: When the node (ℓ, i) is faulty, then if one of the white nodes would also be faulty, one node would have two faulty nodes on incoming edges. If one of the gray nodes would be faulty, then some nodes would have two faulty neighboring nodes.

a "correct" random placement of faulty nodes to drop fast with the number of faulty nodes.

The following formula provides a lower bound for the probability of a correct placement of f faulty nodes, which require k surrounding neighboring nodes to be non-faulty, in a grid with n nodes:

$$
\begin{aligned}
p(k,f,n) &\geq \frac{n \cdot (n-(k+1)) \cdot \ldots \cdot (n-(f-1)(k+1))}{f!\binom{n}{f}} = \\
&= \frac{\prod_{i=0}^{f-1} n-(k+1)i}{\prod_{i=0}^{f-1} n-i} > \left(1 - \frac{(k+1)(f-1)}{n}\right)^f .
\end{aligned}
\tag{3.25}
$$

Equation (3.25) is similar to a (n,f)-combination, but instead of removing only one element per selection from the set of possible elements, $k+1$ elements, the entire neighborhood of the faulty node and the faulty node itself, are removed. To get the probability, the term is divided by the $\binom{n}{f}$ possible combinations of placing f faulty nodes in n nodes.

Figure 3.7 shows a plot of Equation (3.25) with 2500 nodes, which is the grid size used in our simulations in Section 3.4. The probability for a successful placement of 15 faults in a neighborhood of size $k = 18$ is at least 45.5%, which is pretty low given that only 0.6% of the nodes are faulty.

Fortunately, Assumption 3.10 can be weakened. As the HEX grid is directed, the fault model only has to ensure that no node has two faulty nodes on its incoming links.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

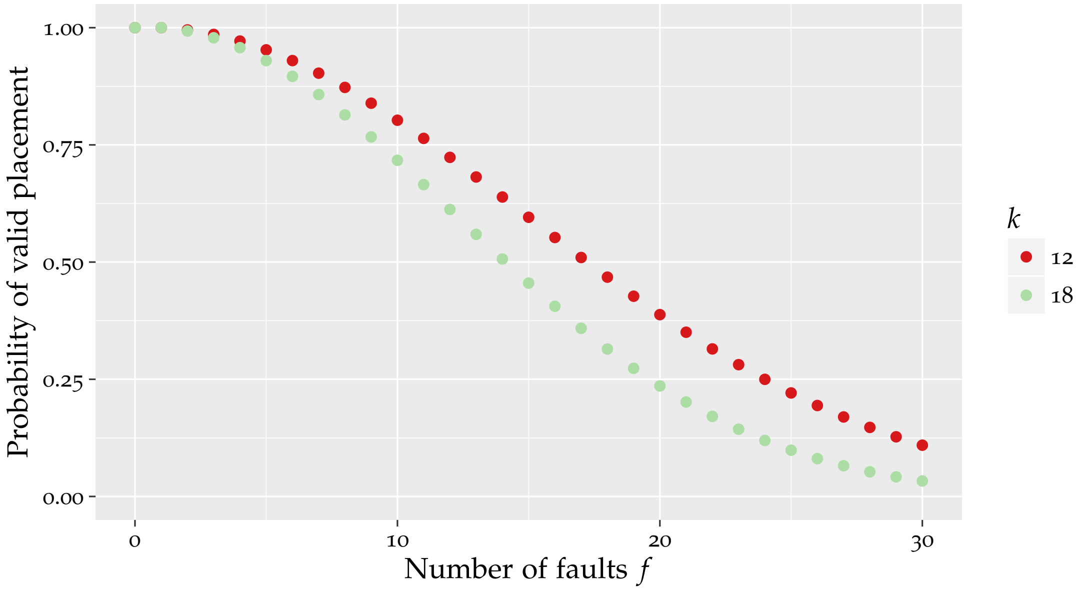


Figure 3.7: Probability of correct placement of faults under the 1–local fault model according to Equation (3.25). The lower bound was plotted with $n = 2500$ nodes and such that $k = 12$ resp. $k = 18$ direct neighbors of a faulty node must be correct.

Definition 3.11 (1–local fault model):
Every node in the HEX grid has at most one faulty neighbor on its incoming links. □

This weakened model reduces the size of the fault-free neighborhood to 12 nodes: Every in-neighbor of the 4 nodes who have a faulty node as in-neighbor must be non-faulty. With this model, the probability of a correct placement of 15 faults has increased to at least 59.5%, for $f = 20$, this probability has increased from at least 23.5% to at least 38.7%.

Clustered Faults In general, a HEX node can only tolerate one faulty in-neighbor. Nonetheless, HEX can withstand some cases where one node has multiple faulty neighbors, at the cost of an increase of the skew. Consider the case where the two intra-layer neighbors of (ℓ, i) are faulty. In this case, node (ℓ, i) would not be affected by the faults, due to the firing rules. Unfortunately, however, if both faulty nodes are fail-silent, then the nodes $(\ell + 1, i - 1)$ and $(\ell + 1, i)$ could not fire, i.e., become mute. Furthermore, this would also prevent node $(\ell + 2, i - 1)$ from firing.

A more extreme scenario would assume that the two adjacent in-neighbors of (ℓ, i) in the layer below, i.e., $(\ell - 1, i)$ and $(\ell - 1, i + 1)$, are faulty. In the case of Byzantine faults, they could force (ℓ, i) to produce spurious pulses, resulting in a spurious wave similar to the scenario shown in Figure 3.3. In the case of fail-silent nodes, the skew of $(\ell, i - 1)$ and $(\ell, i + 1)$ would possibly increase and (ℓ, i) would become mute.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

In the case of three neighboring fail-silent nodes, e.g. $(\ell-1, i-1)$, $(\ell-1, i)$ and $(\ell-1, i+1)$, the nodes $(\ell, i-1)$ and (ℓ, i) would also become mute. Yet, due to the structure of the HEX grid, this "muteness wave" will not expand, but rather vanish in the form of a triangle with increasing layers. Note that this even generalizes to more than three fail-silent neighbors in the same layer.

If the fail-silent nodes would be in adjacent layers and in the same column, then there is no way to cause other nodes to become mute. But a chain of fail-silent faults along a column would slice part of the cylindrical HEX grid open, which would, again, have a bad effect on the skews.

Finally, if we allow both neighbors, the left and the right, of a correct node to fail, we could have every second node in an entire layer failing, which would prevent the propagation of pulses if these nodes do not send messages.

3.2.3 Byzantine Faults

We now extend the analysis from the fault-free case to the case of some faulty nodes in the grid. We still confine our examination to a single pulse; we will show later that the necessary preconditions for this type of analysis will eventually be satisfied, independent of the initial states of the nodes.

With these issues listed above in mind, we arrive at the following sufficient condition for triggering all nodes exactly once per pulse.

Condition 3.12 (Fault Separation): *For each node, the 1–local fault assumption applies, hence no more than one of its incoming links connects to a faulty neighbor.* □

We use the following definition to summarize skews.

Definition 3.13:
For $\ell \in [L+1] \setminus \{0\}$, we say that *layer ℓ has skew at most σ* if, for any two correct neighbors (ℓ, i) and (ℓ', i'), $|t_{\ell,i} - t_{\ell',i'}| \leq \sigma$ with $\ell - 1 \leq \ell' \leq \ell$, and layer $\ell - 1$ has skew at most σ. The skew for layer 0 is given by the used clock generation scheme. If layer L has skew at most σ, i.e., any two correct neighbors have skew at most σ, we say that *the pulse has skew at most σ*. □

If Condition 3.12 is satisfied and all correct nodes have cleared all memory flags before a pulse arrives, it is straightforward to derive a (fairly coarse) skew bound.

Lemma 3.14: *Suppose all correct nodes in layer 0 send trigger messages during $[t_{\min}, t_{\max}]$, Condition 3.12 holds, and no correct node in any layer $\ell' \in [\ell+1]$, where $\ell \in [L+1]$, memorizes a trigger message from another correct node or is sleeping at time $t_{\min} + \ell' d^-$. With $f_\ell \leq f$ denoting the number of layers $\ell' \in [\ell]$ containing some faulty node, all correct nodes on layer ℓ are triggered at times within $[t_{\min} + \ell d^-, t_{\max} + (\ell + f_\ell)\, d^+]$. In particular, the pulse has skew at most $\sigma(f) < t_{\max} - t_{\min} + \varepsilon L + f d^+$.* □

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

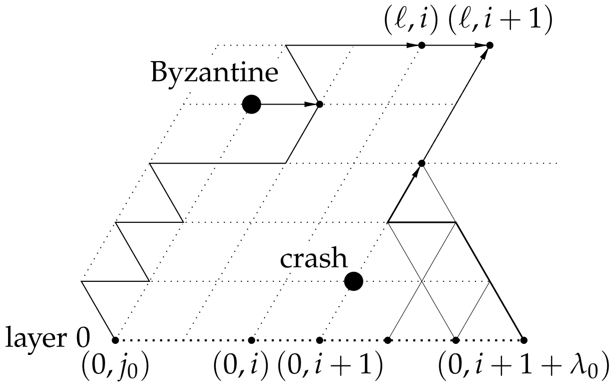


Figure 3.8: Illustration of alternative causal path construction to bypass a Byzantine node (upper left) and circumventing a crashed (or Byzantine) node when arguing why $(\ell,i+1)$ is triggered in a timely fashion (lower right).

Proof: By induction on ℓ. Clearly the statement is true for $\ell = 0$. The step is trivial for the lower bound, since Condition 3.12 implies that each node needs to receive a trigger message from a correct neighbor to be triggered, which is delayed by at least d^- time. If the upper bound is satisfied for $\ell \in [L]$ and all nodes in layer ℓ are correct, certainly all nodes in layer $\ell+1$ are triggered within d^+ time. If there is a faulty node in layer ℓ, the upper bound allows for $2d^+$ time for all nodes on layer $\ell+1$ to be triggered. If a correct node on layer $\ell+1$ has a faulty neighbor on layer ℓ, Condition 3.12 necessitates that either its right or its left neighbor is correct and has only correct neighbors on layer ℓ; the claim hence follows, as the node will be left- or right-triggered within $2d^+$ time.■

This lemma shows that even in the presence of multiple faults the time to complete a pulse increases only moderately. Thus, increasing the time between pulses (originating from layer 0) accordingly will maintain a clean separation of pulses.

While Lemma 3.14 guarantees bounded skew and suggests that actually $\sigma(f) = \sigma_\ell + \mathcal{O}(fd^+)$, where σ_ℓ is the fault-free layer ℓ intra-layer skew given in Theorem 3.9, a more detailed reasoning is required to prove such a bound. Unfortunately, the number of cases that needs to be considered in a formal proof explodes quickly. A large number of cases needed to be examined already in the fault-free case, and dealing with just a single fault becomes tedious, hence we refer to the appendix of [DFL+16] for a consideration of the faulty case. Informally, the reasoning employs the following arguments: For a Byzantine faulty node, there are only two options for increasing the skew between neighbors: (i) "shortcut" a causal path to the fast node and (ii) refrain from triggering nodes to inhibit the propagation of the pulse to the slow node.

Dealing with (i) is straightforward: If during the construction of a causal path we run into a Byzantine node, we follow the *other* incoming causal link of the predecessor node instead, thereby avoiding the Byzantine node, and resume the construction. This is particularly simple for the left zig-zag paths used in Cases 2 and 3 of Lemma 3.7; see

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Figure 3.8 for an example of how this might look like, where the faulty node is located at $(\ell - 1, j_0)$. When dealing with (ii), the situation is similar. Instead of circumventing faulty nodes in the construction of a causal path, we now need to avoid relying on them to trigger correct nodes. Figure 3.8 also gives an example for this, for a faulty node at $(1, i + 1)$.

Consequently, in order to increase skews significantly, *several* faulty nodes need to be in a region that causally affects two neighbors. In a setting where delays are random, it seems unlikely that the elaborate patterns required for large skews will arise, in particular if faults are not in close vicinity of each other. The simulations in Section 3.4 support this view, showing moderate increase of skews despite a significant number of faults.

We point out that *crash failures*, where nodes simply cease operating, are more benign. Instead of breaking the entire system, two adjacent crash failures on some layer just effectively crash their common neighbor in the layer above and affect the skews of the surrounding nodes.

3.2.4 Self-Stabilization

Self-stabilization is the ability of the system to recover from an unbounded number of transient faults [Dij74], which may put the system into an arbitrarily erroneous state. When transient faults cease, the system will resume normal operation within a bounded *stabilization time*—ideally even in the presence of a bounded number of persistent faults. In this section, we will show that HEX is self-stabilizing even in the presence of up to f Byzantine faulty nodes that satisfy Condition 3.12. Under the assumption that correct nodes faithfully execute the HEX algorithm, the pulse distribution will eventually work as specified in Lemma 3.14, even when all nodes start from arbitrary internal states. Note that for the whole grid to be self-stabilizing, the pulse generation at layer 0 must use a self-stabilizing algorithm as well, e.g., the self-stabilizing and Byzantine fault-tolerant FATAL$^+$ [DFLS14].

The analysis in Section 3.2.1 assumed that $T^-_{\mathbf{link}}$, $T^-_{\mathbf{sleep}}$, and the time between pulses are sufficiently large for all correct nodes to clear their memory and complete their sleeping period before the next pulse arrives. In the previous section, we argued that faulty nodes have an adverse, but bounded, effect on the skew and the time to complete a pulse. To ensure self-stabilization, we account for this by some additional slack in the time between the k^{th} and $(k+1)^{\text{th}}$ pulses, enabling nodes to reach consistent states even when the initial states are arbitrary. Any choice of parameters thus represents a trade-off between the frequency at which pulses can be issued and the fault-tolerance properties of the system. The following condition provides conservative bounds for the parameters $T^-_{\mathbf{link}}$, $T^+_{\mathbf{link}}$, $T^-_{\mathbf{sleep}}$, and $T^+_{\mathbf{sleep}}$, as a function of the number of faults and the inaccuracy of the implementation's local time measurements. We stress that, to ensure self-stabilization, the timers must be designed so that they expire within $T^+_{\mathbf{sleep}}$ and $T^+_{\mathbf{link}}$ time, respectively, even when started from an *arbitrary* internal state.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Condition 3.15 (Timing Constraints): *For* $k \in \mathbb{N}$, *define*

$$t^k_{\min} := \min_{\substack{i \in [W] \\ (0,i) \text{ correct}}} \left\{ t^k_{0,i} \right\} \quad \text{and} \quad t^k_{\max} := \max_{\substack{i \in [W] \\ (0,i) \text{ correct}}} \left\{ t^k_{0,i} \right\}.$$

An execution of Algorithm 3.1 has pulse separation time $\mathcal{S}$, *if, for all* $k \in \mathbb{N}$, *it holds that* $t^{k+1}_{\min} \geq t^k_{\max} + \mathcal{S}$. *For a given number of Byzantine faults* f *in the grid with* stable skew $\sigma(f)$, *we define*

$$\begin{aligned}
T^-_{\textbf{link}}(f) &:= \sigma(f) + \varepsilon \\
T^+_{\textbf{link}}(f) &:= \vartheta T^-_{\textbf{link}}(f) \\
T^-_{\textbf{sleep}}(f) &:= 2T^+_{\textbf{link}}(f) + 2d^+ \\
T^+_{\textbf{sleep}}(f) &:= \vartheta T^-_{\textbf{sleep}}(f) \\
\mathcal{S}(f) &:= T^-_{\textbf{sleep}}(f) + T^+_{\textbf{sleep}}(f) + \varepsilon L + f d^+.
\end{aligned}$$

Here, $\vartheta \geq 1$ *bounds the maximum clock drift, in the sense that* $t' - t \leq T' - T \leq \vartheta(t' - t)$ *for all real-times* $t' \geq t$ *with clock readings* T', T. □

In this definition, the *stable skew* $\sigma(f)$ is meant to be a bound on the skew between *any* two correct neighboring nodes, assuming that the system has already "stabilized", i.e., when the preconditions of Lemma 3.14 are satisfied for each pulse. This flexibility allows to plug in either a conservative or a more optimistic skew bound $\sigma(f)$ into the stabilization analysis. Note that the constraints $T^-_{\textbf{link}}(f) \geq \sigma(f) + \varepsilon$ and $\mathcal{S}$ being sufficiently large must be satisfied for any realistic $\sigma(f)$, since it is necessary to ensure that nodes do not "forget" a pulse before they are triggered and wake up on time for the next; recall that these assumptions were also implicit in the analysis in Section 3.2.1.

Before we can cast the algorithm's self-stabilization properties into a theorem, we need to specify what it means for the HEX pulse propagation to have stabilized up to a certain layer.

Definition 3.16 (Stabilized Pulse Propagation):
For an execution of Algorithm 3.1 on the HEX grid, we say that *layer* $\ell \in [L+1]$ *is stable with skew at most* σ *in pulse* k, if:

- All layers $\ell' \in [\ell]$ are stable with skew at most σ in pulse k;
- Node (ℓ, i), $i \in W$, is not sleeping at time $t^k_{\min} + \ell d^-$, and it does not memorize any trigger messages from correct neighbors at this time;
- Layer ℓ has skew at most σ in pulse k. □

Assuming that all parameters are chosen in accordance with Condition 3.15 and $\sigma(f)$ is indeed a valid bound on the stable skew, we can now show that the system will recover from arbitrary initial states.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Theorem 3.17: *Suppose that, given values f and $\sigma(f)$, an execution of Algorithm 3.1 satisfies the following prerequisites:*

- *There are at most f Byzantine faulty nodes satisfying Condition 3.12.*
- *The stable skew is at most $\sigma(f)$.*
- *The parameters $T^-_{\mathbf{link}}$, $T^+_{\mathbf{link}}$, $T^-_{\mathbf{sleep}}$, and $T^+_{\mathbf{sleep}}$ in Algorithm 3.1 are chosen in accordance with Condition 3.15.*
- *The pulse separation time is larger than $\mathcal{S}(f)$, as specified by Condition 3.15.*

Then, each layer $\ell \in [L+1]$ is stable with skew at most $\sigma(f)$ in all pulses $k > \ell$. Moreover, for each $i \in W$ such that (ℓ, i) is correct and each pulse $k > \ell$, there is a unique triggering time $t^k_{\ell,i}$ of (ℓ, i) during $\left[t^k_{\min} + \ell d^-, t^{k+1}_{\min} + \ell d^-\right)$. □

Note that this, in particular, implies that all correct neighbors will satisfy the skew bound $\sigma(f)$ in all pulses $k > L$. In order to prove the theorem, we first show a helper statement saying that if all layers up to layer ℓ are stable in *some* pulse, they will satisfy the claim of the theorem in *all* subsequent pulses.

Lemma 3.18: *Assume that the preconditions of Theorem 3.17 are satisfied and that all layers $\ell \in [L+1]$, $\ell' \in [\ell+1]$, are stable with skew at most $\sigma(f)$ in pulse $k \in \mathbb{N}$. Then, these layers are also stable with skew at most $\sigma(f)$ in pulse $k+1$, and for each correct node (ℓ, i), there is a unique triggering time $t^k_{\ell,i}$ during $\left[t^k_{\min} + \ell d^-, t^{k+1}_{\min} + \ell d^-\right)$.* □

Proof: Since we have that all layers $\ell' \in [\ell+1]$ are stable, Lemma 3.14 shows that all nodes in these layers are triggered during $\left[t^k_{\min} + \ell' d^-, t^k_{\max} + (\ell' + f_{\ell'})\, d^+\right]$. For each node (ℓ', i) in such a layer, let $t^k_{\ell',i}$ be the minimal triggering time in this interval (we still need to establish that there is only one). Since the stable skew is at most $\sigma(f)$, for any node (ℓ', i) with $0 \neq \ell' \leq \ell$, the triangle inequality yields that its correct neighbors on layers ℓ' and $\ell' - 1$ trigger within $2\sigma(f)$ of each other. Hence, all trigger messages from correct neighbors are received within a time window of duration $2\sigma(f) + \varepsilon$. We have $T^-_{\mathbf{sleep}}(f) > 2T^-_{\mathbf{link}}(f) > 2\sigma(f) + \varepsilon$, implying that nodes will not memorize any late pulse k trigger messages from correct neighbors after waking up. We thus conclude that, for each node (ℓ', i), $t^k_{\ell',i} < t^{k+1}_{\min} + \ell' d^-$ is unique. Hence, our assumptions yield

$$t^k_{\max} \leq t^{k+1}_{\min} - \mathcal{S}(f) < t^{k+1}_{\min} - T^+_{\mathbf{sleep}}(f) - T^-_{\mathbf{sleep}}(f) - \varepsilon L - f d^+. \tag{3.26}$$

Hence, the upper bound on the pulse k triggering times of layer ℓ' nodes established in Lemma 3.14 leads to

$$\begin{aligned} t^k_{\ell',i} &\leq t^k_{\max} + \left(\ell' + f_{\ell'}\right) d^+ \\ &< t^{k+1}_{\min} - T^+_{\mathbf{sleep}}(f) - T^-_{\mathbf{sleep}}(f) - \varepsilon L - f d^+ + \left(\ell' + f_{\ell'}\right) d^+ \\ &\leq t^{k+1}_{\min} - T^+_{\mathbf{sleep}}(f) - T^-_{\mathbf{sleep}}(f) + \ell' d^-. \end{aligned} \tag{3.27}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

We thus observe that no node will be sleeping or have memorized any trigger messages from other correct nodes at time $t_{\min}^{k+1} + \ell' d^-$. Since we assumed that $\sigma(f)$ is a bound on the stable skew, the requirements of Definition 3.16 are met for pulse $k+1$ and all layers $\ell' \in [\ell + 1]$, as claimed. ■

With this lemma, the proof of Theorem 3.17 boils down to showing that if layer ℓ is stable in pulse k, then layer $\ell + 1$ is stable in pulse $k+1$.

Proof (Proof of Theorem 3.17): We prove the theorem by induction on ℓ, where the hypothesis is that the claims of the theorem are satisfied by all layers $\ell' \in [\ell + 1]$. For $\ell = 0$, the statement is trivial. For the step from ℓ to $\ell + 1$, repeated use of Lemma 3.18 reveals that it is sufficient to show that layer $\ell + 1$ is stable in pulse $\ell + 2$.

By the induction hypothesis and Lemma 3.14, no correct node in layer ℓ sends trigger messages during $\left[t_{\max}^{\ell+1} + (\ell + f)d^+, t_{\min}^{\ell+2} + \ell d^-\right]$. Using exactly the same derivation as for Equation (3.27) in the proof of Lemma 3.18, we find that any triggering message originating in a correct layer ℓ node must have arrived at any correct layer $\ell + 1$ node before time $t := t_{\min}^{\ell+2} - (T_{\mathbf{sleep}}^{+}(f) + T_{\mathbf{sleep}}^{-}(f)) + (\ell + 1)d^-$. We will complete the proof by showing that this entails that correct nodes $(\ell + 1, i)$ will neither sleep nor memorize trigger messages from correct nodes at time $t_{\min}^{\ell+2} + (\ell + 1)d^-$.

To this end, suppose that starting from the above time t, nodes on layer $\ell + 1$ do not receive trigger messages from correct nodes on the previous layer. By Algorithm 3.1, nodes will forget messages that arrived more than $T_{\mathbf{link}}^{+}(f)$ time ago. Thus, the latest time when a correct node $(\ell + 1, i)$ in layer $\ell + 1$ could be triggered due to a remembered message from some *correct* neighbor on layer ℓ is smaller than $t + T_{\mathbf{link}}^{+}(f)$.

Hence, consider the case that $(\ell + 1, i)$ has a faulty neighbor on layer ℓ. W.l.o.g., assume that the faulty neighbor is node $(\ell, i + 1)$ (the other case is symmetric). By the above reasoning, $(\ell + 1, i)$ can only be left-triggered at or after time $t + T_{\mathbf{link}}^{+}(f)$, which in addition requires a memorized trigger message from $(\ell + 1, i)$. Thus, if neither $(\ell + 1, i)$ nor $(\ell + 1, i + 1)$ are triggered during $\left[t - d^+, t + T_{\mathbf{link}}^{+}(f)\right)$, neither node can be triggered anymore: Condition 3.12 guarantees that both nodes have no other faulty neighbor than $(\ell, i + 1)$, and the above reasoning applies also to $(\ell + 1, i)$.

Therefore, assume that one of them, say $(\ell + 1, i + 1)$, is triggered at time $t_{\ell+1,i+1} \in \left[t - d^+, t + T_{\mathbf{link}}^{+}(f)\right)$. It will not wake up (and therefore not be triggered again) before time $t_{\ell+1,i+1} + T_{\mathbf{sleep}}^{-}(f)$. Node $(\ell + 1, i)$ receives the trigger message from $(\ell + 1, i + 1)$ by time $t_{\ell+1,i+1} + d^+$ and therefore either

(a) forgets it by time $t_{\ell+1,i+1} + T_{\mathbf{link}}^{+}(f) + d^+$ or

(b) is triggered during $\left[t_{\ell+1,i+1}, t_{\ell+1,i+1} + T_{\mathbf{link}}^{+}(f) + d^+\right)$.

If $(\ell + 1, i)$ is triggered, its corresponding trigger message to $(\ell + 1, i + 1)$ arrives by time $t_{\ell+1,i+1} + T_{\mathbf{link}}^{+}(f) + 2d^+ \le t_{\ell+1,i+1} + T_{\mathbf{sleep}}^{-}(f)$. This message thus arrives at $(\ell + 1, i + 1)$ before it wakes up again. When $(\ell + 1, i + 1)$ wakes up, it clears its memory and will not be re-triggered before $(\ell + 1, i)$ is triggered again (or the next pulse arrives

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

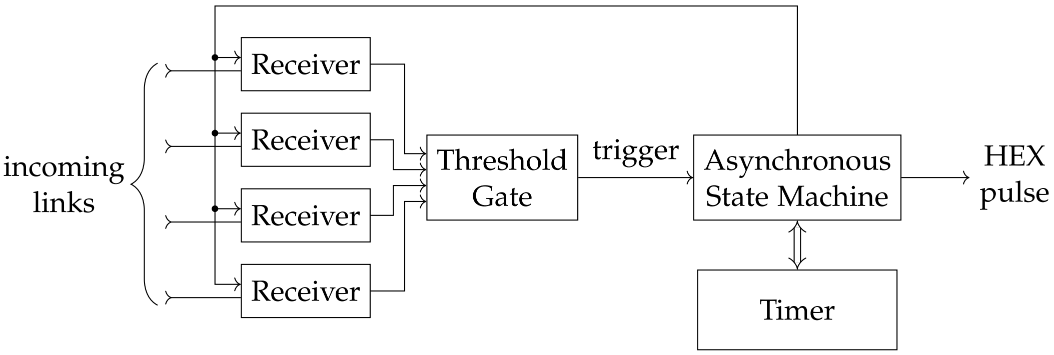


Figure 3.9: Schematic overview of the implementation of a HEX node. On the left are the receivers. Those are connected to the threshold gate, which provides the ASM with a trigger signal when one of the firing condition is satisfied. The ASM performs the state transitions, interacts with the timer, and outputs a HEX pulse when necessary. The reset signal as well as the link timer of the receivers, are not drawn in this figure to avoid clutter.

on layer ℓ). But $(\ell+1, i)$ cannot be triggered again: If it was not triggered then (a) holds and hence it forgot any previous trigger messages. Otherwise, (b) holds and the node will clear its memory upon waking up. Consequently, $(\ell+1, i+1)$ cannot be triggered again either.

Overall, no node on layer $\ell+1$ is triggered after time $t_{\ell+1,i+1} + T^{+}_{\mathbf{link}}(f) \leq t + 2T^{+}_{\mathbf{link}}(f) + d^{+}$ or receives trigger messages from correct nodes after time $t + 2T^{+}_{\mathbf{link}}(f) + 2d^{+} \leq t + T^{-}_{\mathbf{sleep}}(f)$ $\left(\text{until time } t^{k+1}_{\min} + (\ell+1)d^{-}\right)$. Hence, all nodes are awake by time $t + T^{-}_{\mathbf{sleep}}(f) + T^{+}_{\mathbf{sleep}}(f) > t + T^{-}_{\mathbf{sleep}}(f) + T^{+}_{\mathbf{link}}(f)$ and have forgotten all spurious messages. By the previous observations, this concludes the proof. ■

3.3 Implementation

This section focuses on the implementation of a HEX node, with the aim to have a model for simulations of the grid. These simulations are used to verify the previously established theoretical results. The implementation has been done in VHDL and consists of multiple components, as shown in Figure 3.9. The main component is the *asynchronous state machine* (ASM), the primary input (trigger) of which is provided by a threshold gate. The threshold gate takes the outputs of the receivers of the incoming links of the HEX node and checks whether one of the firing conditions, defined by Algorithm 3.1, is satisfied.

The implemented design was synthesized using the UMC 90 nm standard cell library [UMC03], using Synopsis® Design Compiler version C-2009.06-SP4. Note that we used an augmented version of the UMC 90 nm library, which includes a custom Muller C-Gate [Mul55; Sut89] added in the context of the DARTS project [FS11; FS12].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

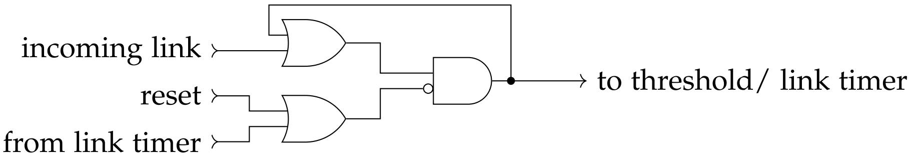


Figure 3.10: The receiver component of a HEX node. A simple memory flag that stores, through a loopback, any trigger message (= rising signal transition) received from a neighbor.

To provide timing characteristics for this gate, the timing information of the AND-Gate of said library was used; the resulting inaccuracy is negligible w.r.t. our purposes.

3.3.1 Receiver

The receivers are implemented as simple resettable memory flags. Their sole purpose is to store an incoming trigger message, which is just a rising signal transition on the single wire of the link, from a neighbor until it is reset. To limit the influence of Byzantine faults, a dedicated link timer is attached to each receiver. This timer is started once the trigger message has received, and resets the storage loop after the timeout of $T_{\mathbf{link}}$. If the ASM issues a reset to the receiver, the receiver itself as well as the link timer are reset. The implementation is shown in Figure 3.10, with a generic state machine depicted in Figure 3.11b. In the current implementation, the input of a link is just tied to the HEX pulse output of the sender node, i.e., a trigger message is just the HEX pulse output of the ASM.

3.3.2 Asynchronous State Machine

The state machine has basically three states, and is similar to the quick cycle state machine of FATAL$^+$ [DFL+14]. The states are the following:

ready The state machine waits in this initial state until one of the firing conditions of Algorithm 3.1 becomes true.

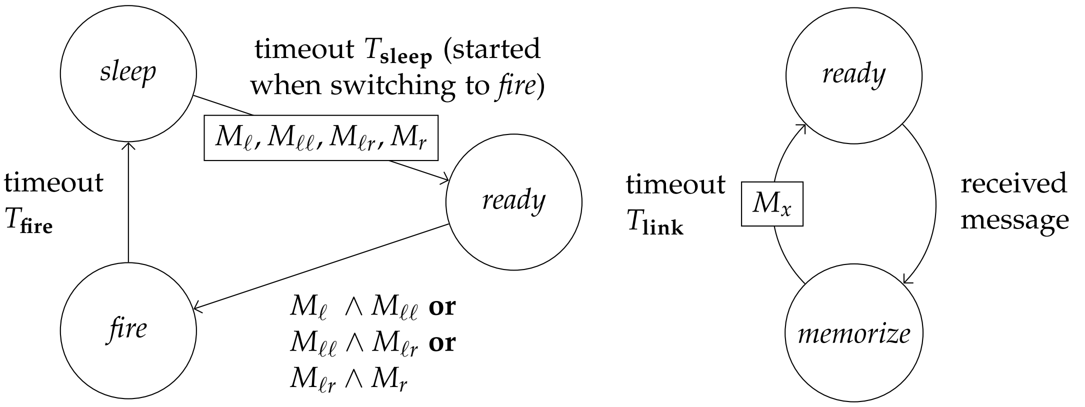
fire In this state, a HEX pulse is generated, and the timer is started. When the timer reaches $T_{\mathbf{fire}}$ a transition to *sleep* is performed.

sleep The state machine sleeps until the timer started in state *fire* reached a time within the interval $[T^-_{\mathbf{sleep}}, T^+_{\mathbf{sleep}}]$. After the timeout, the state machine transitions to *ready*, and resets the receivers.

The transitions between those states are cyclic, i.e. *wait* to *fire*, *fire* to *sleep* and *sleep* to *wait*. A high-level visualization is given in Figure 3.11a, whereas Figure 3.12 refers to the specific signal transitions as used in Figure 3.13.

While the quick cycle state machine in FATAL$^+$ was implemented as a *hybrid state machine* (HSM), the implementation here uses a ASM. A HSM is a combination of an

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.


sleep
timeout $T_{\textbf{sleep}}$ (started when switching to *fire*)
$M_\ell, M_{\ell\ell}, M_{\ell r}, M_r$
timeout $T_{\textbf{fire}}$
ready
fire
$M_\ell \wedge M_{\ell\ell}$ **or**
$M_{\ell\ell} \wedge M_{\ell r}$ **or**
$M_{\ell r} \wedge M_r$
ready
timeout $T_{\textbf{link}}$
M_x
received message
memorize

(a) State machine of the firing algorithm.

(b) State machine for the resettable memory flag M_x.

Figure 3.11: The state machines of a HEX node. Circular nodes represent states connected by transitions, labeled with their transition guards. A boxed node lists the memory flags to be cleared when taking the transition. The right state machine describes the link timers, while the left state machine implements the firing algorithm as defined by Algorithm 3.1.

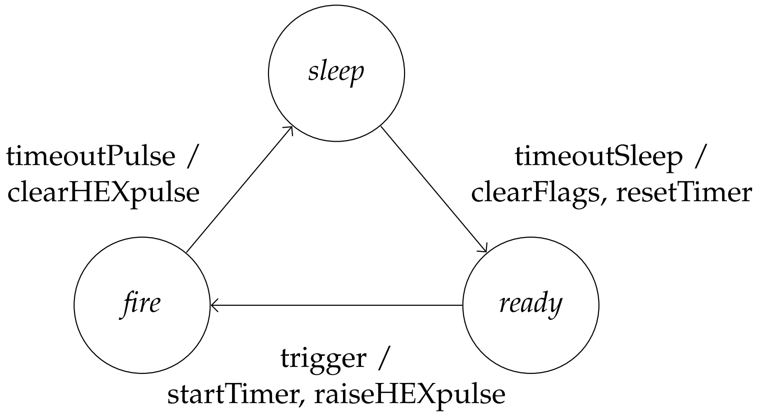

sleep
timeoutPulse / clearHEXpulse
timeoutSleep / clearFlags, resetTimer
fire
ready
trigger / startTimer, raiseHEXpulse

Figure 3.12: The state transition diagram of the ASM. Circular nodes represent states, edges represent the transitions. The transition label includes the guarding signal of the transition separated by a slash from the outputs which are enabled on taking the transition.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

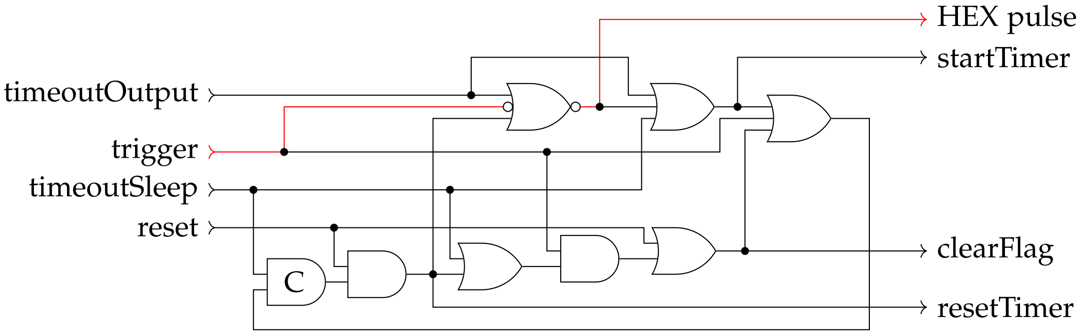


Figure 3.13: The logic elements building the ASM of a HEX node. Input signals are on the left, output signals are on the right. The path from acceptance of a firing condition to generation of a HEX pulse is marked red.

ASM with a synchronous *transition state machine* (TSM). The asynchronous part decides which transition should be taken and starts a local oscillator. This oscillator acts as the clock for the TSM, which then executes the actual state transition. This is done to ensure that only one transition is done at a time, as the asynchronous part could enable multiple transition at the same time. Although a HSM would be a valid approach for the problem at hand, it would suffer from a large delay and area overhead.

Since the state transition diagram in Figure 3.12 is deterministic, i.e., the transition to be taken is predefined in every state and not conditional on certain signals, a pure ASM is sufficient. Compared to the HSM approach, however, an additional timeout, *timeoutPulse*, is required to ensure the proper length of the HEX pulse, which could be ensured by the timed transition sequence *wait* → *fire* → *sleep* of the TSM. The design of this ASM was done with Petrify [CKK+02].

As described in more detail in Section 2.1.2.3, Petrify allows the generation of speed-independent asynchronous circuits. The relationship between the signal transitions were given to Petrify in form of the *signal transition graph* (STG) of Figure 3.14. From this input, Petrify generated a net-list, on which a technology mapping to Verilog can be performed. This feature was used, with a generic technology library augmented by Muller C-Gates, to generate a Verilog net-list, which was then manually converted to VHDL. During the manual conversion, a *reset* signal was incorporated into the design, which ensures the required initial states for the ASM.[7] In Figure 3.15, the timing diagram of an execution of the ASM is shown, while Figure 3.13 shows the circuit diagram of the generated HDL code.

The ASM has as its main input the *trigger* signal, which comes directly from the threshold gate. This signal is responsible for triggering a local pulse, represented by

---

[7] Note that the design is still self-stabilizing. The explicit reset is included, as the multivalued logic system (IEEE 1164 [IEE93]) used by ModelSim may not allow to reach a valid state from an uninitialized start. The testing of the self-stabilizing properties of the HEX grid is not directly influenced by this. We describe in Section 3.4.1.2 how an "uninitialized" valid state can be reached.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

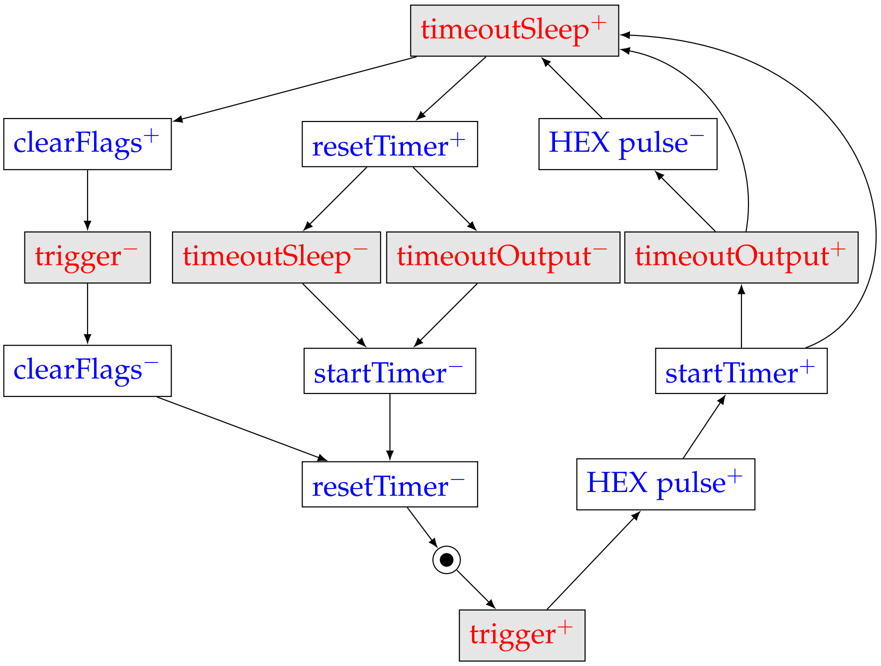


Figure 3.14: The *signal transition graph* (STG) that describes the ASM implementing the HEX algorithm. Nodes with red text, and filled gray represent events on input signals of the state machine, whereas white nodes with blue text represent output signals. A *signal*$^{+}$ as a node label represents a rising transition of the signal, *signal*$^{-}$ a falling transition. More detail on the interpretation of a STG can be found in Section 2.1.2.3.

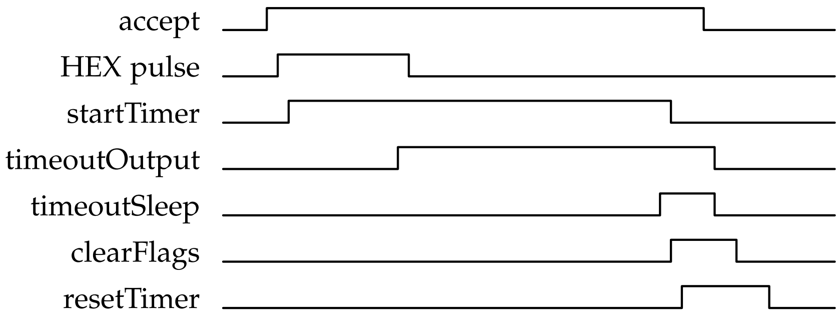


Figure 3.15: Timing diagram of a complete cycle of the ASM

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

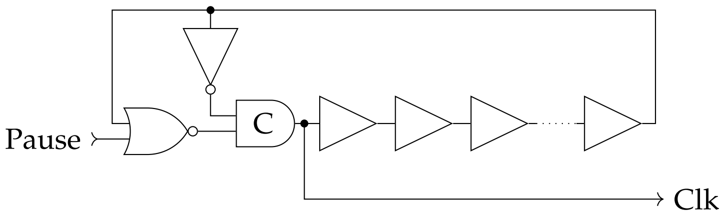


Figure 3.16: Start/stoppable ring oscillator used for generating the clock of the timer. The clock frequency is determined by a series of buffer elements, which determine the delay of a feedback loop that also incorporates a Muller C-Gate, i.e., allows to open/close the feedback loop and hence to start/stop the generation of the clock signal.

the signal *HEX pulse*, if the node is not sleeping. The timer is controlled with *startTimer* and *resetTimer*, while the timeouts are introduced into the ASM with *timeoutOutput*, which defines the length of the HEX pulse signal, and *timeoutSleep*, which wakes up the node after sleeping for some time $T \in \left[T^{-}_{\mathbf{sleep}}, T^{+}_{\mathbf{sleep}}\right]$. Finally, the signal *clearFlags* is used to reset the receivers' memory flags.

Note that it is possible to find multiple correct STGs for the given ASM. They differ in the detailed dependencies of the transitions, e.g., the *clearFlags* could be enabled after the timeout signal was reset. The version given in Figure 3.14 was chosen, as it does not contain any RS-latches in the output generated by Petrify, which could introduce metastability into the system.

3.3.3 Timer and Ring Oscillator

As the ASM needs some kind of time reference, to guarantee the minimal/maximal duration of the HEX pulse and the link/sleep/reset times, multiple timers are needed. The timer design consists of a synchronous counter driven by a start/stoppable ring oscillator, similar to the implementation in FATAL^{+} [DFL+14].

A single timer consists of a register, which stores a value that is incremented with every clock cycle of the ring oscillator, until the oscillator is stopped. When a predefined value is reached, the timer asserts an overflow signal,[8] which represents the timeout, and resets the register to its initial value. Our implementation of a HEX node uses one timer for each receiver and two timers for the state machine. Each receiver uses the same link timer design. The variances in the timeout value, due to differences in the placement and routing on the chip, can be covered by the uncertainty of the interval $\left[T^{-}_{\mathbf{link}}, T^{+}_{\mathbf{link}}\right]$. The timers for the state machine are, first, lower-range timer called *timeoutPulse* that specifies the length of the HEX pulse, which will be generated by the ASM. The second timer of the state machine is a higher-range timer called

---

[8] To reduce the stabilization time of the timer, the overflow signal must be asserted for any value larger than the sought after timeout, even though these value cannot be reached in a correct execution.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

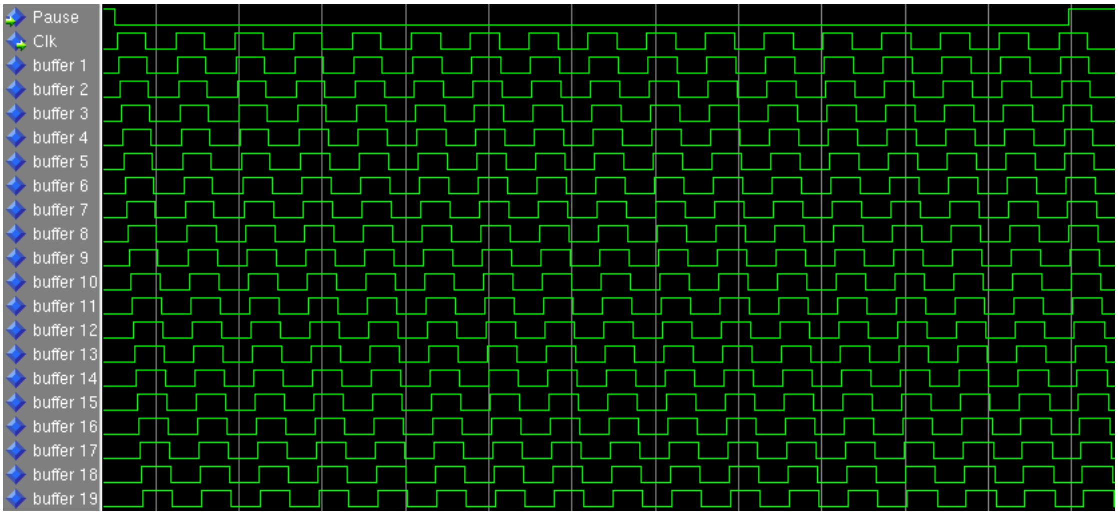


Figure 3.17: A screenshot of a simulation of the ring oscillator. The oscillator has 19 buffer gates, which form the delay chain. On the left, it is also visible how the active high *pause* signal controls the generation of the clock signal.

timeoutSleep: It enables the transition from *sleep* to *wait* in the ASM. The two timers are implemented sharing the same register with two "overflow" signals, as both are started at the same time and the pulse timeout is strictly smaller than the sleep timeout. Hence, only the sleep timeout stops the oscillator of this timer.

The ring oscillator is implemented by a feedback loop consisting of a Muller C-gate, an inverter, and a chain of buffer gates,[9] which delays the signal and thus determines the clock frequency, cf. Figure 3.16. The Muller C-Gate's purpose is to allow the generation of the clock signal to be paused. This is necessary to avoid metastability in the timer: The signal, which starts the timer, could violate the setup/hold-requirements of the timer register when the clock would always be running. An exemplary execution of the ring oscillator with 19 buffer gates is shown in Figure 3.17.

3.3.4 Implementation Characteristics

In the pre-layout simulations, the response time of the HEX node circuit, from applying the last signal of a firing condition at a receiver to the generation of the HEX pulse, was within $[161, 197]$ ps , depending on which firing condition enabled the triggering of the node. The difference is caused by (i) the variance in the delay, and (ii) the difference in length of the signal paths through the threshold gate, as the synthesis used only 2-input gates.

---

[9] An implementation with an odd total number of inverters is feasible as well. However, the buffers have a higher delay in the ModelSim simulation. Thus, we can achieve the same delay with a smaller number of gates, hence, reducing the required resources for the simulations.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

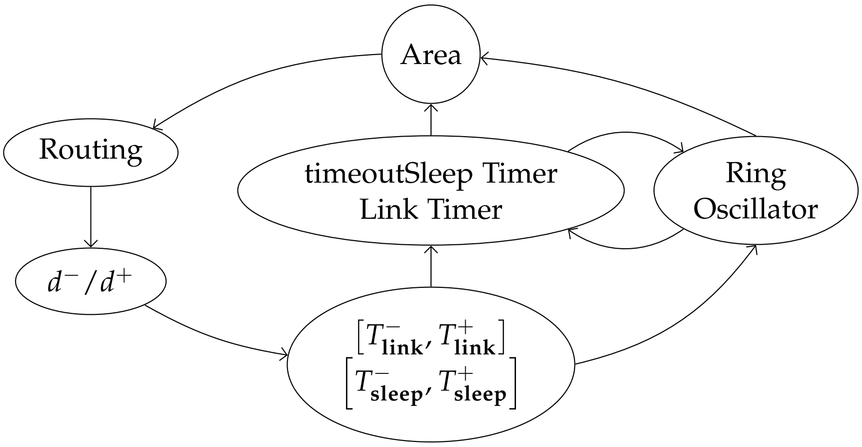

Area
Routing
timeoutSleep Timer
Link Timer
Ring
Oscillator
d^- / d^+
$[T^-_{\text{link}}, T^+_{\text{link}}]$
$[T^-_{\text{sleep}}, T^+_{\text{sleep}}]$

Figure 3.18: Cyclic dependencies of the different factors that influence the required area of the timer and the ring oscillator.

We can expect the response time to increase in the post-layout design, where the wire delays are incorporated. Nonetheless, the post-layout response time will most likely be affected by the actual end-to-end link delays d^- and d^+, which in turn incorporate the response time: As the timeout intervals $[T^-_{\text{link}}, T^+_{\text{link}}]$ and $[T^-_{\text{sleep}}, T^+_{\text{sleep}}]$ depend on those values, recall Condition 3.15, this directly affects the area required for the timer and the ring oscillator, which in turn can influence these response times, due to routing decisions. Due to this cyclic dependency, a comprehensive verification of the timing parameters is imperative to ensure the correct operation of a HEX grid. Cf. Figure 3.18 for a graphical representation of the cyclic dependencies of the parameters just discussed.

In terms of area, the current HEX node design has size of about 117 μm^2, without the timers and with an approximated area for the Muller C-Gates. The timers will, very likely, be the dominant factor in terms of area requirements of the design. As the area of the timers depends on the frequency of the ring oscillator and the bits required to represent the timer value, however, it is not possible to give an exact value for the area of each timer. Clearly, there is an area trade-off between the size of the counter and the size of the oscillator as a higher frequency (smaller oscillator) requires a larger counter. In the designs used for the simulations later on, the area of the timer was, unsurprisingly, multiple times larger than the area of the other components of the HEX node together. With respect to the post-layout design, we note that the missing routing delays will significantly increase the delay of the feedback loop of the ring oscillator, which allows to reduce the number of buffer gates.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

3.4 Simulation Experiments

As mentioned earlier, the analytic results obtained in Section 3.2 are limited to worst-case skews. Hence, we conducted extensive simulation experiments to complement them with representative statistical data. This data will reveal the following major facts about HEX:

(C1) *Average skews are considerably smaller than worst-case.* The quite fancy scenarios required for establishing the worst-case skews in Theorem 3.9 already suggested that they are very unlikely to occur in practice. Indeed, we were not able to generate neighbor skews that came close to the worst case under uniformly and independently distributed link wire delays.[10]

(C2) *Small, localized fault effects.* As argued in Section 3.2.4, HEX should implicitly confine the effects created by a faulty node to a small neighborhood. Our simulation results revealed that faults typically affect direct neighbors only, and become invisible after one additional hop. This is true even for clustered crash faults and multiple (separated) Byzantine faults.

(C3) *Stabilization times are much smaller than guaranteed by theoretical analysis.* The layer-wise stabilization used for bounding the worst-case stabilization time in Section 3.2.4 assumes an involved worst-case scenario *on each layer*, which appears to be very unlikely in practice. Simulations of the original version [DFL+13] of the HEX algorithm already revealed typical stabilization times that are much smaller than the predicted worst case. The link timeouts added in Algorithm 3.1 cause HEX to reliably stabilize within two clock pulses, even in scenarios with multiple faults.

To acquire the necessary data, a simulation framework has been develop specifically for this purpose.

3.4.1 Simulation Framework

Our simulation framework has been built around Mentor Graphics® ModelSim 10.1d, a well-known environment for simulation and verification of hardware designs, which allows accurate digital timing simulations of circuit designs specified in VHDL, among others.

Our simulation environment was build upon a testbed which was written in the Haskell programming language [Mar10]. Its top-level features is to generate a testbench for a test configuration and then start ModelSim to perform the simulation. The results of the conducted pre-layout simulation of the testbench, with timing annotations of the gates of the HEX node, where stored by ModelSim in a format that a separate

---

[10] With link wire delay, we refer to the delay between nodes in the grid. While these delays can be configured in the simulations, the processing delay of the HEX node is inherently fixed by the hardware design of the same.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

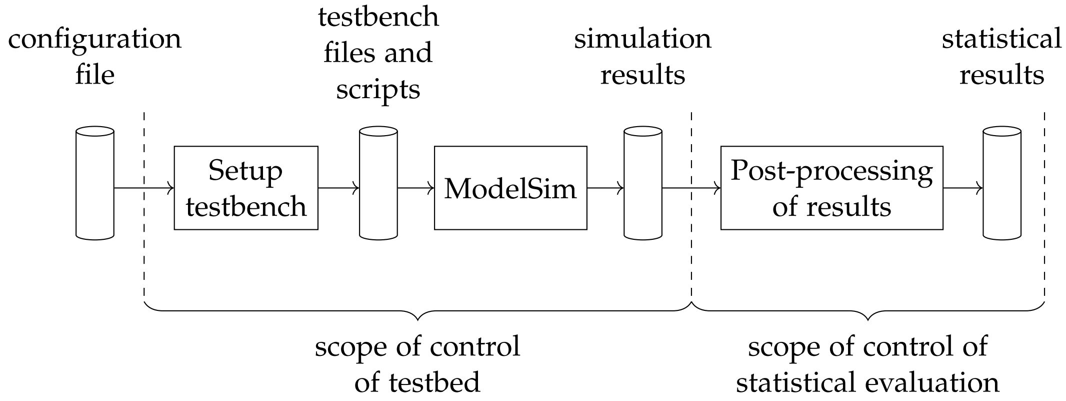

configuration file
testbench files and scripts
simulation results
statistical results
Setup testbench
ModelSim
Post-processing of results
scope of control of testbed
scope of control of statistical evaluation

Figure 3.19: This figure visualizes, from left to right, the process from a configuration file to statistical data. Starting with the configuration file, the testbed generates testbench files for each run, which are then simulated with ModelSim. The simulation results are then further evaluated by a statistical evaluation software, which provides the final results.

custom-build application used for the statistical evaluation. Figure 3.19 shows an overview of the major components used for the simulation of a specific setting.

A *set* describes a combination of parameters of the HEX grid that shall be simulated. As some parameters are randomly distributed, e.g., the link wire delays or the placement of the faults, multiple executions of the same set were used to acquire meaningful data. We call every such execution a *run*. For every run, the testbed part of the process shown in Figure 3.19 is executed, i.e., a new testbench is generated and then simulated.

We also distinguish between two types of sets: Sets which observe only one pulse in the HEX grid, and ones that observe multiple pulses, where *pulses* refers to the number of clock pulses generated in layer 0 that are simulated in a run. For the stabilization experiments, multiple pulses are obviously required, whereas for the skew evaluation a single pulse is sufficient.

Every set is generated from a configuration file provided by the user. From this configuration file, the testbed generates files that build the testbench for a run of that set. The testbench is then used as an input by ModelSim for the simulation of the HEX grid.

We will now explain each component of Figure 3.19 in detail.

3.4.1.1 Configuration File

The configuration file describes a specific set by defining the HEX grid, which includes the following properties:

- the size of the HEX grid.
- the interval in which the link wire delays shall be distributed.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

- the number of runs that shall be simulated.
- the number of pulse per run.
- the properties of the layer 0 clock source, i.e., the skew between these nodes.
- whether the link wire delays shall change on a per-pulse basis.
- whether an arbitrary initial state of the HEX nodes shall be enforced before the first pulse. This was used for the stabilization experiments.
- the number and types of the faults which should be placed in the HEX grid, without violating the 1–local fault assumption. The types of faults include
  - fail-silent faulty nodes.
  - Byzantine faulty nodes, which behave, on a per-link basis, either fail-silent or send constantly a (fast) trigger message.
  - link-faulty nodes, i.e., correct nodes where a few outgoing links are faulty.

  The behavior of the Byzantine and the link-faulty nodes are chosen randomly on a per-link basis, in every run.
- the initial seed used for the random number generator.

3.4.1.2 Testbench

As stated above, the testbed is provided with a configuration file for a specific test configuration, from which multiple files are generated. These files are needed to build a complete testbench cf. Figure 3.20. Further, they setup ModelSim to simulate the specific run of the given configuration. The files are

- the *Testbench*. This design entity acts as a clock source for the layer 0 pulses, selects the link wire delays, monitors the HEX grid and applies the reset signal to the other component of the testbench on startup. The actual monitoring of the system is not done within the testbench, but rather uses ModelSim facilities. Nonetheless, the signals are mapped to allow later enhancements, e.g., controlling more intelligent faults.
- the *Network* configuration, which is responsible for connecting the HEX nodes, thereby forming the HEX grid.
- the *Grid*, which is not generated by the testbed, but is rather a given entity that is instantiated in the Network and configured to generate the HEX nodes needed. I.e., the Grid instantiates all HEX nodes and faulty nodes and provides the signals connecting the nodes to the network.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

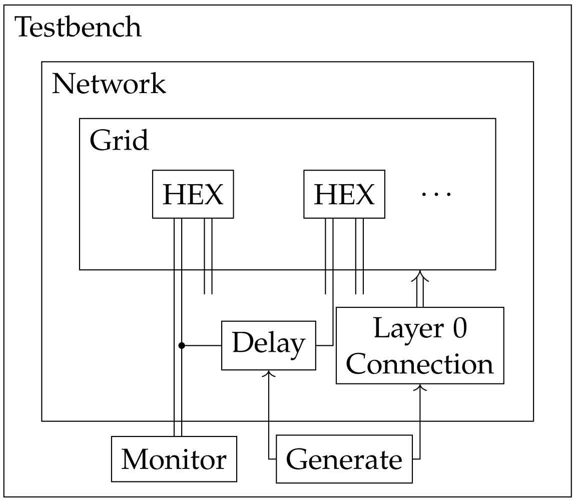


Figure 3.20: Simplified overview of the testbench generated by the testbed. The *Testbench* monitors the signals applied to and generated by the HEX nodes. It also acts as a clock source for the layer 0 pulses and selects the delays applied to the links, if they should change with every run. The *Network* generates the interconnect links between the nodes and thereby constructs the HEX grid. These interconnection links are equipped with delay elements, which simulate the link delays. The *Grid* instantiate the HEX nodes, or the selected faulty nodes, at the specified places in the HEX grid.
To avoid clutter, this figure only shows two HEX nodes, with one connected link. Also only two signals from one node are connected to the *Monitor*.

- an optional *Tool Command Language* (TCL) script [OJ09], which is executed by ModelSim after the reset signal was applied, to force an arbitrary state of all HEX nodes. This is done by forcing certain signals in the HEX nodes to randomly selected values, wait for a certain time until the values have propagated through the nodes, and then releasing the signals and let the nodes progress. The generation of this TCL script is aware of every forceable internal structure of the components, e.g., the number of register-bits in the timers.
- a script that configures ModelSim to monitored specific signals. These are dumped after the simulation concluded and stored as the simulation result. The dumping is done via the list feature of ModelSim, which proved to be less data and time consuming, w.r.t. post-processing, than the processing of the waveform files.

3.4.1.3 ModelSim

ModelSim is an environment for simulation and verification of hardware designs developed by Mentor Graphics®. ModelSim is frequently used in industrial practice and allows simulations from structural to post-layout level. The software allows extensive

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

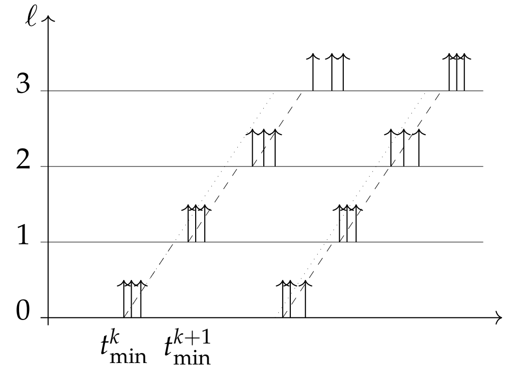


Figure 3.21: Exemplary multi-pulse execution. The arrows mark the triggering of certain nodes, the ℓd^- lines are dotted and the dashed lines mark the respective $\min_{i\in[W]}\left(t^k_{\ell,i}\right) + d^-$ line.

control over the signals in the design, monitoring of the same and, if applicable, timing checks on the components of the design.

3.4.1.4 Statistical Evaluation

The results, generated by ModelSim, are finally analyzed by a statistical evaluation application also written in Haskell. The operations conducted on the results are data intensive but, for most parts, trivial. One operation stands out as being more complex, though: In the run with multiple pulses, it must be determined whether some generated HEX pulse corresponds to pulse k or $k+1$. The trivial approach of counting from the last pulse recorded backwards does not work in these experiments, as, due to faults and stabilization, nodes can fire multiple times during the period of one pulse, which must be correctly detected. In settings with sufficient pulse separation time, the simple approach of splitting the timeline for node (ℓ, i) into intervals of the form $\left[t^k_{\text{min}} + \ell d^-, t^{k+1}_{\text{min}} + \ell d^-\right)$, with t^k_{min} being the time when the first node in layer 0 triggers pulse k (cf. Condition 3.15), is sufficient.

With decreasing pulse separation time, however, it is possible that $t^k_{\ell,i} \geq t^{k+1}_{\text{min}} + \ell d^-$, due to the inevitable jitter $\ell\varepsilon$ of the triggering times, i.e., there is no way to unambiguously assign a HEX pulse to pulse k or pulse $k+1$. To handle this case, a more involved approach is required. Therefore, instead of the theoretical lower interval bound, the practical lower interval bound is taken: A node in layer $\ell+1$ cannot trigger pulse k before the first node[11] in layer ℓ has done so plus the minimal link delay d^-, i.e., $\min_{i\in[W]}\left(t^k_{\ell+1,i}\right) \geq \min_{i\in[W]}\left(t^k_{\ell,i}\right) + d^-$. A visualization of the difference between those two approaches is shown in Figure 3.21.

---

[11] In the fault-free case, the triggering of the second node would be sufficient.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Below, we present selected results supporting the three claims stated at the beginning of this section:

(C1) Average skews are considerably smaller than worst-case.

(C2) Small, localized fault effects.

(C3) Stabilization times are much smaller than guaranteed by theoretical analysis.

Since we argue about different executions of HEX in this section, for a set R of executions, we denote by $t^k_{\ell,i,\rho}$ the time when (ℓ, i) forwards the k^{th} pulse in execution $\rho \in R$. Using the simulation framework, we conducted the following two types of simulation experiments:

Statistical evaluation of the neighbor skews These experiments require simulation runs involving only a single pulse propagating through the HEX grid. The primary quantities of interest in a simulation run ρ are:

- the (absolute) skews $|t_{\ell,i,\rho} - t_{\ell,i-1,\rho}|$ between neighbors (ℓ, i) and $(\ell, i-1)$ in the same layer ℓ,
- the (signed) skews $t_{\ell,i,\rho} - t_{\ell-1,i,\rho}$ and $t_{\ell,i,\rho} - t_{\ell-1,i+1,\rho}$ of every node (ℓ, i), $\ell > 0$, relative to its direct layer $\ell - 1$ neighbors $(\ell - 1, i)$ resp. $(\ell - 1, i + 1)$.

We remark that the former is defined in terms of the absolute values due to the symmetry of the topology (and thus skews) within a layer, whereas the latter respects the sign and thus correctly captures the non-zero bias (of at least d^-) in the inter-layer neighbor skew.

To support the claims (C1) the observation of average resp. median compared to the maxima is enough. However, in combination with claim (C2), the tightness, so to say the majority of the skews is of relevance. Hence, we introduce the *inter-quantile distance at level p* (IQD^p). The inter-quantile distance of a level p is defined as the difference between the quantiles $1 - p$ and p, i.e., $q_{1-p} - q_p$. Thus, we consider here also the values relevant to the IQD^1 and IQD^5, as there are good measures for the tightness of a distribution.

For op $\in \{q_1, q_5, \text{avg}, q_{50}, q_{95}, q_{99}, \max\}$, we define the 1%-quantile, 5%-quantile, average, median, 95%-quantile, 99%-quantile, and maximum (absolute)

- *layer ℓ intra-layer skew in run ρ*:

$$\sigma^{\text{op}}_{\ell,\rho} := \text{op}_{i \in [W]} \{|t_{\ell,i,\rho} - t_{\ell,i+1,\rho}|\}; \tag{3.28}$$

- *intra-layer skew in run ρ*:

$$\sigma^{\text{op}}_{\rho} := \text{op}_{i \in [W], \ell \in [L+1] \setminus \{0\}} \{|t_{\ell,i,\rho} - t_{\ell,i+1,\rho}|\}; \tag{3.29}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

- *intra-layer skew in simulation set* R:

$$\sigma^{\mathrm{op}} := \mathrm{op}_{i\in[W],\ell\in[L+1]\setminus\{0\},\rho\in R}\{|t_{\ell,i,\rho} - t_{\ell,i+1,\rho}|\}. \tag{3.30}$$

Further, with $T_{\ell,i,\rho} := \{t_{\ell,i,\rho} - t_{\ell-1,i,\rho}, t_{\ell,i,\rho} - t_{\ell-1,i+1,\rho}\}$ defining the inter-layer neighbors of a node, for $\mathrm{op} \in \{\min, q_1, q_5, \mathrm{avg}, q_{50}, q_{95}, q_{99}, \max\}$ we define the minimal, 1%-quantile, 5%-quantile, average, median, 95%-quantile, 99%-quantile, and maximum (signed)

- *inter-layer skew between layer* ℓ *and* $\ell - 1$ *in run* ρ:

$$\hat{\sigma}^{\mathrm{op}}_{\ell,\rho} := \mathrm{op}_{i\in[W]}\, T_{\ell,i,\rho}; \tag{3.31}$$

- *inter-layer skew in run* ρ:

$$\hat{\sigma}^{\mathrm{op}}_{\rho} := \mathrm{op}_{i\in[W],\ell\in[L+1]\setminus\{0\}}\, T_{\ell,i,\rho}; \tag{3.32}$$

- *inter-layer skew in simulation set* R:

$$\hat{\sigma}^{\mathrm{op}} := \mathrm{op}_{i\in[W],\ell\in[L+1]\setminus\{0\},\rho\in R}\, T_{\ell,i,\rho}. \tag{3.33}$$

Statistical evaluation of the stabilization time These experiments require multiple pulses. Essentially, the system is started with every node in an arbitrary state, and then attempts to forward a sequence of pulses generated at layer 0 (with bounded skew and a certain separation time). Using post-processing of the recorded triggering times, we compute the stabilization time as the number of pulses needed for the intra- and inter-layer skews to persistently fall below a layer-dependent threshold.

Both types of experiments were performed with and without faulty nodes of different types. Note that the triggering times of faulty nodes are of course not considered when computing the inter- and intra-layer skews.

3.4.2 Simulation Results: Fault-free Case

We conducted a suite of simulations that complement the analytic intra- and inter-layer worst-case skew bounds given in Theorem 3.9 by statistical data. To visualize the progression of a pulse wave through a grid, we provide plots like Figure 3.22. There, the grid (sliced between width $W-1$ and $0 \equiv W$ for readability, and truncated to the first 30 layers) lies in the $(\ell \in [L+1], i \in [W])$ plane and the z-dimension shows the triggering time $t_{\ell,i}$ of the corresponding node (ℓ, i). To further improve readability, we connected all points $(\ell, i, t_{\ell,i})$ and $(\ell, i+1, t_{\ell,i+1})$, $i \in \{0, \dots, W-2\}$.

First, Figure 3.22 resp. Figure 3.23 show a 3D plot of a typical pulse wave propagation in a fault-free grid with $L = 50$ and $W = 20$, and layer 0 skews all 0 resp. ramping up/down by d^+. It is apparent that the wave propagates evenly throughout the grid, nicely smoothing out the initial skew differences. The end-to-end delays in

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

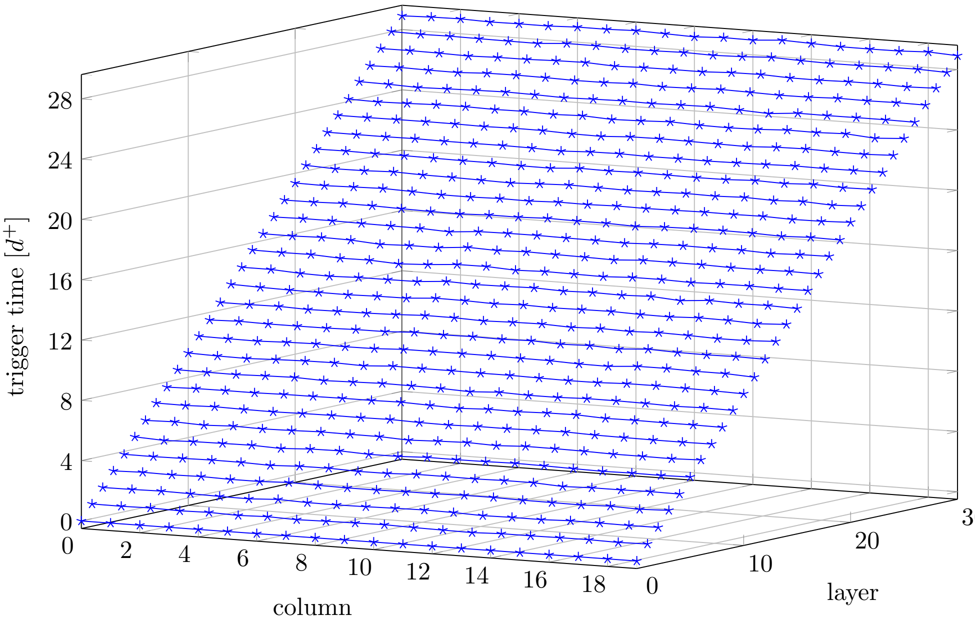


Figure 3.22: Pulse wave propagation with layer 0 skews of 0.

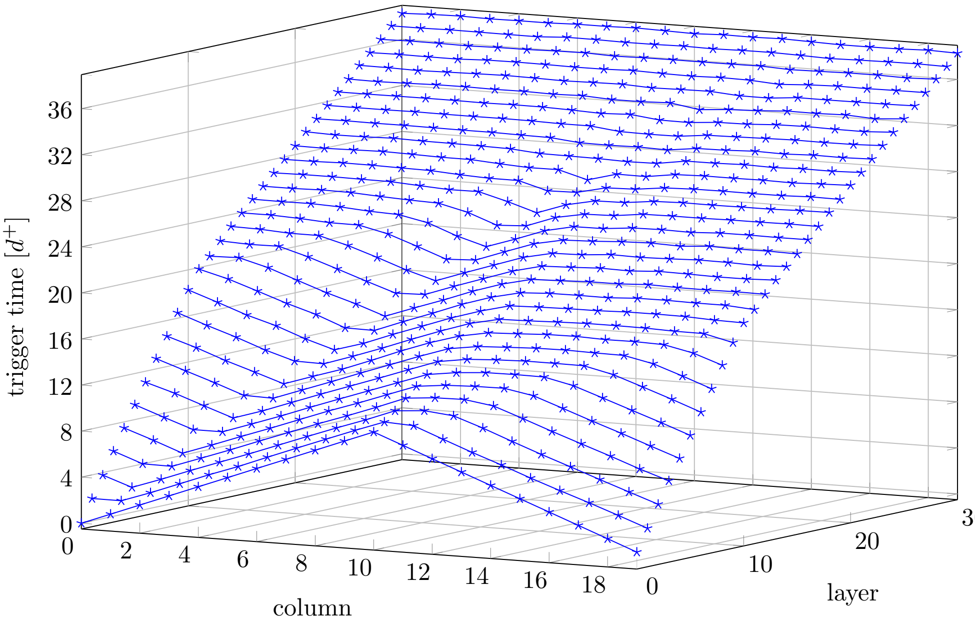


Figure 3.23: Pulse wave propagation with layer 0 skews ramping up/down by d^+.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

$[7.161, 8.197]$ ns ($\varepsilon = 1.036$ ns) result from combining the assumed link wire delays with the processing delay of the HEX node. The latter was determined in Section 3.3.4, to lie within $[0.161, 0.197]$ ns. For the link wire delays, we more or less arbitrarily assumed a value within $[7, 8]$ ns. Since the absolute values do not really matter for our simulations, our choice just reflects the facts that (i) communication delays dominate switching delays and (ii) that the delay uncertainty ε is not too large compared to d^- in modern hardware devices. Throughout this section, we hence assume delay values that are uniformly distributed within $[d^-, d^+] = [7.161, 8.197]$ ns.

To observe the behavior of the grid, four sets have been defined by their choice for the layer 0 skews. The choices for the triggering times for the layer 0 nodes $t_{0,i}$ are:

(i) all 0, resulting in $\sigma_0 = 0$ and skew potential $\Delta_0 = 0$.

(ii) uniformly in $[0, d^-]$, i.e., $\sigma_0 \approx d^-$ and $\Delta_0 = 0$.

(iii) uniformly in $[0, d^+]$, i.e., $\sigma_0 \approx d^+$ and $\Delta_0 \approx \varepsilon$.

(iv) ramping-up/down by d^+: $t_{0,i+1} = t_{0,i} + d^+$ for $0 \leq i < W/2$ and $t_{0,i+1} = t_{0,i} - d^+$ for $W/2 \leq i < W - 1$, i.e., $\sigma_0 = d^+$ and $\Delta_0 \approx W\varepsilon/2 = 10.36$.

Note that the sets (iii) resp. (iv) reasonably model the average case and worst-case input provided by a layer 0 clock generation scheme with neighbor skew bound d^+, respectively.

Table 3.1 shows 1%-quantile (σ^{q_1}) ,5%-quantile (σ^{q_5}), average (σ^{avg}), median ($\sigma^{q_{50}}$), 95%-quantile ($\sigma^{q_{95}}$), 99%-quantile ($\sigma^{q_{99}}$), and maximal ($\sigma^{\max}$) intra-layer and minimal ($\hat{\sigma}^{\min}$), 1%-quantile ($\hat{\sigma}^{q_1}$), 5%-quantile ($\hat{\sigma}^{q_5}$), average ($\hat{\sigma}^{\text{avg}}$), median ($\hat{\sigma}^{q_{50}}$), 95%-quantile ($\hat{\sigma}^{q_{95}}$), 99%-quantile ($\hat{\sigma}^{q_{99}}$), and maximal ($\hat{\sigma}^{\max}$) inter-layer skews, respectively, in the absence of faulty nodes and taken over all nodes and 250 simulation runs, in the sets described above. Inspecting Table 3.1 reveals that not a single instance in the collected data showed a skew $\sigma^{\max} > d^+ = 8.197$ ns resp. $\hat{\sigma}^{\max} > 2d^+ = 16.394$ ns. In sets (i) to (iii), $\hat{\sigma}^{\min} \approx d^-$, i.e., all nodes were always triggered by their lower neighbors (obviously, this property is violated in set (iv) due to the excessive initial skews). A comparison with the worst-case results of Theorem 3.9, which bound $\sigma^{\max} \leq 21.63$ ns and $[\hat{\sigma}^{\min}, \hat{\sigma}^{\max}] \subseteq [-14.47, 29.83]$ ns for sets (i) and (ii), reveals a much better typical skew in every set.

The histograms of the skew distributions of sets (i), (ii), and (iii) are shown in Figures 3.24 to 3.26 respectively. One observes a sharp concentration with an exponential tail, and the similarity between the sets (i), (ii), and (iii). Only in set (iv), as already indicated by the large values of q_{95} in Table 3.1, there is a visible cluster near the end of the tail that is again caused by the large initial skews, see Figure 3.27.

Given the considerable differences between the minimum ($\hat{\sigma}^{\min}$) and maximum ($\hat{\sigma}^{\max}$) inter-layer skew in Table 3.1, in conjunction with its non-zero bias, the question of layer-dependence arises. Figure 3.28 provides $\hat{\sigma}_\ell^{\min}$, $\hat{\sigma}_\ell^{\text{avg}}$ and $\hat{\sigma}_\ell^{\max}$, along with their standard deviations, over the layers $\ell \in [30] \setminus \{0\}$ in case of set (iii) resp. (iv). The diagrams reveal that the fairly discrepant skews observed in lower layers start to smooth

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

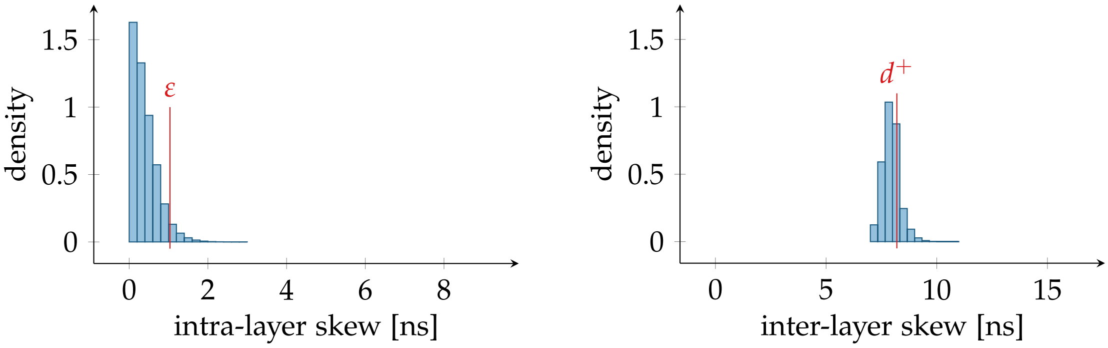


Figure 3.24: Cumulated skew histograms, from 250 simulation runs in set (i).

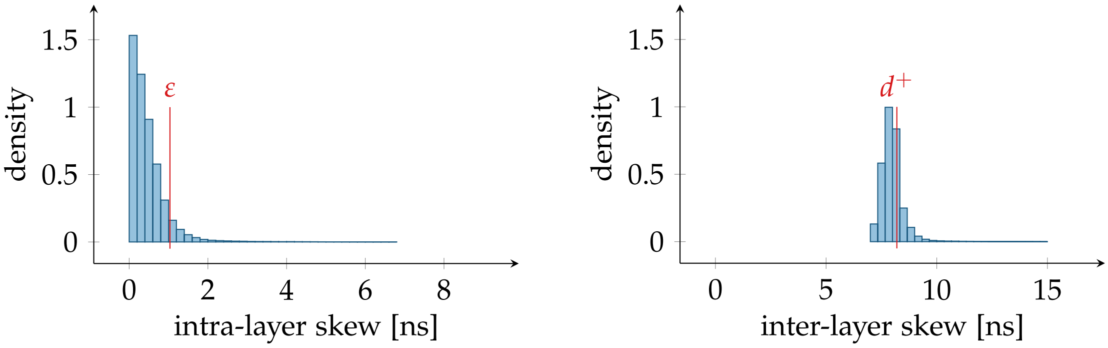


Figure 3.25: Cumulated skew histograms, from 250 simulation runs in set (ii).

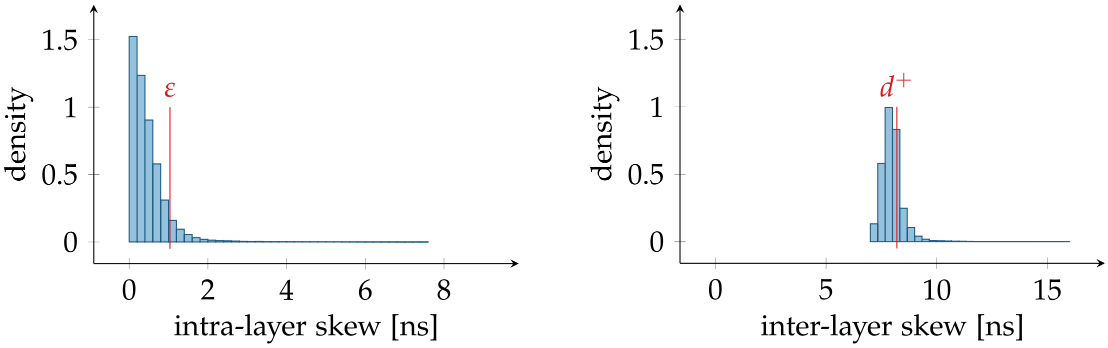


Figure 3.26: Cumulated skew histograms, from 250 simulation runs in set (iii).

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

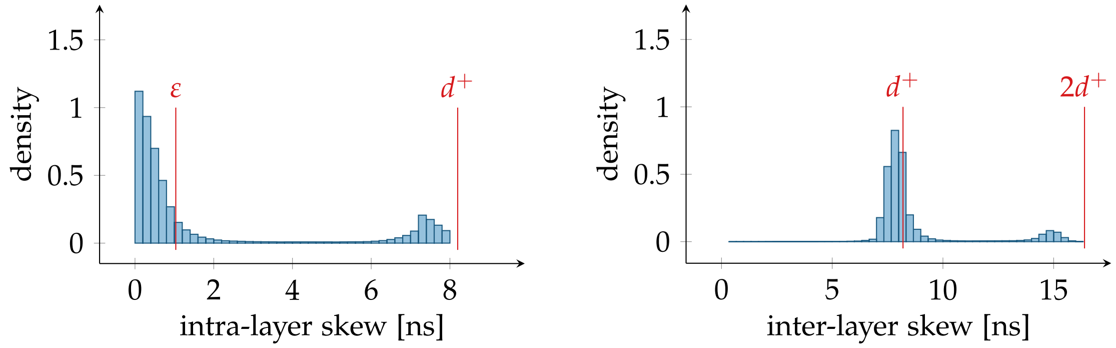
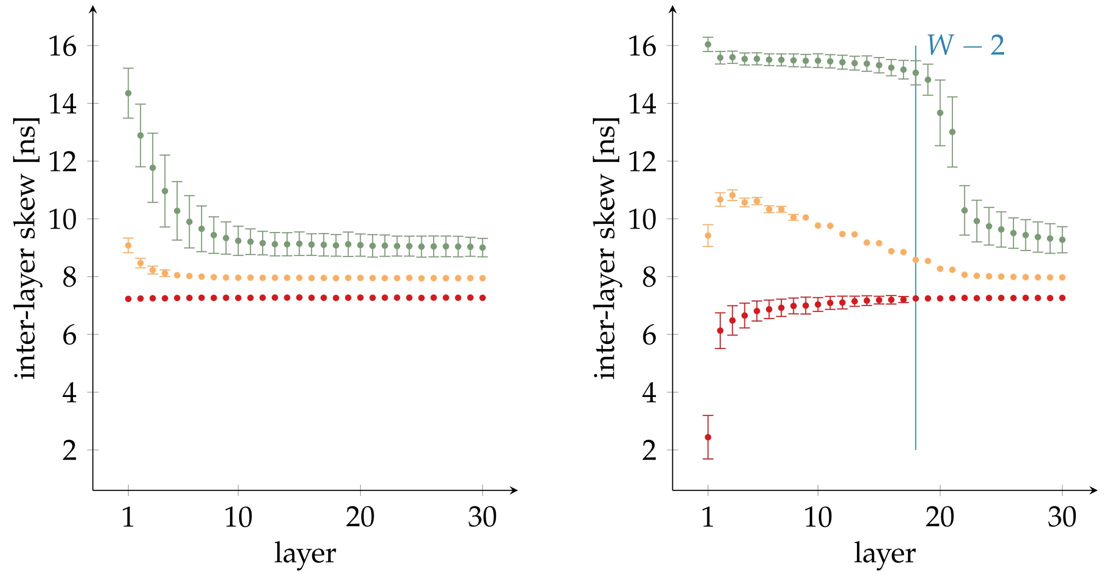


Figure 3.27: Cumulated skew histograms, from 250 simulation runs in set (iv).



(a) Set (iii)

(b) Set (iv)

Figure 3.28: Visualization of the inter-layer skews in set (iii) and (iv), truncated to 30 layers. Averages and standard deviations over 250 simulation runs are plotted on a per-layer basis, to avoid clutter the standard deviation was dropped if it was below 0.1 ns. The top data series (•) shows $\hat{\sigma}_\ell^{\max}$, the middle (•) $\hat{\sigma}_\ell^{\text{avg}}$, and the lower (•) $\hat{\sigma}_\ell^{\min}$. Note that the reduction of the maximal skew after layer $W-2$ in set (iv) is in accordance with the vanishing initial skew dependence predicted by Lemma 3.6.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 3.1: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in 250 simulation runs without faults on a 50 × 20 grid in the scenarios (i) to (iv).

	intra-layer σ^{op}							inter-layer $\hat{\sigma}^{\text{op}}$							
Set	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
(i)	0.01	0.03	0.39	0.33	1.00	1.42	3.10	7.16	7.23	7.36	7.94	7.93	8.63	9.06	11.03
(ii)	0.01	0.03	0.46	0.35	1.23	2.27	6.89	7.16	7.23	7.35	7.99	7.94	8.80	9.97	15.20
(iii)	0.01	0.03	0.47	0.35	1.26	2.52	7.79	7.16	7.22	7.35	8.00	7.94	8.81	10.23	16.22
(iv)	0.01	0.04	1.86	0.52	7.64	8.01	8.19	0.36	6.53	7.26	8.64	7.96	14.83	15.39	16.39

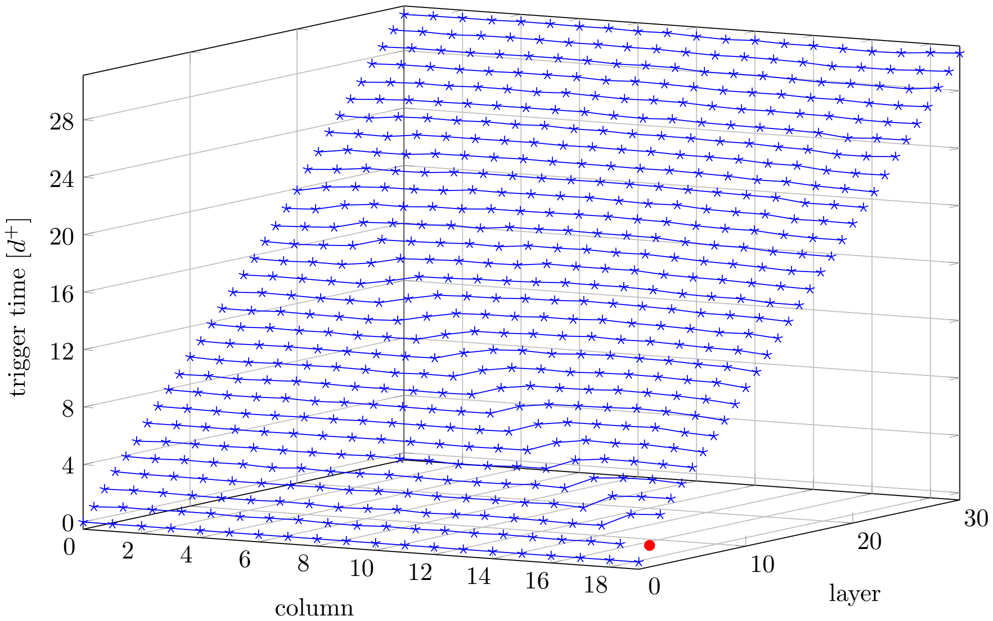


Figure 3.29: Pulse propagation for set (i), with one Byzantine node (marked as red dot) at $(1, 19)$. It sends a constant 1 to its left and right neighbors, and a constant 0 to both upper-layer neighbors. Observe how the increase in skews emanating from the faulty node fades with the distance from the fault location.

out after layer $W - 2$, in accordance with Lemma 3.6, which shows that the behavior observed in Figure 3.23 is very typical. To avoid cluttering the diagrams, we omitted $\hat{\sigma}_\ell^{q_5}$ and $\hat{\sigma}_\ell^{q_{95}}$, which are close to $\hat{\sigma}_\ell^{\min}$ resp. $\hat{\sigma}_\ell^{\max}$.

3.4.3 Simulation Results: Failures

Next, we back up the results of our analysis in Section 3.2.4: We consider f uniformly placed failures under the constraint that Condition 3.12 holds. In each run, each Byzantine node randomly selects its behavior on each outgoing link as either constant 0 (fail-silent) or constant 1. A typical pulse with a single faulty node $(1, 19)$, marked as a red dot, is shown in Figure 3.29; Figure 3.30 depicts another example pulse with 5 Byzantine faulty nodes.

In Table 3.2, we list the statistical results for the case of $f = 1$ Byzantine fault. Additionally, we included the parameter h, which defines the outgoing h-hop neighborhood of a fault that is discarded in the skew analysis. That is, for $h = 0$ the faulty node itself is discarded, for $h = 1$ the outgoing neighbors of the fault are also discarded. One observes that the behavior for sets (i) to (iii) is very similar, same as it was for the fault-free case. Hence, we will showcase the main points by examining sets (iii) and (iv) more closely.

In Figure 3.31a resp. 3.31c, we show box plots of minimum, 5%-quantile, average,

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

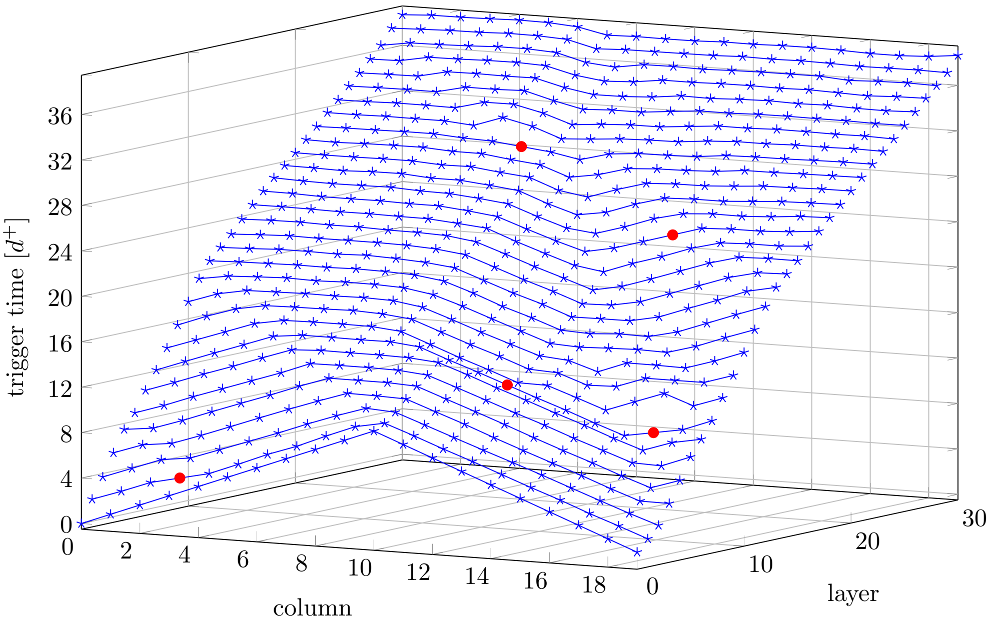


Figure 3.30: Pulse propagation for set (iv), with five Byzantine nodes (marked as red dots).

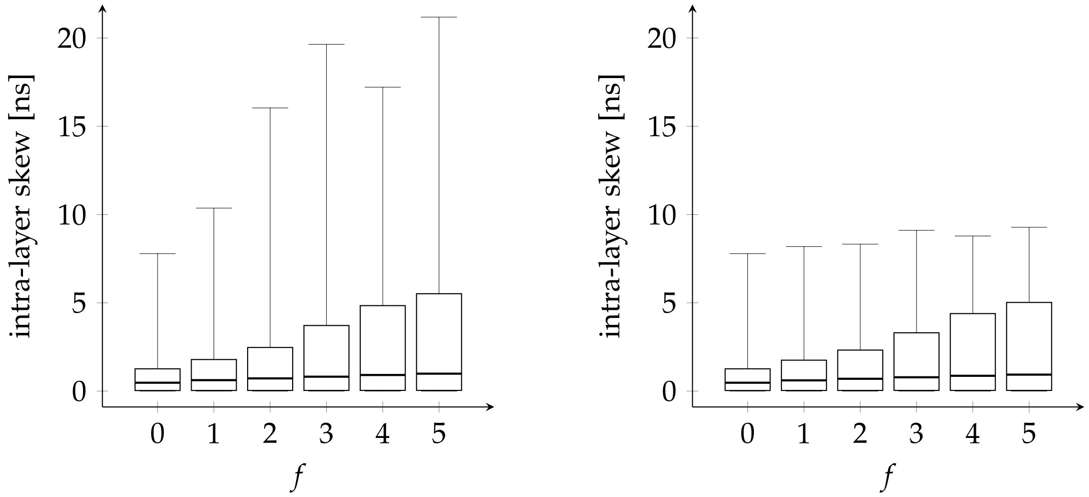
95%-quantile, and maximum intra- resp. inter-layer skews from 250 runs with $f \in [6]$ faults for set (iii).[12] A comparison of values $f > 0$ with $f = 0$ reveals that skews increase moderately. In particular, the skews increase substantially slower than the derived worst-case bound of roughly $5fd^+$. Furthermore, Figures 3.31b and 3.31d show the same data with $h = 1$. Here, a comparison of the case $f = 0$ and any $f > 0$ reveals that all fault effects have essentially disappeared or are mitigated notably, which confirms claim (C2): HEX exhibits a strong fault-locality. This has already been visible in Table 3.2, where with $h = 1$ the IQD$^\text{p}$ reduced for all level (0, 1, and 5). Concerning fail-silent nodes, all results are qualitatively similar, albeit with smaller skews as can be seen in Table 3.3.

Figure 3.32 gives the same plots for set (iv). Apart from the expected increase in skews, we observe two points worth mentioning: First, a single fault can essentially causes the "worst-case" skew. This demonstrates that, as already indicated by set (iii), skew effects of multiple faults do not accumulate, or do so in a very limited way. Second, the maximal intra-layer skews typically *exceed* the inter-layer skews. An explanation for this behavior is provided in Figure 3.33: Intuitively, for "ramped" triggering times generated at layer 0, the pulse propagates "diagonally" (cf. Figure 3.23) instead of "vertically" (= layer by layer). This limits the power of the HEX grid to mitigate Byzantine behavior, as it is implicitly optimized for propagating pulse waves vertically.

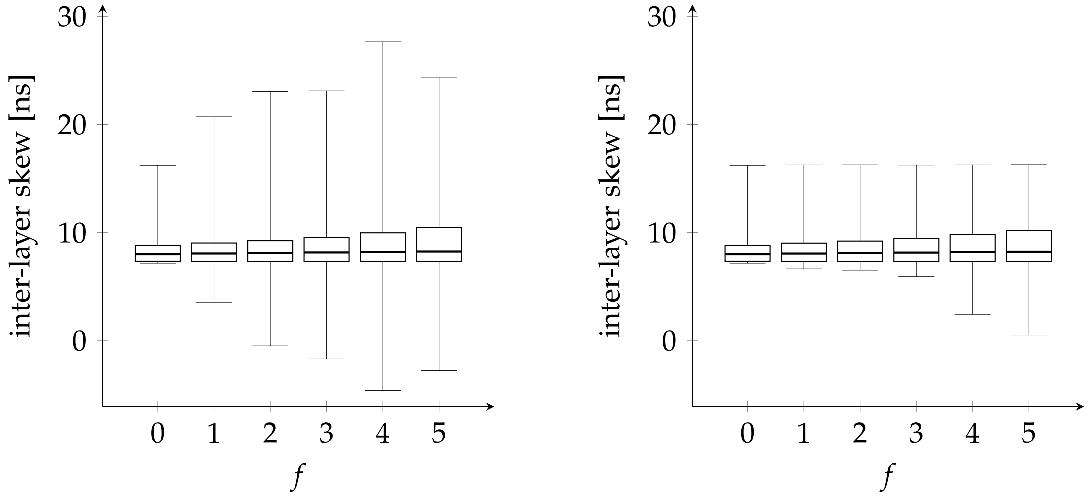
---

[12] Note that for the absolute values of the intra-layer skew, the minimal and 5%-quantile values are close to zero and, thus, of little relevance.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



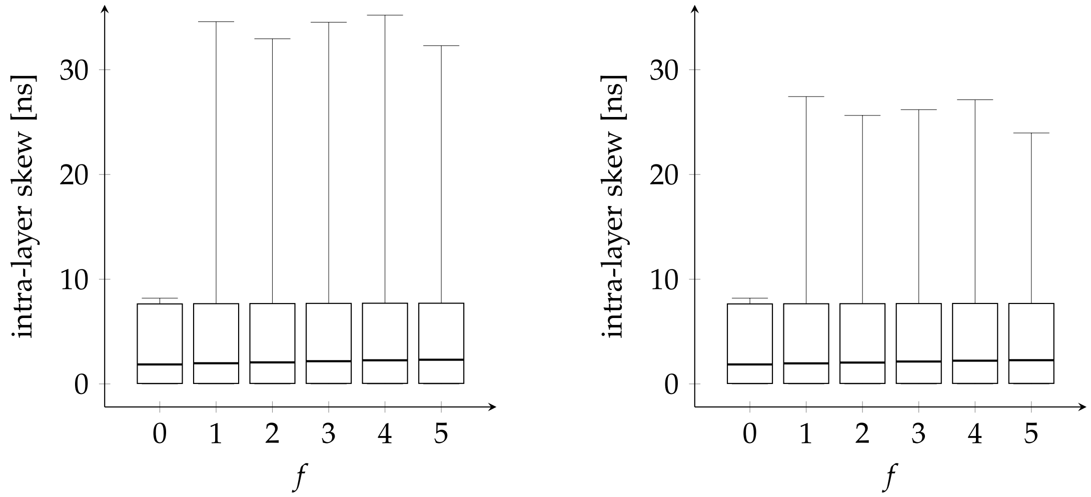
(a) intra-layer skew, $h = 0$ (b) intra-layer skew, $h = 1$



(c) inter-layer skew, $h = 0$ (d) inter-layer skew, $h = 1$

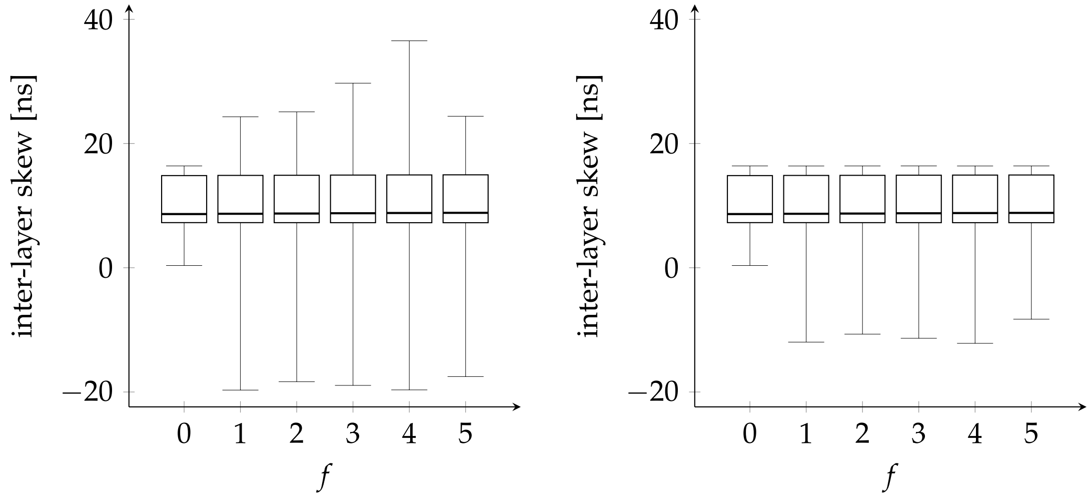
Figure 3.31: Box plots of inter-layer and intra-layer skews $\hat{\sigma}^{\mathrm{op}}$ from 250 runs in set (iii), with h-hop outgoing neighbors of the f Byzantine faulty nodes removed. Note that the whiskers represent the minimum/maximum, and the box spans over the IQD[5].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skew, $h = 0$

(b) intra-layer skew, $h = 1$



(c) inter-layer skew, $h = 0$

(d) inter-layer skew, $h = 1$

Figure 3.32: Box plots of inter-layer and intra-layer skews $\hat{\sigma}^{\text{op}}$ from 250 runs in set (iv), with h-hop outgoing neighbors of the f Byzantine faulty nodes removed. Note that the whiskers represent the minimum/maximum, and the box spans over the IQD[5].

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

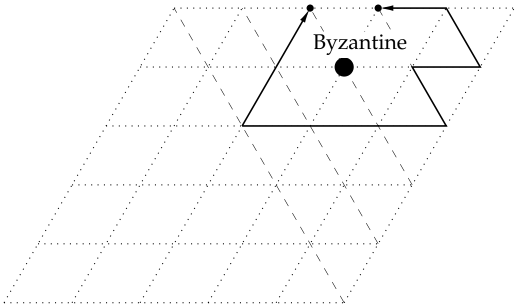


Figure 3.33: A single Byzantine node in set (iv) maximizes the skew between its upper neighbors. All delays are assumed to be d^+ and the triggering times of the nodes in the smallest layer are increasing from left to right, by d^+ per hop. Hence, in the absence of faults, all nodes on the diagonals in the left-up direction would be triggered at identical times. The marked causal paths originating at the same node illustrate that the generated skew is $5d^+$; the inter-layer skew is smaller by d^+.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 3.2: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\mathrm{op}}$ (in [ns]) in 250 simulation runs on a 50×20 grid with a Byzantine node in the sets (i) to (iv).

		intra-layer σ^{op}							inter-layer $\hat{\sigma}^{\mathrm{op}}$							
Set	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
(i)	0	0.01	0.03	0.54	0.35	1.34	5.85	10.39	5.58	7.23	7.35	8.01	7.93	8.76	11.28	17.55
	1	0.01	0.03	0.53	0.35	1.32	5.53	8.19	6.58	7.23	7.35	8.00	7.93	8.75	10.89	16.25
(ii)	0	0.01	0.03	0.61	0.38	1.72	5.85	10.12	4.21	7.22	7.34	8.06	7.94	9.00	12.21	20.03
	1	0.01	0.03	0.60	0.38	1.68	5.52	8.19	6.64	7.22	7.35	8.05	7.94	8.99	12.02	16.26
(iii)	0	0.01	0.03	0.62	0.38	1.79	5.93	10.36	3.52	7.22	7.34	8.07	7.94	9.03	12.48	20.72
	1	0.01	0.03	0.61	0.38	1.75	5.60	8.19	6.64	7.22	7.34	8.06	7.94	9.02	12.28	16.26
(iv)	0	0.01	0.05	1.97	0.56	7.66	8.02	34.59	−19.70	6.46	7.26	8.69	7.96	14.87	15.43	24.31
	1	0.01	0.05	1.96	0.56	7.65	8.01	27.44	−11.97	6.48	7.26	8.69	7.96	14.86	15.42	16.38

Table 3.3: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\mathrm{op}}$ (in [ns]) in 250 simulation runs on a 50×20 grid with a fail-silent node in the sets (i) to (iv).

		intra-layer σ^{op}							inter-layer $\hat{\sigma}^{\mathrm{op}}$							
Set	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
(i)	0	0.01	0.03	0.58	0.36	1.52	6.38	8.19	6.43	7.23	7.35	8.03	7.93	8.82	12.62	17.55
	1	0.01	0.03	0.57	0.35	1.48	6.07	8.16	6.57	7.23	7.35	8.02	7.93	8.81	12.22	16.25
(ii)	0	0.01	0.03	0.65	0.39	1.95	6.40	8.19	6.38	7.22	7.34	8.08	7.95	9.09	12.89	18.86
	1	0.01	0.03	0.64	0.39	1.89	6.06	8.16	6.55	7.22	7.34	8.08	7.95	9.08	12.64	16.26
(iii)	0	0.01	0.03	0.66	0.39	2.05	6.43	8.19	6.38	7.22	7.34	8.09	7.95	9.13	13.08	19.34
	1	0.01	0.03	0.65	0.39	1.98	6.11	8.16	6.55	7.22	7.34	8.08	7.95	9.11	12.86	16.26
(iv)	0	0.01	0.05	2.02	0.58	7.66	8.02	27.66	−13.62	6.49	7.26	8.72	7.96	14.88	15.44	24.31
	1	0.01	0.05	2.00	0.57	7.65	8.01	20.94	−5.95	6.50	7.26	8.71	7.96	14.87	15.42	16.38

Table 3.4: Assumed stable skews σ and corresponding timeout values (in [ns]) used in stabilization experiments.

Set	σ	$T^-_{\mathbf{link}}$	$T^+_{\mathbf{link}}$	$T^-_{\mathbf{sleep}}$	$T^+_{\mathbf{sleep}}$	$\mathcal{S}$
(i)	28.48	31.98	33.58	83.56	87.74	264.08
(ii)	31.16	34.66	36.39	89.18	93.64	275.60
(iii)	31.75	35.25	37.01	90.42	94.94	278.14
(iv)	40.64	44.14	46.34	109.08	114.53	316.40

3.4.4 Simulation Results: Self-stabilization

We now present the results of the multi-pulse simulations conducted for evaluating stabilization time statistics. The same sets as in our single-pulse experiments were used here. For the timeouts and the pulse separation time $\mathcal{S}$, we used nominal values that are compatible with the (set-dependent) maximum skew observed for $f \in [6]$ Byzantine or fail-silent nodes. These skews where determined via the previous simulations, plus a slack of d^+ accounting for the fact that we work with a statistical sample showing an exponential tail. Timeouts and pulse separation times were computed according to (a modified[13] version of) Condition 3.15, assuming $\vartheta = 1.05$. The results are shown in Table 3.4.

For each combination of set, fault-number, and fault-type, we executed 250 simulation runs. For each run, f faulty nodes were placed uniformly at random under the constraint that Condition 3.12 held. Then, starting with all non-faulty nodes in random initial states, 10 consecutive pulses were generated, where placement and behavior of faulty nodes remained fixed during an individual run. For each run ρ, the firing times $t^k_{\ell,i,\rho}$ of all pulses $1 \leq k \leq 10$ were recorded.

The stabilization time estimate for each run is computed (off-line) as the minimal pulse k with the property that the maximal layer ℓ intra- resp. inter-layer skew, for every layer $\ell \in [L+1]$, is below the a priori chosen skew bound $\hat{\sigma}(f,\ell)$ resp. $\sigma(f,\ell)$. As the former directly depends on the latter (cf. Theorem 3.9), we used 4 different choices $C \in \{0,1,2,3\}$ for the skew bound, obtained by setting $\sigma(f,\ell) = (4-C)d^+$ for $C \in \{1,2,3\}$ resp. the very conservative skew bounds resulting from Lemma 3.14 for $C = 0$.

Due to the fact that our simulations only cover the first 10 pulses, it could occur that some runs do not stabilize (yet). Even worse, we cannot rule out the possibility that a run which seems stable violates the skew bounds in some later pulse. However, our actual experiments showed that non-stabilizing runs occur only in scenarios in which the a priori chosen skew bound $\sigma(f,\ell)$ was smaller than the maximal skews reported in Table 3.2, i.e., where $\sigma(f,\ell)$ was too small. We are hence confident that the stabilization time estimates can be considered representative for realistic skew bounds.

---

[13] Condition 3.15 does not take into account that triggering signals in our HEX implementation have non-zero duration, resulting in slightly increased values.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Figure 3.34a resp. Figure 3.35a shows both the average and the average + standard deviation of the stabilization times for set (iii) resp. (iv) for Byzantine faults. Similarly, Figure 3.34b resp. Figure 3.35b shows set (iii) resp. (iv) in case of fail-silent faults. By and large, unless C is chosen aggressively large, resulting in too small $\sigma(f,\ell)$, HEX usually stabilizes after the very first pulse. For large C, the average stabilization time estimates go up moderately. Given that the number of simulation runs that did not stabilize within 10 pulses was low ($< 25\%$ even in the most unfavorable scenario), however, we can safely infer that the typical stabilization time of HEX is indeed much lower than the worst-case stabilization time bound $L+1$ established in Theorem 3.17. The results verify that the already good self-stabilization properties of the original HEX algorithm reported in [PSSL13] are further improved by the additional link timeouts.

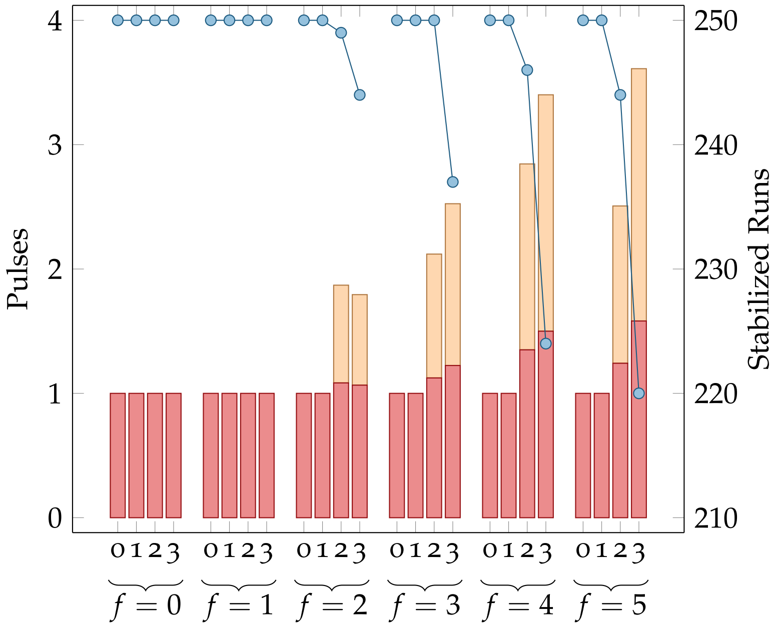
Finally, we also computed the stabilization time estimates for the very same sets after additionally discarding the single-hop outgoing neighbors of faulty nodes ($h=1$). In these cases, shown in Figures 3.36 and 3.37, HEX turned out to stabilize after the first pulse in almost every run. The "artifacts", e.g., the non-zero standard deviation in Figure 3.37 for $f=\{2,3\}$, are expected to appear for a higher number of faults as well with an increase in simulation runs. Nonetheless, this again confirms the strong fault locality property claimed for the HEX grid.

3.5 Fast Clocks

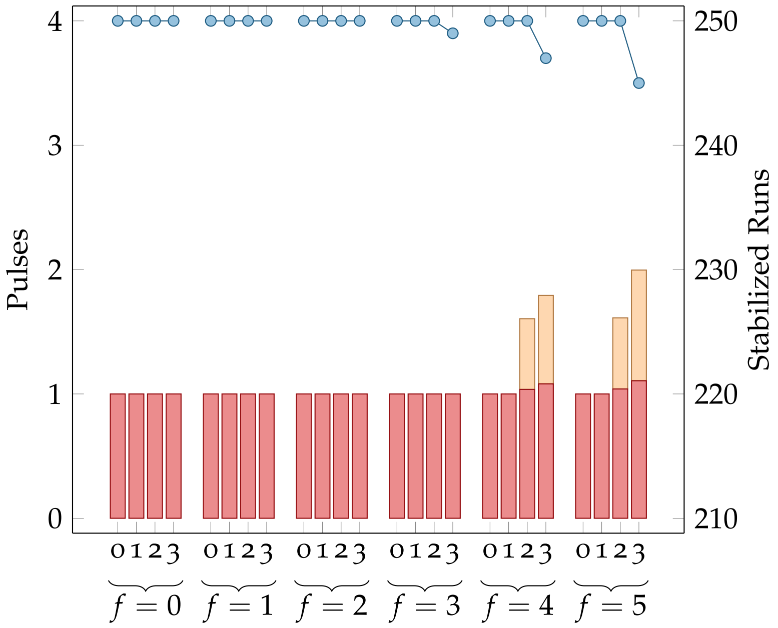
Summarizing the results of the analysis (Section 3.2) and the simulations (Section 3.4), HEX provides each node with a local clock signal that is well-synchronized even in a system with multiple persistent faults. Furthermore, HEX is able to recover from an unbounded number of concurrent transient faults. It should come as no surprise that these impressive fault-tolerance properties come at a cost: HEX requires a fairly low input frequency from its clock source to ensure that the nodes become ready for the next, well-synchronized pulse. Hence, a conservative pulse separation time must be granted to "flush out" spurious pulses from the system; otherwise, we might observe a phenomenon similar to "ventricular fibrillation". Furthermore, the locally generated HEX clock signal has the same low frequency as the clock source, which, however, suffers from a large jitter, due to the skew. This makes HEX unsuitable for applications that need to determine real-time durations very accurately. One such negative example is clock multiplication based on PLLs, which can sustain only moderate input jitter. For a more thorough introduction into clock multiplication, we refer to Section 2.4.4. Fortunately, as will be shown in Chapter 6, many applications do not have such stringent requirements.

However, these undesirable properties of HEX are inherent to this type of clock distribution system, if not even inevitable under the given design goals. Recall that Byzantine or fail-silent nodes locally affect the triggering times by (i) triggering a pulse early, or (ii) cutting off or delaying the clock distribution signal on a (shortest) path to some node. Because of (i), a node cannot locally trigger a pulse just based on one of its lower neighbors. Due to (ii), a faulty lower-left or lower-right neighbor entails that the

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



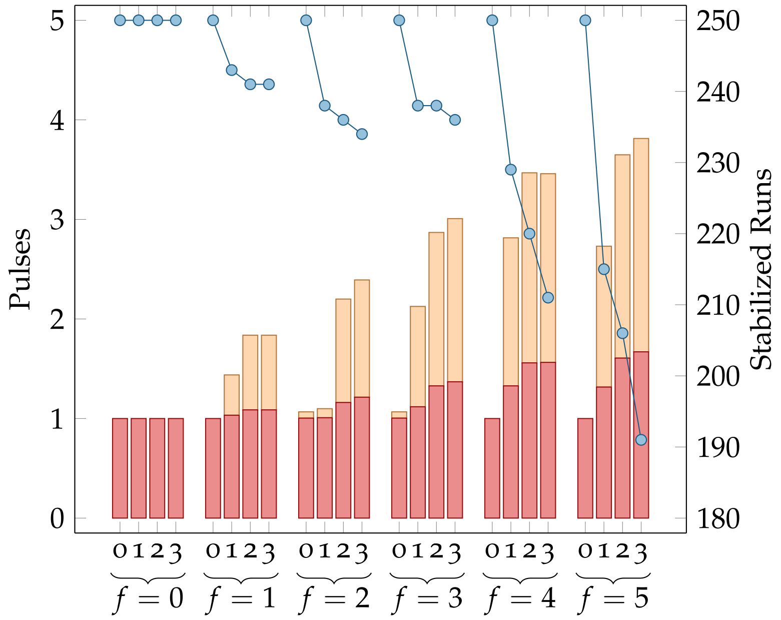
(a) Byzantine faults



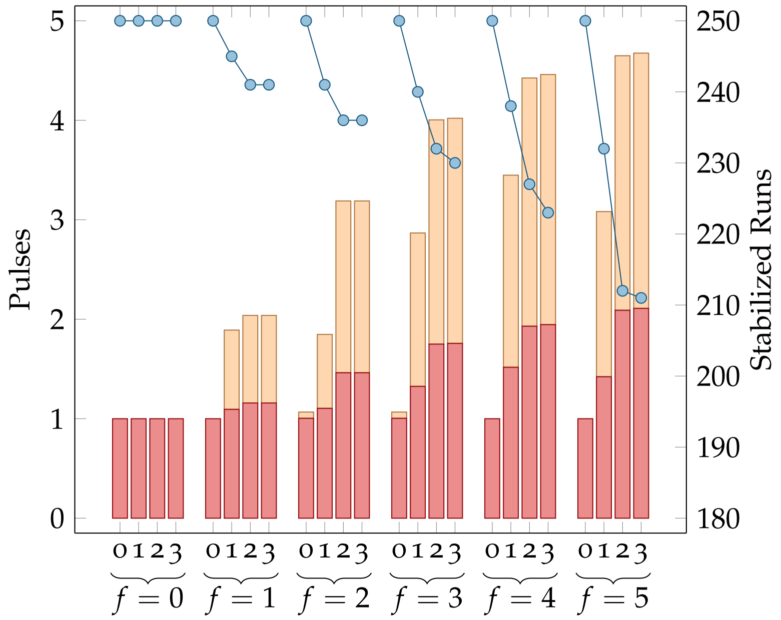
(b) Fail-silent faults

Figure 3.34: Plot of the average (🟥), the average + standard deviation (🟧) of the estimated stabilization time and the number of stabilized runs (—●—), under set (iii) in 250 10-pulse simulations. For each number f of faulty nodes, $C \in \{0, \ldots, 3\}$ encodes the (decreasing) choice of $\sigma(f, \ell)$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



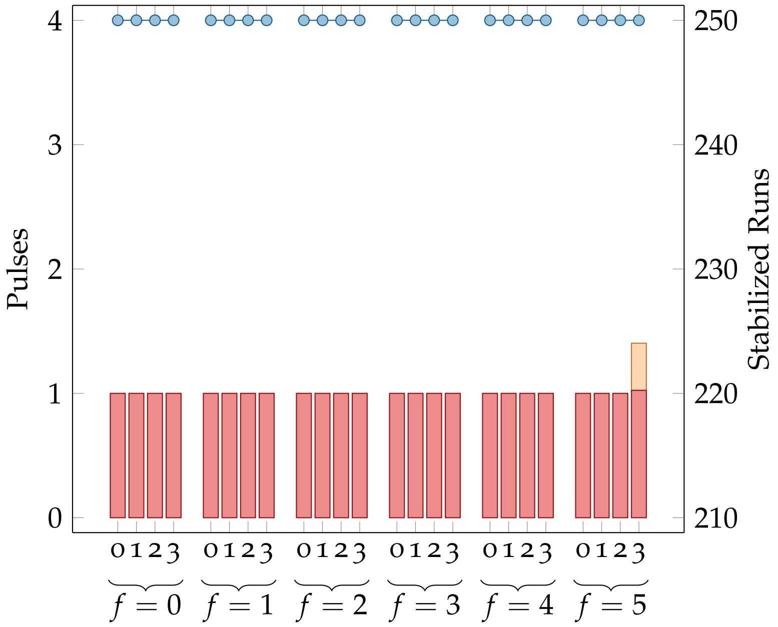
(a) Byzantine faults



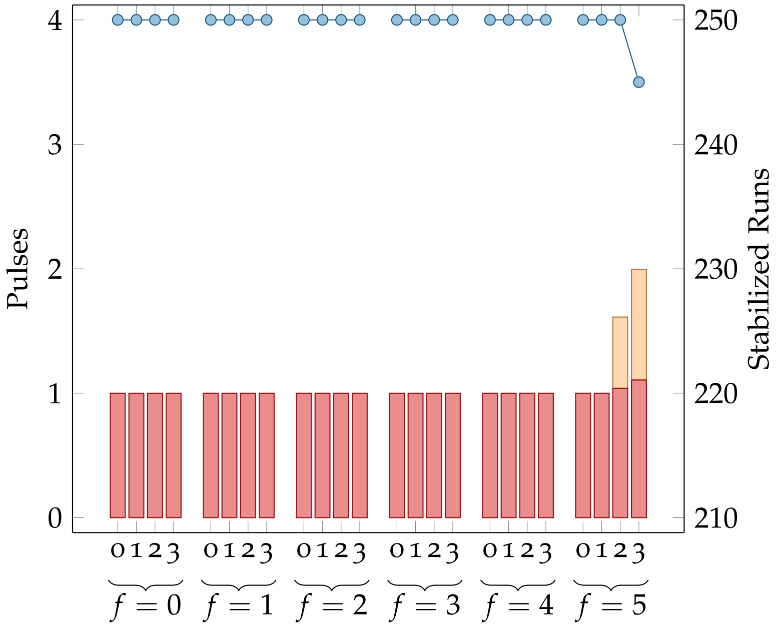
(b) Fail-silent faults

Figure 3.35: Plot of the average (🟥), the average + standard deviation (🟧) of the estimated stabilization time and the number of stabilized runs (—●—), under set (iv) in 250 10-pulse simulations. For each number f of faulty nodes, $C \in \{0, \ldots, 3\}$ encodes the (decreasing) choice of $\sigma(f, \ell)$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



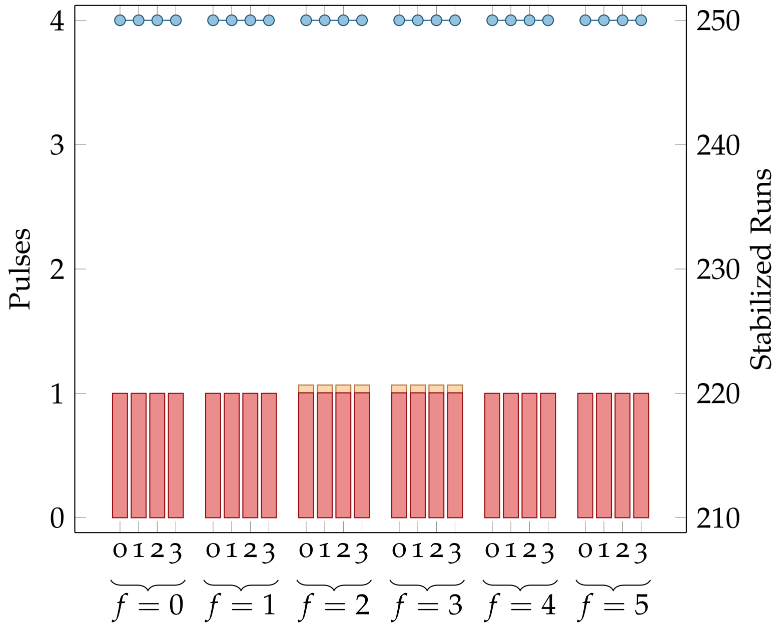
(a) Byzantine faults



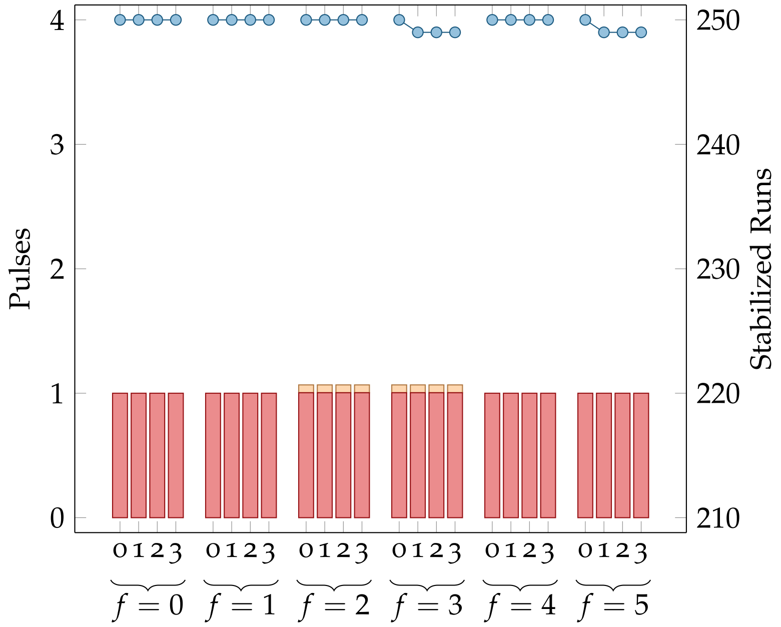
(b) Fail-silent faults

Figure 3.36: Plot of the average (🟥), the average + standard deviation (🟧) of the estimated stabilization time and the number of stabilized runs (—●—), under set (iii) in 250 10-pulse simulations. For each number f of faulty nodes, $C \in \{0, \ldots, 3\}$ encodes the (decreasing) choice of $\sigma(f, \ell)$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) Byzantine faults



(b) Fail-silent faults

Figure 3.37: Plot of the average (🟥), the average + standard deviation (🟧) of the estimated stabilization time and the number of stabilized runs (—●—), under set (iv) in 250 10-pulse simulations. For each number f of faulty nodes, $C \in \{0, \ldots, 3\}$ encodes the (decreasing) choice of $\sigma(f, \ell)$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

node must be triggered with the help of a node in the same layer, thereby increasing the length of a causal chain [Lam78] involved in triggering the node compared to the fault-free case. Consequently, the time when the node is triggered may be affected notably, and this effect may accumulate over several layers leading to an increased layer-global difference in triggering times resp. skew potential, e.g., between $(\ell, 0)$ and $(\ell, W/2)$. Albeit, this may only have a minor impact on the skew itself, which is the local difference of triggering times.

From these observations, we can conclude that simultaneously guaranteeing tolerance of Byzantine faults and a stable clock frequency would require a different topology. However, in anticipation of the results of Chapter 4, a stronger connectivity of the grid and thus larger node degrees do not alleviate large skews completely. Furthermore, higher node degrees increase the complexity of nodes—and thus the likelihood that an individual node fails—as well as the probability that two neighbors are faulty (even for a fixed failure probability).

3.5.1 Local Clocks

While increasing the frequency stability of the HEX clock signal is challenging in a general setting (cf. Chapter 4), there is an obvious solution to the low clock frequency issue: frequency multiplication. By equipping each node with a high-frequency oscillator that is synchronized to the HEX clock, one can generate well-synchronized high-frequency clocks (termed *fast clocks* in the sequel).

However, synchronizing a fast clock to an unstable HEX clock involves a trade-off:

(a) If a stable fast clock frequency is desired, the HEX clock's jitter must be amortized over $m > 1$ pulses. For example, this could be implemented by first dividing the HEX clock by m and then using a PLL clock multiplier to generate the desired fast clock; m must be chosen appropriately to guarantee an acceptable PLL input clock jitter. Unfortunately, this approach may increase the skew of the fast clock considerably if the high-frequency oscillators driving the fast clocks have a large drift.

(b) If a fast clock with minimal skew is desired, which is our major objective, the HEX clock jitter must be amortized within a single pulse. Unfortunately, using solutions like [ON03] are complex and inherently lead to considerable frequency fluctuations of the fast clock.

With this in mind, we propose a simple, robust approach for (b) that achieves a high and reasonably stable fast clock frequency with good skews, at the expense of *burstiness*. For each HEX pulse, the fast clock generates a fixed number B of fast clock pulses, similar to [ONM+00], as shown in Figures 3.38 and 3.39.

3.5.1.1 Implementation

The schematic of the used implementation is shown in Figure 3.40. The clock multiplier here consists of a start/stoppable ring oscillator, a cycle counter, and a few logic gates.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

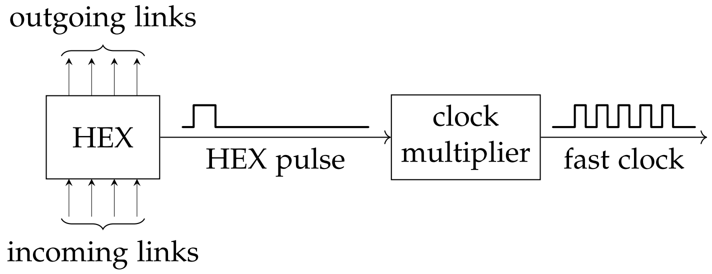


Figure 3.38: The composition of HEX with a clock multiplier. The HEX pulse generated from the HEX node is used by the clock multiplier to generate a fast clock.

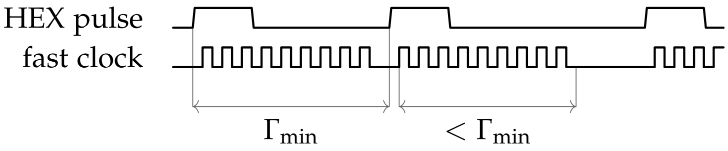


Figure 3.39: A pulse generated by HEX will result in a fixed number of clock cycles being generated. The number of clock cycles and their frequency has to be chosen so that the required time span for them is less than the minimal pulse separation time at correct nodes (Γ_{min}).

The output of the fast clock signal is enabled by the running signal, allowing the cycle counter to control the burst length. For the ring oscillator we may make use of the same oscillator design already used for the timeouts $T_{\mathbf{link}}$ and $T_{\mathbf{sleep}}$ of the HEX state machine (Figure 3.11a). The oscillator is started with the rising edge of a HEX pulse and stops after the completion of the last cycle of the burst, which is determined by the count of the cycle counter: Starting from a counter value of 0, the running signal is enabled. When the predefined counter value, i.e., the burst length B, is reached, running is disabled. Then, the counter is reset to 0 again.

Note that this can be implemented in a way that entirely avoids metastability, assuming that the HEX pulse performs no metastability-inducing behavior. Furthermore, the HEX pulse must be wide enough so that the running signal is able to keep the oscillator running before the HEX pulse's falling transition. See Figure 3.41 for a timing diagram of the clock multiplier, and Figure 3.42 for the result of a timing simulation.

The clock frequency of the ring oscillator, and thus of the fast clock, must be chosen so that the running signal of the cycle counter is stable after half the clock period. Otherwise, the first or last clock cycle of the fast clock could suffer from an incorrect timing that, in the worst case, could cause metastability or other faults in the circuit using the fast clock. Note that this is also the reason why the cycle counter runs with the inverted oscillator clock: Otherwise, the output of the cycle counter could enable the fast clock at some time in the high-phase of the clock cycle of the oscillator.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

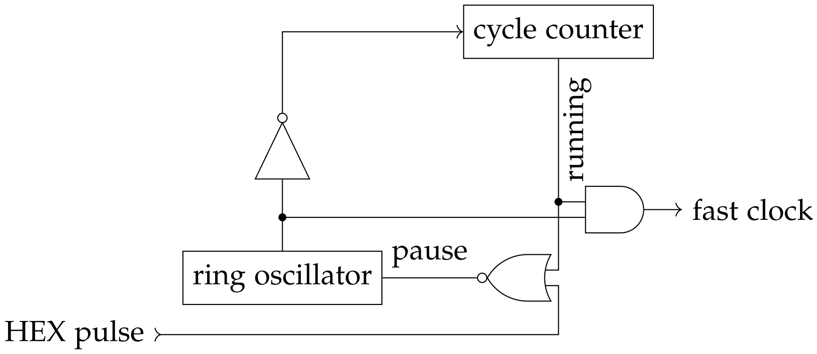


Figure 3.40: The logic elements building the fast clock. The HEX pulse enables the oscillator, which in turn drives the cycle counter. The cycle counter enables the output of the oscillator as the fast clock, for a predefined amount of clock cycles, and stops the oscillator after the burst has been completed.

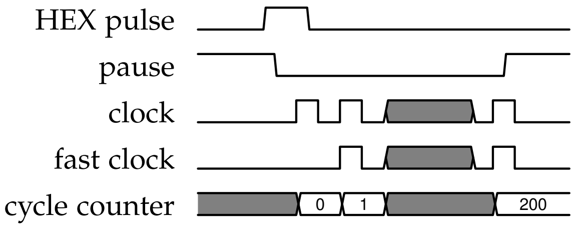


Figure 3.41: Timing diagram of the fast clock generated by the clock multiplier. After the HEX pulse is received, the pause signal is disabled and the ring oscillator starts generating a clock. This clock starts the cycle counter, which then enables the fast clock output. This is done until the predefined value of the cycle counter, in this case 200, is reached. Then, pause will be reenabled and the generation of the fast clock stops.

3.5.2 Analysis

First, we determine conservative timing constraints under the assumption that the system is fault-free. In this setting, the feasible number of clock cycles in each burst can be precisely expressed in terms of the *minimal pulse separation time*

$$\Gamma_{\min} := \inf_{\substack{i\in[W],\ \ell\in[L+1]\setminus\{0\},\ k\in\mathbb{N}\\ (\ell,i)\text{ correct}}} \left\{ t_{\ell,i}^{k} - t_{\ell,i}^{k-1} \right\}. \tag{3.34}$$

Assuming that the fast clock sources are guaranteed to run at least with frequency $f_{\min}$, the number $B \in \mathbb{N}$ of generated ticks per pulse must satisfy

$$B \leq f_{\min}\Gamma_{\min}. \tag{3.35}$$

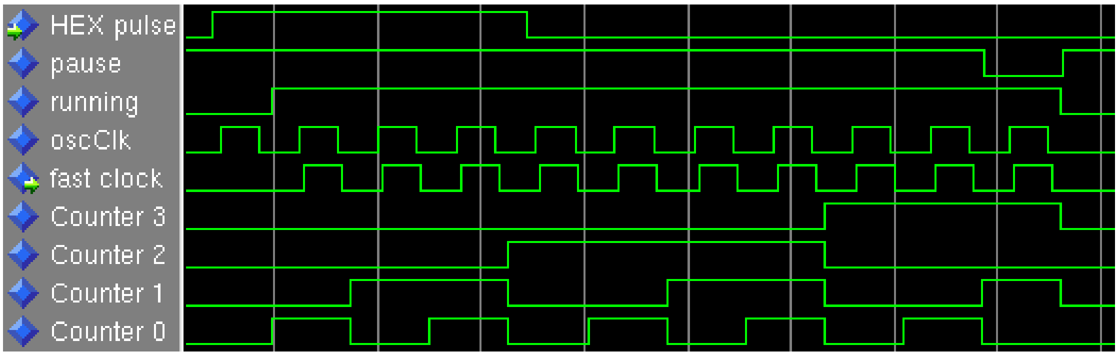


Figure 3.42: A screenshot of a simulation of the clock multiplier circuit generating the fast clock. On top is the HEX pulse signal, which disables the high-active pause signal of the ring oscillator. The cycle counter, whose activity is represented by the running signal, starts counting as soon as a clock by the oscillator is provided. The running signal holds the pause signal low and enables the output of the fast clock. Note that the cycle counter is running with the inverted clock of the oscillator, thus signal running goes high when oscClk is low. This is necessary to avoid shortened fast clock cycles. Below the fast clock, the bits of the cycle counter are shown.

Note that it is sufficient here if the frequency $f_{\min}$ is not a bound but holds amortized over $\Gamma_{\min}$ time. This, however, is not sufficient for the system presented in Chapter 6, where $f_{\min}$ must be a strict bound. To understand the performance implications resulting from the fact that the average pulse separation time is typically higher than the minimal pulse separation time, we will set those into perspective. Therefore, we assume for simplicity that this bound is an integer, i.e., we neglect that we may lose a "fractional tick", and choose $B = f_{\min}\Gamma_{\min}$. First, we check the long-term average frequency of the system. Clearly, we cannot guarantee a larger amortized frequency than $f_{\min}$. Of course, in addition $\Gamma_{\min}$ could be smaller than the long-term average time between pulses

$$\Gamma_{\text{avg}} := \lim_{k\to\infty} \frac{t^k_{\ell,0} - t^0_{\ell,0}}{k}. \tag{3.36}$$

Note that choosing the reference node to be $(\ell, 0)$ in this definition is arbitrary, as using every other node must lead to the same result: Since the skew in the fault-free case is bounded, the influence of the difference of the triggering times vanishes in the limit (assuming, of course, that the clock generation algorithm employed on layer 0 has bounded skew in the absence of faults). With this definition, the amortized frequency of the generated high-frequency clock of node (ℓ, i) can be expressed as

$$\lim_{k\to\infty} \frac{Bk}{t^k_{\ell,i} - t^0_{\ell,i}} = B \lim_{k\to\infty} \frac{k}{t^k_{\ell,0} - t^0_{\ell,0}} = f_{\min}\frac{\Gamma_{\min}}{\Gamma_{\text{avg}}}, \tag{3.37}$$

which equals $f_{\min}$ exactly if $t^{k+2}_{\ell,i} - t^{k+1}_{\ell,i} = t^{k+1}_{\ell,i} - t^k_{\ell,i} = \Gamma_{\min}$ for all i, ℓ, and k, i.e., the HEX clock is perfectly stable. The "frequency loss" is determined by three factors:

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

(i) The frequency stability of the clock generation algorithm at layer 0.

(ii) The skew between the layer 0 nodes.

(iii) The variance in the speed at which pulses propagate through the grid, i.e., ε.

The first two factors are determined by the clock generation algorithm and thus not attributable to HEX. We remark that since we need to keep pulses well-separated, (i) is likely to dominate (ii). Further, the frequency of the fast clocks will be $f_{\min}$ multiplied by $\Gamma_{\min}/\Gamma_{\mathrm{avg}}$. Since the difference between $\Gamma_{\min}$ and Γ_{avg}, in the limit, is caused by the skews, it is independent from the pulse separation time. Hence, the influence of the skews vanishes in the limit, whereby the resulting frequency of HEX reflects the frequency provided by the clock generation algorithm at layer 0. Therefore, a large pulse separation time will mitigate the influence of (iii).

These are good news, showing that a large pulse separation time does not hurt in terms of the overall frequency at which the system will be clocked. On the contrary, a large pulse separation time ensure the maximum frequency we can achieve by reducing the impact of the skews. Further, even with fairly small pulse separation times, the system will run at a constant fraction of the frequency $f_{\min}$.

In the following we will analyze the fast clock skew of adjacent nodes. Since pulses are anonymous, we interpret the generated local fast clock $F_{\ell,i}$ of node (ℓ, i) as a clock modulo B with the initial value being 0. For simplicity, we assume that the clock is continuous, i.e., it is real-valued from $[0, B)$, increasing modulo B at (possibly varying) rates from within $[f_{\min}, f_{\max}]$, where $f_{\max}$ is its maximal frequency. The actual discrete clock reading at time t then is simply $\lfloor F_{\ell,i}(t) \rfloor$. With these definitions, the clock value at time $t \in \left[t^k_{\ell,i}, t^{k+1}_{\ell,i}\right]$ satisfies

$$\min\left\{f_{\min}\left(t - t^k_{\ell,i}\right), B\right\} \leq F_{\ell,i}(t) \leq \min\left\{f_{\max}\left(t - t^k_{\ell,i}\right), B\right\}, \tag{3.38}$$

since the clock is halted until the next pulse once it reaches value $B \equiv 0$.

Observe that at times where two adjacent nodes both locally triggered pulse k, but not pulse $k+1$, the worst possible clock skew is attained if (i) the difference in the triggering times is maximal, (ii) the clock of the node that triggered first runs at frequency $f_{\max}$ until its clock halts, and (iii) the other node's clock runs at frequency $f_{\min}$. For the skew between two adjacent nodes in layer ℓ, we thus get at times $t \in \left[t^k_{\ell,i}, t^{k+1}_{\ell,i}\right] \cap \left[t^k_{\ell,i+1}, t^{k+1}_{\ell,i+1}\right]$ that

$$\begin{aligned} |F_{\ell,i}(t) - F_{\ell,i+1}(t)| &\leq B - f_{\min}\left(\frac{B}{f_{\max}} - \left|t^k_{\ell,i} - t^k_{\ell,i+1}\right|\right) \\ &\leq \frac{f_{\max} - f_{\min}}{f_{\max}} \cdot B + f_{\min}\sigma^{\max}_{\ell} = \varrho B + f_{\min}\sigma^{\max}_{\ell}, \end{aligned} \tag{3.39}$$

where $\varrho := (f_{\max} - f_{\min})/f_{\max}$ is the *relative drift* of the high-speed clocks.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

Now consider a time $t \in \left[t_{\ell,i}^{k+1}, t_{\ell,i}^{k+2}\right] \cap \left[t_{\ell,i+1}^{k}, t_{\ell,i+1}^{k+1}\right]$, i.e., node (ℓ, i) already triggered pulse $k+1$, but node $(\ell, i+1)$ has not done so yet. Since node i starts its clock—which in the worst case runs fast while the clock of node $i+1$ runs slow—at time $t_{\ell,i}^{k+1}$ again, a worst-case bound on the clock skew (modulo B) is obtained by comparing the clocks just before time $t_{\ell,i+1}^{k+1}$. However, this bound is subsumed by the previous one, as it covers the case $t = t_{\ell,i+1}^{k+1}$. Finally, the case $t \in \left[t_{\ell,i}^{k}, t_{\ell,i}^{k+1}\right] \cap \left[t_{\ell,i+1}^{k+1}, t_{\ell,i+1}^{k+2}\right]$ is symmetrical. This covers all cases and therefore provides a worst-case bound for the intra-layer skew of the constructed fast clocks; replacing $\sigma_\ell^{\max}$ by σ_ℓ^{avg} in this result yields an average-case bound.

To also derive good bounds for the inter-layer skew, more care is necessary. Of course, we could simply replace $\sigma_\ell^{\max}$ by $\hat{\sigma}_\ell^{\max}$ and repeat the same analysis, but this would provide overly conservative results as it fails to leverage the known bias of the inter-layer skew. We will hence use a refined analysis for slightly modified clocks to improve the results.

On average, nodes at layer ℓ trigger roughly $\hat{\sigma}_\ell^{\text{avg}}$ after those at layer $\ell - 1$ (recall that $\hat{\sigma}_\ell^{\text{avg}}$ respects the sign). Therefore, we can compensate the resulting bias in the inter-layer skew of the fast clocks by *shifting* the fast clocks of the nodes at layer $\ell > 1$ accordingly with respect to layer $\ell - 1$. Similarly, to obtain a worst-case bound, we can consider $\hat{\sigma}_\ell^{\min}$ and $\hat{\sigma}_\ell^{\max}$ and apply a shift corresponding to $\bar{\sigma}_\ell := (\hat{\sigma}_\ell^{\max} + \hat{\sigma}_\ell^{\min})/2$.

We examine the latter case here; the former can be treated similarly. In $\bar{\sigma}_\ell$ time, a running fast clock makes progress of at least $f_{\min}\bar{\sigma}_\ell$. Therefore, we map the clock values $F_{\ell,i}(t)$ for all i and t to $F'_{\ell,i}(t) := (F_{\ell,i}(t) + f_{\min}\bar{\sigma}_\ell) \bmod B$. Analogous to the intra-layer skew, we now can bound

$$\begin{aligned} F_{\ell-1,i}(t) - F'_{\ell,i}(t) &\le B - f_{\min}\left(\frac{B}{f_{\max}} - \left(t_{\ell,i}^{k} - t_{\ell-1,i}^{k}\right) + \bar{\sigma}_\ell\right) \\ &\le \frac{f_{\max} - f_{\min}}{f_{\max}} \cdot B + f_{\min}\left(\hat{\sigma}_\ell^{\max} - \bar{\sigma}_\ell\right) \\ &= \varrho B + f_{\min} \cdot \frac{\hat{\sigma}_\ell^{\max} - \hat{\sigma}_\ell^{\min}}{2} = \varrho B + f_{\min}\upsilon_\ell, \end{aligned} \tag{3.40}$$

where $\upsilon_\ell := \left(\hat{\sigma}_\ell^{\max} - \hat{\sigma}_\ell^{\min}\right)/2$ corresponds to $\sigma_\ell^{\max}$. The latter can be seen by noting that the signed variant of the (symmetrical, i.e., absolute value) inter-layer skew is just $-\sigma_\ell$. Likewise,

$$\begin{aligned} F'_{\ell,i}(t) - F_{\ell-1,i}(t) &\le B - f_{\min}\left(\frac{B}{f_{\max}} + \left(t_{\ell,i}^{k} - t_{\ell-1,i}^{k}\right) - \bar{\sigma}_\ell\right) \\ &\le \frac{f_{\max} - f_{\min}}{f_{\max}} B + f_{\min}\left(\bar{\sigma}_\ell - \hat{\sigma}_\ell^{\min}\right) = \varrho B + f_{\min}\upsilon_\ell, \end{aligned} \tag{3.41}$$

which together with the bound on $F_{\ell-1,i}(t) - F'_{\ell,i}(t)$ implies that $\left|F'_{\ell,i}(t) - F_{\ell-1,i}(t)\right| \le \varrho B + f_{\min}\upsilon_\ell$. An analogous reasoning shows that also $\left|F'_{\ell,i}(t) - F_{\ell-1,i+1}(t)\right| \le \varrho B + f_{\min}\upsilon_\ell$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Wien Bibliothek Your knowledge hub

Obviously, shifting all clocks in a layer by the same value will not affect the intra-layer clock skews. Applying the clock shifts inductively, i.e., shifting the clocks in each layer $\ell \geq 1$ by $\sum_{\ell'=0}^{\ell-1} \bar{\sigma}_{\ell'}$, we can bound the inter-layer skews on each pair of consecutive layers as above. We thus can conclude with the worst-case bounds

$$\varrho B + f_{\min}\sigma_\ell^{\max} \leq f_{\min}(\varrho\Gamma_{\min} + \sigma_\ell^{\max}) \text{ and} \tag{3.42}$$

$$\varrho B + f_{\min}\upsilon_\ell \leq f_{\min}(\varrho\Gamma_{\min} + \upsilon_\ell) \tag{3.43}$$

on the intra- and inter-layer skews of the (shifted) fast clocks. Note that they are tight for $B = f_{\min}\Gamma_{\min}$, which maximizes the amortized frequency of the fast clocks.

It can be seen that even if ϱ is fairly large, e.g., 10%, and $\Gamma_{\min}$ is significantly larger than $\sigma_\ell^{\max}$, e.g., up to factor 10, the maximal clock skew is still dominated by the term $f_{\min}\sigma_\ell^{\max}$. If one is willing to invest in more stable fast clock sources, e.g., $\varrho \approx 1\%$, large pulse separation times are feasible without incurring a significant impact on the skew. We stress that $f_{\min}$ and $f_{\max}$ in our bounds are—unless the clocks are extremely unstable—the amortized frequency upper and lower bounds over B fast clock pulses; large skews require a consistent difference in frequency for roughly $B/f_{\min}$ time. Moreover, ϱ captures the *relative drift* of the clocks. Any frequency change that applies to adjacent nodes in roughly the same way, i.e., a system-wide change in temperature or supply voltage, will not have noticeable effects on the skews.

Regarding the sizing of ϱ, we refer to the work of Sundaresan, Allen, and Ayazi [SAA06], where the effects of PVT on a ring oscillator where studied. While the compensated oscillator performed better ($\varrho < 2\%$ in simulations and $\varrho < 3\%$ in measurement), the uncompensated was above the 10% used above: the measurements lead to a $\varrho < 12\%$ while the simulations showed a much higher uncertainty of $\varrho < 32\%$.

3.5.3 Determination of $\Gamma_{\min}$

So far, we neglected two key obstacles in our approach for constructing fast clocks:

- $\Gamma_{\min}$ is not known, and therefore we do not know an appropriate choice for B.
- Due to persistent or transient faults, $\Gamma_{\min}$ may be very small or not even well-defined.

Concerning the first issue, there are basically three options: We can rely on analytical bounds, simulation results, or experimental data. All approaches have pros and cons; for a final system, experimental data is the most valuable, but it is also the most difficult and expensive to create, as one needs a chip and a suitable apparatus for measurements first. From Section 3.2 and bounds for the layer 0 clock generation algorithm employed, analytical worst-case bounds can be derived. The results from Section 3.4 and Chapter 4 provide insight into the skew distributions for an average-case setting under some fairly conservative assumptions on the parameters.

Regarding the second issue, we clearly must exclude faulty nodes as well as triggering time differences from the self-stabilization period when estimating (an

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

equivalent to) $\Gamma_{\min}$ in the general setting. Note that there is a trade-off: We can consider a non-faulty node that experiences large skews as incorrect, permitting the use of a less conservative bound on $\Gamma_{\min}$; in turn, we are losing the guarantee that this node will be able to complete each clock burst before the next pulse arrives, even after stabilization of HEX. The results from Section 3.4 give an idea on the behavior of the system under faults. It should be noted, though, that the specific values are highly dependent on the implementation of the HEX nodes, the layer 0 clock generation algorithm, and the used technology.

Once a suitable estimate $\tilde{\Gamma}_{\min}$, taking the role of $\Gamma_{\min}$, has been determined, one can choose B according to $B \leq f_{\min}\tilde{\Gamma}_{\min}$. Note that using a smaller value for B may be desirable in some situations, as this also leads to a smaller skew. On the downside, $B < f_{\min}\tilde{\Gamma}_{\min}$ is equivalent to using a smaller value of $\tilde{\Gamma}_{\min}$, which leads to a lower amortized frequency of the fast clocks. The resulting decrease in the amortized frequency can of course be mitigated by increasing the frequency of the layer 0 pulse generation algorithm, which reduces $\tilde{\Gamma}_{\min}$, Γ_{avg} and the gap $f_{\min}\tilde{\Gamma}_{\min} - B$.

3.6 Conclusions

We proposed a candidate for a scalable and fault-tolerant alternative for clock distribution in VLSI circuits, multi-core processors, and other hardware devices. Our approach supports multiple synchronized clock sources and is self-stabilizing even in the presence of a large number of non-clustered Byzantine failures of both clock sources and HEX nodes. Theoretical worst-case analysis and elaborate simulation experiments, using a hardware-level implementation, have been utilized to show that HEX guarantees a small skew between neighbors in the grid. We have also proposed a way to locally generate a high-frequency signal based on HEX, and analyzed the resulting skew towards the neighbors in the grid.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

CHAPTER 4

ALTERNATIVE TOPOLOGIES FOR CLOCK DISTRIBUTION GRIDS

> You can't always get what you want
> But if you try sometime you find
> You get what you need
>
> — MICK JAGGER AND KEITH RICHARDS,
> You Can't Always Get What You Want

AS WE HAVE SEEN in Chapter 3, HEX solves the clock distribution problem in a rather simple way. However, in the presence of faults, the skews increases significantly compared to the fault-free case, even under well-behaved input skews. Hence, in this chapter we focus on evaluating alternative topologies and firing algorithms for clock distribution grids. The goal is to find a suitable replacement for HEX which may sacrifice some parts of HEX's simplicity in favor of reducing the skews in certain cases. The improvements we seek are, for one, an improved behavior under faults, especially under well-behaved input skews. Furthermore, an answer to the question whether a different topology, with reduced hardware requirements, can deliver a performance similar to HEX in certain settings.

As the analytical consideration of a specific combination, consisting of topology and algorithm, is quite complex, even for the rather simple HEX approach, we focus on simulations. These simulations allow us to gather average-case results that provide a first but nevertheless representative evaluation of the fitness of each combination.

This chapter is structured as follows: First, we will introduce the considered algorithms and topologies in Section 4.1. Then, Section 4.2 presents the used simulation

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

and evaluation framework. Finally, the results of the simulations are presented and reviewed in Section 4.3, with the conclusion being drawn in Section 4.4.

4.1 Alternative Algorithms and Topologies

We will investigate the effects on the skews in a grid caused by different edge sets of the communication graphs, and different algorithms that determine when a node propagates a pulse. As will be shown, the selection of the grid topology has a large impact on the properties of the CDN. From a high-level viewpoint, one expects that a higher connectivity allows to propagate pulses faster, hence reducing the skew due to the possibility of an earlier triggering time. Alternatively, one could opt to tolerate more faults at the cost of a slower pulse propagation. However, as the results show, this hypothesis on the increased connectivity does not hold up with reality in every case. Moreover, when considering VLSI implementations, reality strikes: Differences in wire-lengths require careful tuning of the resistance to equalize delays, and crossing wires can promote crosstalk. Hence, a complex topology needs to be noticeable ahead of a simpler one to cover the costs of the implementation and the additional chip area required.

4.1.1 Topologies

We consider a set of nodes that executes a pulse generation and forwarding algorithm by communicating via message passing over a communication network whose underlying undirected communication graph forms a cylindrical grid. Formally, we define a grid via a directed communication graph (V, E): Letting $L \in \mathbb{N}$ denote its length and $W \in \mathbb{N}$ its width, the set of nodes V is the set of tuples $\{(\ell, i) \in [L+1] \times [W]\}$. Here, $[L+1] := \{0, \dots, L\}$ denotes the row index set, referred to as *layers*, and $[W] := \{0, \dots, W-1\}$ the column index set of the nodes in the grid.

For each node $(\ell, i) \in V$, $\ell \in [L+1] \setminus \{0\}$, $i \in [W]$, the edges in E define the directed links to other nodes in the grid. We call incoming and outgoing links to such neighboring nodes of the same layer, e.g., from (ℓ, i) to $(\ell, i-1 \mod W)$, intra-layer links. Whereas, links to/from nodes in different layers are called inter-layer links. In general, inter-layer links will be directed from a lower layer to a higher layer.[1] Further, we require that each fault-free link guarantees a communication delay, i.e., the time between sending and receiving a trigger message, within $[d^-, d^+] \subset (0, \infty)$, where $\varepsilon := d^+ - d^-$.

Nodes at layer 0 are special as they act as primary clock sources, i.e., they execute a pulse generation algorithms like [DFLS14; FS12] that generates synchronized and well-separated consecutive initial trigger messages. We denote with $\sigma_\ell := \max_{i \in [W]} \{|t_{\ell,i} - t_{\ell,j}| \mid ((\ell, j), (\ell, i)) \in E\}$ the maximum skew of layer ℓ. For layer 0, σ_0 characterizes the synchrony of the pulse generation algorithm, which is assumed

---

[1] Otherwise, the logic of the nodes would have to ensure that no stray/duplicate pulses are distributed in the grid.

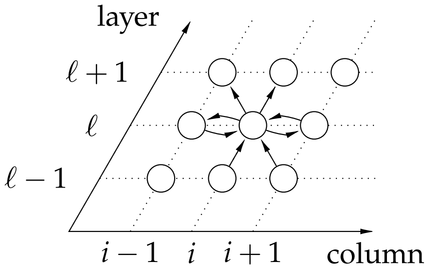


Figure 4.1: Node (ℓ, i) and its links in the cylindrical hexagonal grid topology of $\mathcal{G}_H$. Column coordinates are modulo W and layer coordinates (rows) are between 0 and L.

to reside at layer 0. In this chapter, we focus on a single pulse and hence omit the discussion regarding the requirements on the pulse generation, as has already been done in Chapter 3.

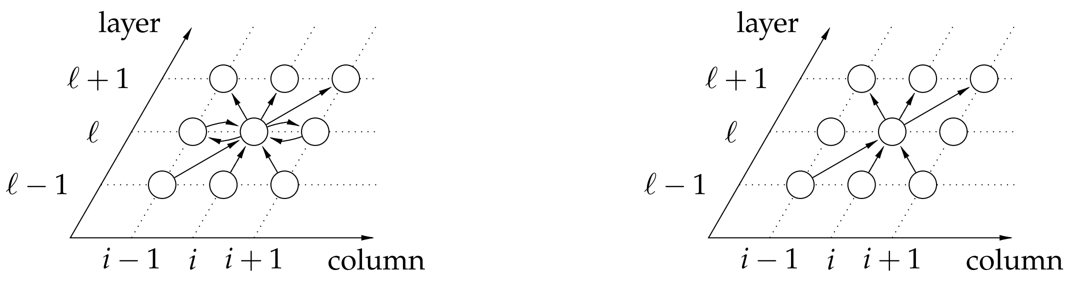
The grid used by HEX, as introduced in Chapter 3, will be referred to as $\mathcal{G}_H$ in this chapter. $\mathcal{G}_H$ has a simple and planar communication structure, as shown in Figure 4.1, where a node (ℓ, i) has links to its two adjacent intra-layer neighbor and to its upper $(\ell+1, i)$, and upper-left $(\ell+1, i+1 \mod W)$ inter-layer neighbor. Note that nodes in layer L do not have outgoing inter-layer links.

Under the assumption that connections to local neighbors are advantageous, an extended version of $\mathcal{G}_H$, named $\mathcal{G}_A$, has been contrived that adds an additional inter-layer neighbor, cf. Figure 4.2a. Furthermore, a version of $\mathcal{G}_A$ without intra-layer links, $\mathcal{G}_B$ (Figure 4.2b), was considered as well. The grids $\mathcal{G}_C$ (Figure 4.2c) and $\mathcal{G}_D$ (Figure 4.2d) are extensions of $\mathcal{G}_A$ and $\mathcal{G}_B$ with two additional inter-layer neighbors per node. Further, $\mathcal{G}_E$ (Figure 4.3a) was considered, as a variant of $\mathcal{G}_C$, where the intra-layer links reach to the $(\ell, i \pm 2 \mod W)$ neighbors, instead of the direct neighboring columns.

Finally, $\mathcal{G}_F$ (Figure 4.3b) differentiates itself from the other grids as it has a long inter-layer link spanning two layers. To model this longer link faithfully, we define that it has a communication delay different from the other links. This delay is denoted in the algorithm's description, e.g., $\mathcal{G}_F^{13,15}$ means that the longer delay is distributed in the interval $[13, 15]$ ns. The longer link, however, is not suitable for the nodes in layer 1. Hence, to connect the pulse generation in layer 0, the longer link is replaced by one to a neighbor in layer 0, i.e., node $(1, i)$ has $(0, i-1)$ as its third incoming neighbor.

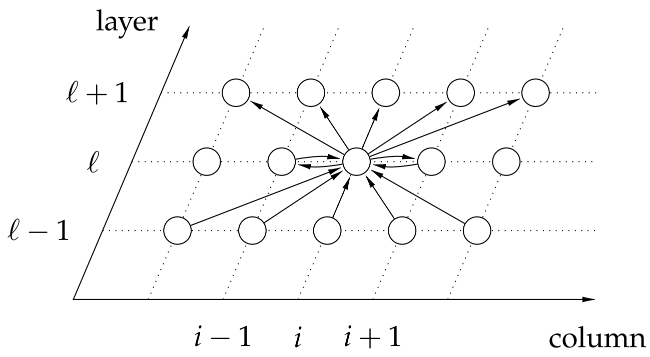
Due to the quite significant differences between the topologies, the question of achieving a fair comparison arises. Using the same neighborhood for the communication as is used for the clock distribution is questionable, as the resulting overhead of the large communication grid may not be justifiable in a general setting. Hence, communication links towards the direct physical neighbors seems favorable, not to mention from the standpoint of routing complexity. Obviously, the skew is only relevant between nodes that interact with each other. Thus, for every skew evaluation except $\mathcal{G}_F$, we only considered the neighbors toward which $\mathcal{G}_A$ has a link. For $\mathcal{G}_F$, the $\mathcal{G}_A$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

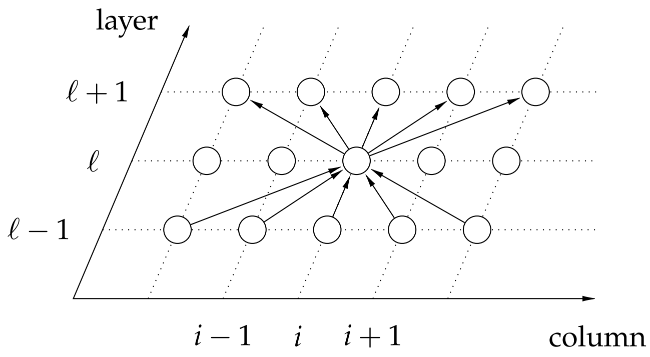


(a) Grid topology of $\mathcal{G}_A$.

(b) Grid topology of $\mathcal{G}_B$.



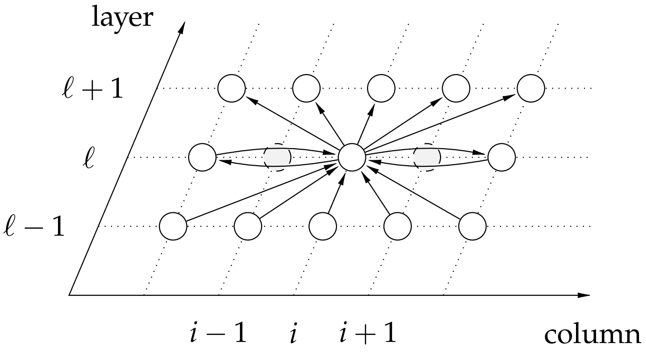
(c) Grid topology of $\mathcal{G}_C$.



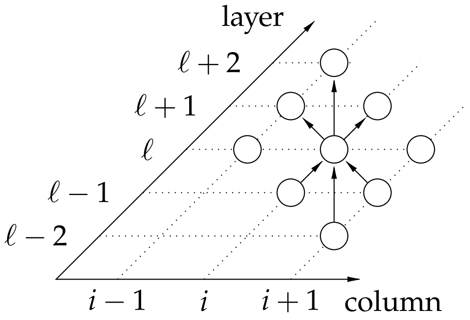
(d) Grid topology $\mathcal{G}_D$.

Figure 4.2: Overview of node (ℓ, i) and its links in the different grid topologies. Column coordinates are modulo W and layer coordinates (rows) are between 0 and L.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) Grid topology of $\mathcal{G}_E$.



(b) Grid topology of $\mathcal{G}_F$.

Figure 4.3: Overview of node (ℓ, i) and its links in the different grid topologies. Column coordinates are modulo W and layer coordinates (rows) are between 0 and L. The shown neighbors of $\mathcal{G}_F$ in (b) are those considered in the skew evaluation.

neighborhood would include links to nodes farther away, leading to an unfavorable comparison. Thus, $\mathcal{G}_F$ is evaluated with the neighborhood of $\mathcal{G}_H$, which covers the nodes in the direct neighborhood.

4.1.2 Algorithms

We call a *combination* a pairing of an algorithm with a grid topology. In every combination, we assume that each node on layer $\ell > 0$ in the grid runs the same algorithm that can broadcast *trigger messages* (representing clock pulses) over its outgoing links, as well as receive trigger messages over its incoming links. Nodes at layer 0 execute a pulse generation algorithm, which generates trigger messages. The input skew bound guaranteed by this pulse generation specifies the considered *set* of a combination. Depending on the given topology, specifically the number of incoming links, different algorithms are applicable. From an implementation viewpoint, each node has access to a (possibly inaccurate) clock to measure timeouts. In the simulations conducted here, however, we consider only a single pulse propagating through the grid. Nonetheless,

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

Algorithm 4.1: Two-adjacent pulse forwarding algorithm $\mathcal{A}^2_{\text{adj}}$ for nodes in layer $\ell > 0$.

```
upon receiving trigger message from neighbor do
    memorize message for $\tau \in [T^-_{\text{link}}, T^+_{\text{link}}]$ time;
upon having memorized trigger messages from two adjacent neighbors do
    broadcast trigger message; // produce pulse
    sleep for $\tau \in [T^-_{\text{sleep}}, T^+_{\text{sleep}}]$ time;
    forget previously received trigger messages;
```

Algorithm 4.2: Two-any pulse forwarding algorithm $\mathcal{A}^2_{\text{any}}$ for nodes in layer $\ell > 0$.

```
upon receiving trigger message from neighbor do
    memorize message for $\tau \in [T^-_{\text{link}}, T^+_{\text{link}}]$ time;
upon having memorized trigger messages from two neighbors do
    broadcast trigger message; // produce pulse
    sleep for $\tau \in [T^-_{\text{sleep}}, T^+_{\text{sleep}}]$ time;
    forget previously received trigger messages;
```

the algorithms presented are of course capable of propagating consecutive pulse waves.

Since clock pulses, i.e., trigger messages, are anonymous in this model, i.e., carry no information except their sole occurrence, care must be taken in order to prevent the generation of multiple trigger messages for a single pulse. The simple solution, as already used by HEX (cf. Section 3.2 on Page 50pp), is to rely on a sufficiently large separation between pulses, which relieves us from locally keeping track of pulse counts. To achieve this, a node will memorize each link's recent trigger message for some time between T^-_{link} and T^+_{link}, $T^+_{\text{link}} \geq T^-_{\text{link}}$, and then clear the associated memory flag, where the slack $T^+_{\text{link}} - T^-_{\text{link}}$ accounts for inaccurate local measurements of time. After producing a pulse, a node goes to sleep, i.e., will not locally trigger a further pulse, for some time between T^-_{sleep} and $T^+_{\text{sleep}} \geq T^-_{\text{sleep}}$, and clears all its memory flags upon waking up again.

The algorithm used by HEX, called the two-adjacent algorithm $\mathcal{A}^2_{\text{adj}}$ here (Algorithm 4.1), acts as the baseline we seek to improve. Basically, in $\mathcal{A}^2_{\text{adj}}$, a node forwards pulse k once it memorized trigger messages for pulse k from two adjacent neighbors, e.g., in case of $\mathcal{G}_H$ from $(\ell, i-1)$ and $(\ell-1, i-1)$. Hence, this type of algorithm can be seen as pulse-based. Due to its simplicity, $\mathcal{A}^2_{\text{adj}}$, as well as the other presented pulse-based algorithms, can easily be implemented by means of an asynchronous state machine and a few ring oscillators, as has been shown in Section 3.3 on Page 71pp.

Besides $\mathcal{A}^2_{\text{adj}}$, we also consider the obvious extension $\mathcal{A}^3_{\text{adj}}$ that waits for the reception of trigger messages from three of its neighbors. In the topologies described above, $\mathcal{A}^3_{\text{adj}}$ is only applicable to $\mathcal{G}_C$ and $\mathcal{G}_E$: In the other grids, the inter-layer neighbor $(\ell-1, i)$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Algorithm 4.3: Averaging algorithm $\mathcal{A}^{d,f}_{\mathrm{avg,o}}$ for nodes in layer $\ell > 0$. Where o is a given delay, n is the number of neighbors of a node in the grid, d is a parameter to reduce the effect of late neighbors, and f the number of neighboring failures the algorithm shall tolerate.

```
upon receiving trigger message from neighbor do
    timestamp received message;
    memorize message for $\tau \in [T^-_{\mathbf{link}}, T^+_{\mathbf{link}}]$ time;
upon having memorized at least $n - f - d$ trigger messages do
    let $m$ be the set of timestamps of the received trigger messages in ascending
      temporal order;
    let $w$ be $o + \mathrm{avg}\{m_f, m_{f+1}, \ldots, m_{n-f-d}\}$;
    sleep until $w$;
    broadcast trigger message; // produce pulse
    sleep for $\tau \in [T^-_{\mathbf{sleep}}, T^+_{\mathbf{sleep}}]$ time;
    forget previously received trigger messages;
```

Algorithm 4.4: Midpoint algorithm $\mathcal{A}^{d,f}_{\mathrm{mid,o}}$ for nodes in layer $\ell > 0$. Where o is a given delay, n is the number of neighbors of a node in the grid, d is a parameter to reduce the effect of late neighbors, and f the number of neighboring failures the algorithm shall tolerate.

```
upon receiving trigger message from neighbor do
    timestamp received message;
    memorize message for $\tau \in [T^-_{\mathbf{link}}, T^+_{\mathbf{link}}]$ time;
upon having memorized at least $n - f - d$ trigger messages do
    let $m$ be the set of timestamps of the received trigger messages in ascending
      temporal order;
    let $w$ be $o + \frac{m_f + m_{n-f-d}}{2}$;
    sleep until $w$;
    broadcast trigger message; // produce pulse
    sleep for $\tau \in [T^-_{\mathbf{sleep}}, T^+_{\mathbf{sleep}}]$ time;
    forget previously received trigger messages;
```

would be in every firing-relevant subset of the neighborhood, and hence a single point of failure. Due to its dependence on three neighbors, $\mathcal{A}^3_{\mathrm{adj}}$ can sustain two faulty neighbors, i.e., a 2–local fault model can be utilized. In this case, we add $2f$ as a superscript to the algorithm, i.e., $\mathcal{A}^{3,2f}_{\mathrm{adj}}$.

Additionally, the algorithms $\mathcal{A}^2_{\mathrm{any}}$ (Algorithm 4.2) respectively $\mathcal{A}^3_{\mathrm{any}}$ were also considered. They differ from $\mathcal{A}^2_{\mathrm{adj}}$ respectively $\mathcal{A}^3_{\mathrm{adj}}$ in their consideration of the received triggering messages: Only the number of memorized trigger messages is of relevance and not by whom they were sent.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

We further consider time-based algorithms that require complexer nodes than the pulse-based algorithms presented above. By including facilities for sorting, arithmetic operations, and timestamping incoming trigger messages, we can consider the following algorithms: The averaging algorithm $\mathcal{A}_{\text{avg,o}}^{d,f}$ (cf. Algorithm 4.3) reduces the set of received messages by dropping the first and last f received trigger messages. The reception time of the remaining $n-2f$ messages is then averaged. This average is added to a constant offset o to determine the point in time when the node will fire. This constant offset is required as the average is in general well before the reception of the $(n-f)^{\text{th}}$ trigger message, leading to a calculated triggering time in the past, and is annotated in the subscript if needed. Note that we assume here that this constant offset is not influenced by any variance, i.e., this timer of the node has access to a perfect clock.

A closer consideration of this algorithm in combination with $\mathcal{G}_C$ and $f=1$ reveals that the system will deadlock: For a node to fire, one of the two intra-layer neighbors is needed to receive the required $n-f^{\text{th}}$ trigger messages. However, this cannot happen as every node in layer ℓ would require a triggered intra-layer neighbor, hence leaving every node waiting. To circumvent this issue, the parameter d has been introduced. This parameter allows to configure the last trigger message for which the algorithm waits before starting the averaging, i.e., the last trigger message is $n-f-d$ instead of $n-f$. We annotate d in the superscript of the algorithm, e.g., $\mathcal{A}_{\text{avg}}^{1,f}$ denotes $d=1$. Note that we omit the parameter f if it is 1, i.e., when we consider the 1–local fault model.

Similarly to $\mathcal{A}_{\text{avg}}$, the midpoint algorithm $\mathcal{A}_{\text{mid}}$ (cf. Algorithm 4.4) also works on the set of messages between the f^{th} and $n-f-d^{\text{th}}$ received message. However, instead of calculating the average over the set, only the f^{th} and $n-f-d^{\text{th}}$ message are averaged, hence the term midpoint.

Sadly, both algorithms have a weakness in some combinations, e.g., $\mathcal{G}_C$ with $d=1$, which can lead to a calculated triggering times in the past. Consider the setting shown in Figure 4.4 for $\mathcal{A}_{\text{mid}}$, which is derived from a simulation experiment. There, the nodes between $(\ell, i-4)$ and $(\ell, i+5)$ require an intra-layer neighbor to trigger, as they all have a faulty inter-layer neighbor. Hence, given fitting trigger times for the node in layer $\ell-1$, a ramp-like triggering will happen for these nodes in layer ℓ: The ramp will start at $(\ell, i-5)$ and $(\ell, i+6)$ from where it will move towards (ℓ, i) and $(\ell, i+1)$. As this is a time consuming process, due to the additive offset o, the averaging will deliver triggering times that are before the arrival time of the $n-f-d^{\text{th}}$ message at some nodes. Moreover, we have found that even with an offset of $o \gg 50d^+$ the simulations resulted in a triggering time in the past, even though in other combinations $o=3.125d^+$ is sufficient. Besides, as soon as an intra-layer node is required for the triggering of a node, the skew will increase tremendously due to o affecting the triggering time.

The effect of a calculated triggering time in the past is the immediate triggering of the node, as we consider a causal system. It is not clear whether this has an impact on the performance of the system. Furthermore, it is unclear whether the origin of this effect lies in the fault placement or in a fundamental deficit of this combination. Note that it is quite possible that a strengthening of the failure model would allow to

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Wien Bibliothek Your knowledge hub

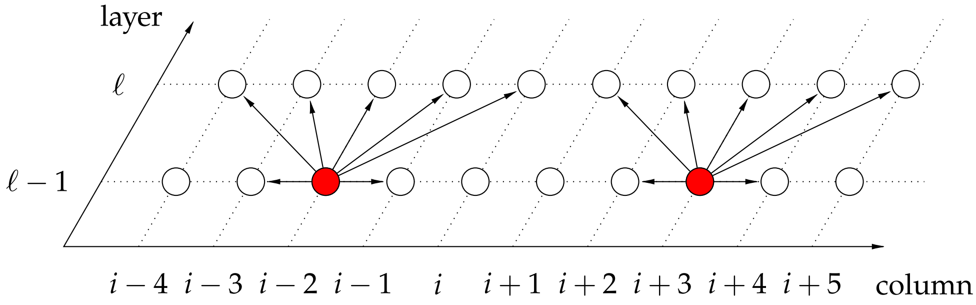


Figure 4.4: A setting with two faults (red nodes) that allows for multiple nodes to trigger after their calculated triggering time, due to causality. This setting requires that the nodes from $(\ell-1, i-1)$ to $(\ell-1, i+2)$ trigger early, and that the nodes to the left respectively right of the faults trigger later. This causes the nodes in layer ℓ to trigger one after the other while gradually moving towards column i respectively $i+1$. As the nodes from $(\ell, i-2)$ to $(\ell, i+3)$ have received early messages from multiple nodes, the calculated triggering time will gradually decrease. Its margin, initially given by the offset o, to the time of calculation thus decreases and can even become negative!

exclude such problematic scenario. However, further restricting the 1–local fault model in a setting with an increased node complexity lowers the assumption coverage.

Even though the performance of $\mathcal{A}_{\text{mid}}$ is quite good in some settings, an in-depth analysis should not be dismissed when considering this combination for actual use in a chip. It is quite possible that the analysis shows that a smaller value of o does not negatively impact the performance of the system, hence making this combination a viable option.

4.2 Methodology

A worst-case bound on the skew is essential to ensure correctness of a chip that relies on a distributed clock signal. However, due to the complexity of the worst-case analysis, as already demonstrated in the case of HEX, we limit our attention here to simulation experiments. Indeed, the determination of average-case results via simulations is an efficient method for an initial evaluation of the fitness of a combination. Hence, we conducted extensive simulation experiments to acquire representative statistical data for gaining insight on the two key points:

1. What is the average-case skew of a combination?
2. How do the different combinations behave in the presence of faults?

The simulation framework of HEX has been partially reused, as some of the simulation targets are the same. Contrary to HEX, however, the implementation of the simulation

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

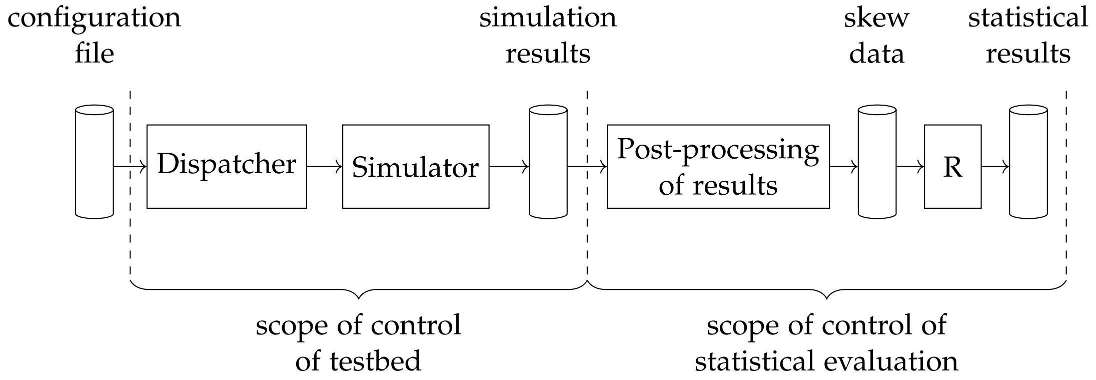

configuration file
simulation results
skew data
statistical results
Dispatcher
Simulator
Post-processing of results
R
scope of control of testbed
scope of control of statistical evaluation

Figure 4.5: This figure visualizes, from left to right, the process from a configuration file to the sought after statistical data. Starting with the configuration file, the dispatcher starts, for every run, a new instance of the simulator. From the simulation results the skews are calculated, which are then processed with R to gather the final evaluations.

has not been done in VHDL with ModelSim. For one, the newly developed custom C++ [Cpp11] simulation tool performs our simulations significantly faster. Furthermore, it allows a clear inspection of the actual behavior of a node with minimal overhead. Additionally, alternative grids and further evaluations can be implemented with less effort compared to the VHDL approach.

Figure 4.5 shows an overview of the dataflow between the major components involved in the simulation of a specific setting. A *set* describes a combination of parameters of a grid topology that shall be simulated. As some parameters are randomly distributed, e.g., the link delays or the placement of the faults, multiple executions of the same set were used to acquire meaningful data. We call every such execution a *run*. For each run, the testbed part of the process, shown in Figure 4.5, is executed, i.e., the simulator is started with a chosen seed for the *pseudo random number generator* (PRNG), to acquire reproducible results.

We will now explain each component of Figure 4.5 in detail.

4.2.1 Configuration File

The configuration file describes a specific set by defining the parameters of a combination. The used combination, however, is determined during compile-time of the simulator to keep backwards-compatibility with the framework used for the VHDL simulations. The configuration file allows to specify the following properties:

- Size of the grid.
- Interval in which the link delays shall be distributed.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

- Interval in which the longer link delays shall be distributed; this is only relevant for $\mathcal{G}_F$.

- Properties of the layer 0 clock source, e.g., the input skew σ_0.

- Number and types of the faults which should be placed in the grid without violating the 1–local fault assumption (cf. Section 3.2.2 on Page 61). The types of faults include

    - fail-silent faulty nodes, and
    - Byzantine faulty nodes, which behave, on a per-link basis, either fail-silent or send constantly a (fast) trigger message.

  The behavior of each of the Byzantine faulty nodes is chosen randomly on a per-link basis, in every run.

- The constant offset used by $\mathcal{A}_{\text{avg}}$ and $\mathcal{A}_{\text{mid}}$.

4.2.2 Dispatcher

The dispatcher is used to distribute the execution of sets over multiple computers. For every run, the dispatcher starts the simulator and provides a deterministic seed for the PRNG to the simulator. Furthermore, the dispatcher also allows to distribute the post-processing of the results.

4.2.3 Simulator

Initially, the simulator used here was designed to follow a reaction/diffusion-style approach for abstracting the propagation of the pulse in the entire grid. This simulation principle splits every time step into two distinct phases. First, the nodes of the grid interact (diffuse) with their environment, determining their next state. In the second phase, the reaction phase, the nodes perform the state-change determined in the diffusion phase. Due to this split into distinct phases, we could compute all nodes in parallel during each phase, as there are no side-effects between the nodes in each phase. This design was been chosen as it is, in general, applicable for a *graphical processing unit* (GPU) as the primary computing unit. This road promised a reduction in simulation times, compared to the VHDL simulations, due to the high parallelism enabled by the independence of the nodes in each phase. However, it turned out that after the reaction phase a memory transfer from the GPU to the *central processing unit* (CPU) was required to perform the termination check and record the triggering times of the nodes. Although the actual simulation benefited from the massive parallelism of the GPU, the overheads induced by the required synchronization between the phases and the memory transfer caused a massive penalty compared to a multi-threaded CPU implementation.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

This time-based simulation paradigm is, however, inefficient for the discrete clock distribution systems at hand. Hence, the simulator has been implemented in an event-list approach where the reception of messages and the triggering of the nodes are queued and processed. This decouples the simulation time from the delays used in the simulation and thereby reduced the simulation time by multiple magnitudes down into the millisecond range.

4.2.4 Skew Evaluation

In order to support a flexible evaluation, the custom simulation control and evaluation infrastructure developed for HEX has been modified to cover the changed requirements. That is, the evaluation tool is only used to calculate the neighbor skews based on the (new) grid topology here. For the calculation of the statistical parameters of the skews, scripts for the R programming language [RC018] have been used, which utilize R's extensive statistics and visualization features [Wic16].

The conducted experiments require only a single pulse in a simulation run ρ. Nonetheless, the primary quantities of interest in a simulation run ρ in a set R for a particular combination are the same as for HEX:

- the (absolute) intra-layer skews between neighboring nodes in the same layer ℓ, and
- the (signed) inter-layer skews of every node (ℓ, i), $\ell > 0$, relative to its neighbors in layers ℓ', $\ell' < \ell$.

Note that "neighboring nodes" means here a connection via a link in the grid used for the evaluation. We remark that the intra-layer skew is defined in terms of the absolute values due to the symmetry of the considered topologies (and thus skews) within a layer, whereas the latter respects the sign and thus correctly captures the non-zero bias, which depends on the combination. Due to the introduction of a run ρ of a set, we denote by $t_{\ell,i,\rho}$ the triggering time of node (ℓ, i) in run $\rho \in R$.

For $\text{op} \in \{q_1, q_5, \text{avg}, q_{50}, q_{95}, q_{99}, \max\}$, we define the 1%-quantile, 5%-quantile, average, median, 95%-quantile, 99%-quantile, and maximum (absolute)

- layer ℓ intra-layer skew in run ρ:
$$\sigma_{\ell,\rho}^{\text{op}} := \text{op}_{i,j\in[W]} \left\{ |t_{\ell,i,\rho} - t_{\ell,j,\rho}| \mid ((\ell,j),(\ell,i)) \in E \right\}; \tag{4.1}$$
- intra-layer skew in run ρ:
$$\sigma_{\rho}^{\text{op}} := \text{op}_{i,j\in[W],\ell\in[L+1]\setminus\{0\}} \left\{ |t_{\ell,i,\rho} - t_{\ell,j,\rho}| \mid ((\ell,j),(\ell,i)) \in E \right\}; \tag{4.2}$$
- intra-layer skew in set R:
$$\sigma^{\text{op}} := \text{op}_{i,j\in[W],\ell\in[L+1]\setminus\{0\},\rho\in R} \left\{ |t_{\ell,i,\rho} - t_{\ell,j,\rho}| \mid ((\ell,j),(\ell,i)) \in E \right\}. \tag{4.3}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Likewise, for op $\in \{\min, q_1, q_5, \text{avg}, q_{50}, q_{95}, q_{99}, \max\}$ we define the minimum, 1%-quantile, 5%-quantile, average, median, 95%-quantile, 99%-quantile, and maximum (signed)

- layer ℓ inter-layer skew in run ρ:

$$\hat{\sigma}^{\text{op}}_{\ell,\rho} := \text{op}_{i,j\in[W],m\in[L]} \left\{ t_{\ell,i,\rho} - t_{m,j,\rho} \mid ((m,j),(\ell,i)) \in E \wedge m \neq \ell \right\}; \tag{4.4}$$

- inter-layer skew in run ρ:

$$\hat{\sigma}^{\text{op}}_{\rho} := \text{op}_{i,j\in[W],\ell\in[L+1]\setminus\{0\},m\in[L]} \left\{ t_{\ell,i,\rho} - t_{m,j,\rho} \mid ((m,j),(\ell,i)) \in E \wedge m \neq \ell \right\}; \tag{4.5}$$

- inter-layer skew in set R:

$$\hat{\sigma}^{\text{op}} := \text{op}_{i,j\in[W],\ell\in[L+1]\setminus\{0\},m\in[L],\rho\in R} \left\{ t_{\ell,i,\rho} - t_{m,j,\rho} \mid ((m,j),(\ell,i)) \in E \wedge m \neq \ell \right\}. \tag{4.6}$$

The experiments were performed with and without faulty nodes. Note that the triggering times of faulty nodes are of course not considered when computing the inter- and intra-layer skews.

4.2.5 Skew of the Clock Input

Besides the actual combination, the other main parameter of the simulation is the input skew σ_0. The inputs skews were selected such that they define sets which are in line with the optimization targets. For adverse inputs, we have chosen a *high input skew* of $\sigma_0 = d^+$, i.e., every pair of neighboring nodes will trigger d^+ apart. These initial skews are quite high, although not unreasonable for a pulse generation affected by faults. Further, it provides a glimpse on the ability of the system to cope with very high skews induced by faults. For the *well-behaved input skews*, we conducted simulations with $\sigma_0 = \{\varepsilon, 2\varepsilon\}$. Note that a typical clock synchronization is able to provide a $\mathcal{O}(\varepsilon)$ worst-case skew between all nodes; for gradient clock synchronization, we have a neighbor skew bound in that range [FL06]. Finally, we use a benign setting with a *low input skew* of $\sigma_0 = 0$, which allows to study the best-case behavior of the combinations.

4.2.6 Simulation Parameters

We consider a grid with $L = 50$ and $W = 25$. The end-to-end delays are within $[d^-, d^+] = [7.161, 8.197]$ ns, to keep some comparability with Chapter 3. However, the timing added for the HEX node's internal logic ($[0.161, 0.197]$ ns) falls now also under the uniform distribution of the link delays.[2] We found that 1500 runs are sufficient for a qualitative evaluation of the observed parameters: By comparing the results for 500 runs with those for 1500 runs, only minor deviations of the maxima and in lower decimal digits of the other values were observed.

---

[2] As these delays were caused by asymmetry in the hardware design, this part of the delay was actually discretely distributed.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

4.2.6.1 Statistical Visualization

Whereas a quantitative approach for determining the fitness of a combination would have been favorable, determining a suitable model a priori was impossible due to unclear/varying optimization goals. Hence, a visual comparison was chosen. As the visualization, combined with additional constraints on the interesting sets, allowed a sufficient qualitative and quantitative ranking of the results, the generation of a mathematical fitness model has been omitted altogether.

During the analysis of the skew data, we realized that the usual metrics (minimum, quantiles, maximum) provide only a limited view of the behavior of a combination. A box plot, as has been used in Chapter 3, is sufficient to compare different sets of the same combination. Yet, the visualization of the skew distribution in this way provides inconclusive results for comparing different combination. Thus, we decided to utilize violin plots, as shown, e.g., in Figure 4.6a, which are basically mirrored kernel density plots.[3] The violin plots were created with R and allow a visual comparison of the distributions of the skews between multiple data sets in an aesthetically satisfactory way. The violins have horizontal marks to indicate the 1%, 5%, 50%, 95% and 99% quantile. Due to the structure of the data, for the intra-layer skews, the 1% quantile is not shown, as it is at, or close to, the minimum. The inner area of each violin is the same within a plot, hence the violins appearance is independent of the number of samples. This is relevant, as skews from/to faulty nodes are not considered for the skew calculations and hence the number of skews in two different violins can be different.

4.2.6.2 Relationship to the Simulations of HEX

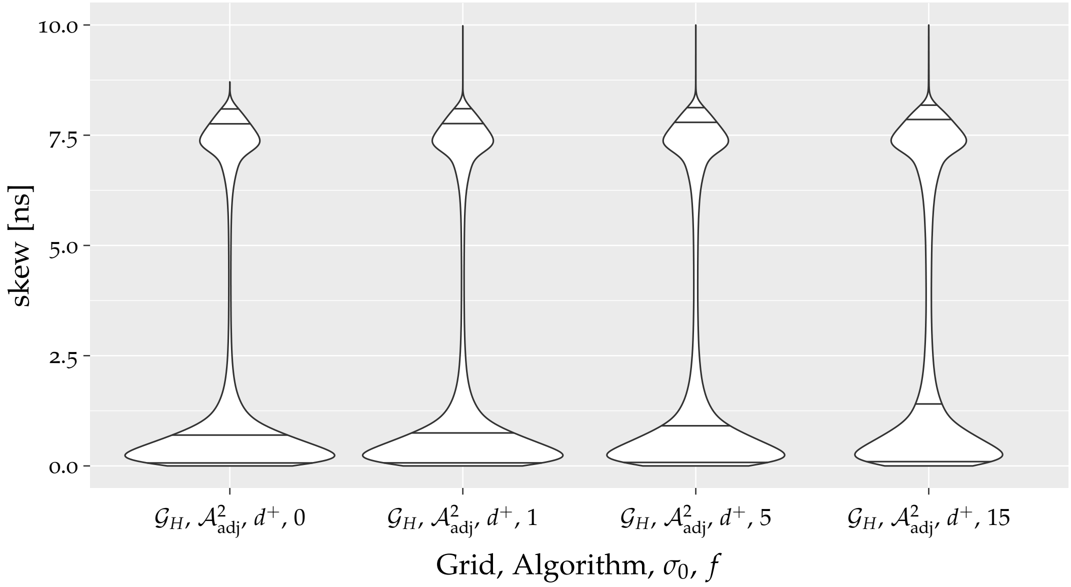
Due to the close relationship of the simulation environment to the one in Section 3.4, a comparison, for the sake of validation, seems appropriate. However, recall that in this chapter, the skew evaluation happens towards the neighbors of $\mathcal{G}_A$. Hence, a two-hop neighbor of the $\mathcal{G}_H$ topology is additionally considered in the skew calculation. Owing to this, the inter-layer skews exhibit a more stretched distribution, with higher maxima, compared to the data presented in Chapter 3. However, the evaluation with respect to $\mathcal{G}_A$ is of no relevance for the intra-layer skews. The different simulation methods, and the increased number of simulations runs, lead only to minor differences in the data acquired. For example, the results for set (iv) in Figure 3.32a and Tables 3.1 and 3.2 (Pages 90, 94 and 96) with $f = 0$ respectively $f = 1$ in Figure 4.6a and Table 4.1 are matching very well.

Figure 4.6a also reveals the strength of the violin plot: The box plot in Figure 3.32a shows a minor positive correlation between the number of faults f and the median, with otherwise relatively constant values, besides the maximum. This is also visible in the violin plot. However, additionally, two clusters of skews can be seen for which

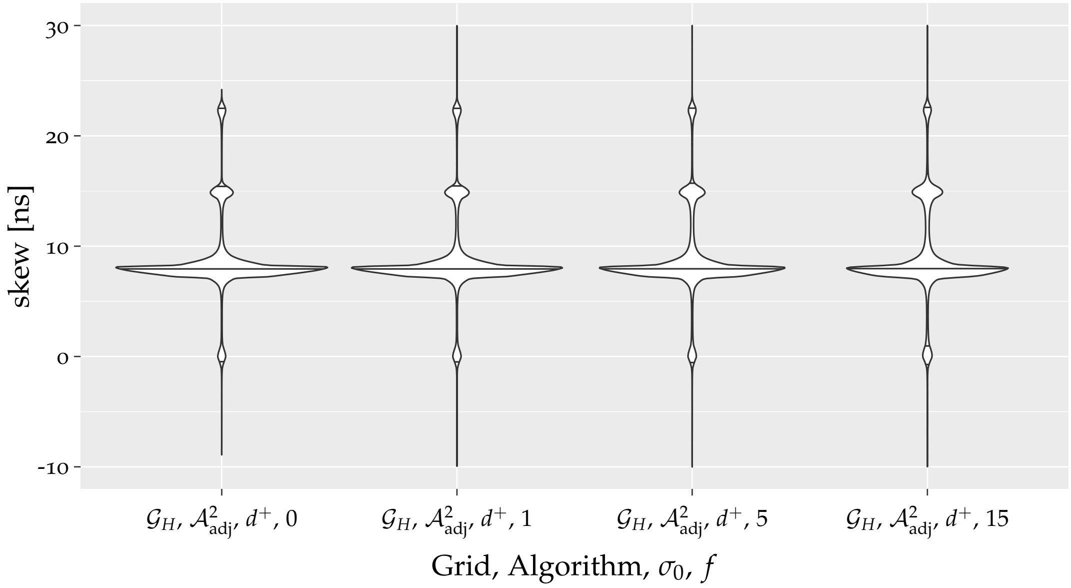
---

[3] A kernel density estimates the probability density function. That is, the plot represents a continuous and smooth version of the histogram of the provided data set. Note that the estimation may cause the minimum and maximum to differ from the provided data set.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



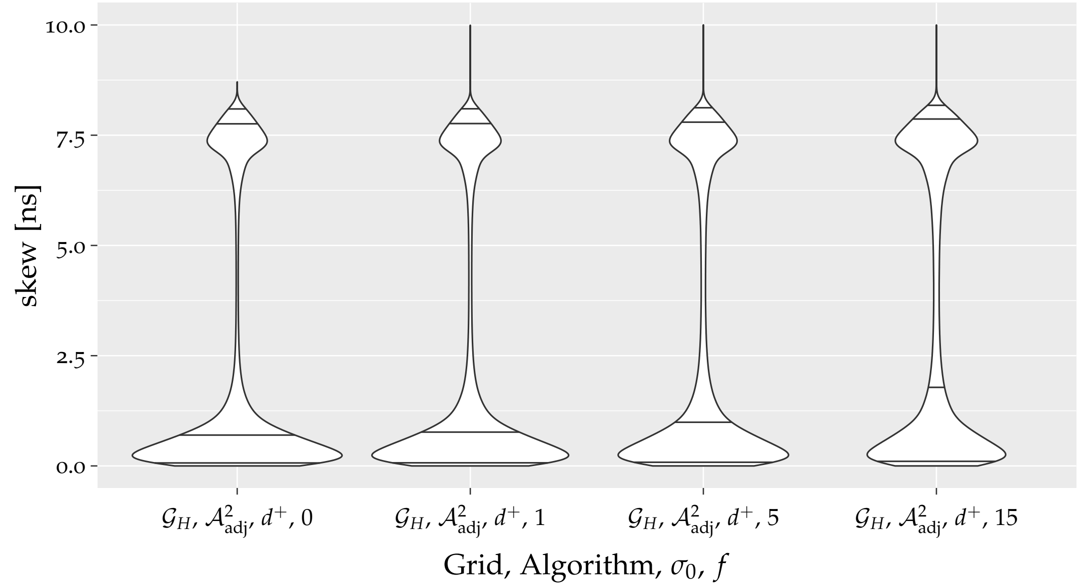
(a) intra-layer skews



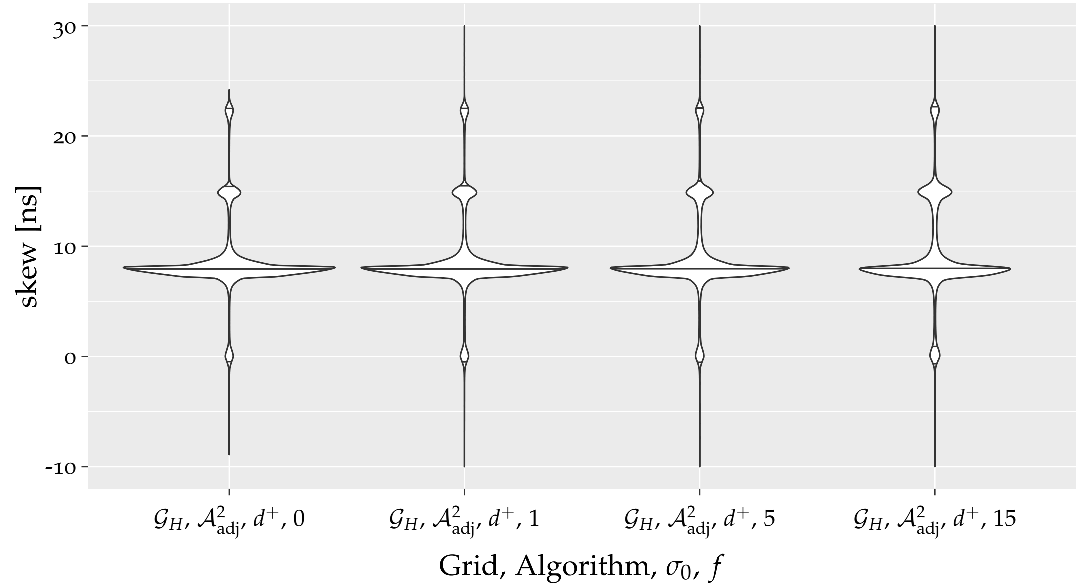
(b) inter-layer skews

Figure 4.6: The development of the skews in $\mathcal{G}_H$ under increasing number f of Byzantine faults, evaluated under the neighborhood of $\mathcal{G}_A$. The plots have been cropped to focus on the majority of the skews: (a) has been cropped at 10 ns, and (b) has been limited to be within $[-10, 30]$ ns. See Table 4.1 for the corresponding data.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skews



(b) inter-layer skews

Figure 4.7: The development of the skews in $\mathcal{G}_H$ under increasing number f of fail-silent faults, evaluated under the neighborhood of $\mathcal{G}_A$. The plots have been cropped to focus on the majority of the skews: (a) has been cropped at 10 ns, and (b) has been limited to be within $[-10, 30]$ ns. See Table 4.2 for the corresponding data.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.1: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\mathrm{op}}$ (in [ns]) in 1500 simulation runs on a 50×25 grid with f Byzantine nodes and an initial skew of $\sigma_0 = d^+$.

	intra-layer						inter-layer							
f	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
0	0.01	0.05	2.30	0.63	7.73	8.06	8.71	−8.90	−0.44	5.23	8.82	7.97	15.45	22.53	24.18
1	0.01	0.05	2.39	0.67	7.74	8.06	38.10	−30.54	−0.45	4.47	8.86	7.98	15.49	22.53	42.70
5	0.01	0.06	2.68	0.81	7.78	8.10	38.10	−30.54	−0.52	2.43	8.98	7.99	15.74	22.55	45.48
15	0.01	0.07	3.23	1.24	7.86	8.16	41.32	−33.37	−0.72	0.96	9.20	8.01	16.66	22.63	46.76

Table 4.2: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\mathrm{op}}$ (in [ns]) in 1500 simulation runs on a 50×25 grid with f fail-silent nodes and an initial skew of $\sigma_0 = d^+$.

	intra-layer							inter-layer							
f	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
0	0.01	0.05	2.30	0.63	7.73	8.06	8.71	−8.90	−0.44	5.23	8.82	7.97	15.45	22.53	24.18
1	0.01	0.05	2.42	0.69	7.74	8.06	30.40	−22.83	−0.45	4.33	8.88	7.98	15.52	22.54	44.55
5	0.01	0.06	2.80	0.88	7.78	8.10	32.82	−24.15	−0.50	2.23	9.07	8.00	15.98	22.58	48.50
15	0.02	0.08	3.46	1.55	7.88	8.18	32.82	−24.65	−0.65	0.90	9.42	8.03	18.72	22.72	48.50

the faults cause a transition of some skews from the first cluster, around the median, towards the second cluster. This confirms the intuition that there is a general increase of the skews in the presence of faults.

4.2.6.3 Determination of Specific Simulation Scenarios

As we have already seen in the simulation results of HEX, not every scenario imaginable is worth considering. Hence, we need to reduce the data sets presented to a representative subset. A beneficial reduction is arguably achievable by focusing on a specific fault setting.

First, we select the number of faults to consider by examining Figures 4.6 and 4.7 respectively Tables 4.1 and 4.2. We see that a majority of the skews are within the same interval, independently of the number of faults f. However, the tendency for higher skews grows with the number of faults. The most noticeable difference occurs for the outliers. Hence, we focus on the visualizations of the cases with $f = 0$ and $f = 15$, delivering a favorable spread between a benign setting and one with the tendency for large skews. As we consider grids of $50 \times 25 = 1250$ nodes with $f = 15$ faulty nodes, the observed setting corresponds to a failure rate of 1.2%. We believe that this is a sufficiently high failure rate to consider under “normal” operation. Furthermore, we limit the presentation to the settings with Byzantine faults due to the fact that, in a majority of the cases, this fault-type has larger outliers. Note that, in some cases, fail-silent faults cause even higher maximum values, yet the shape of the major part of the violin is the same.

4.3 Simulation Results

To provide an overview of the simulation results obtained, we focus on each grid and provide a comparison of the different applicable algorithms. Listing each of the above selected settings in detail, however, would impair our aim to provide a clear and concise presentation of the results. Hence, only selected plots are presented here. The complete violin plots of the selected settings, for each set, and some tables, have been deferred to Appendix A.

First, we consider each topology with their applicable pulse-based algorithms. In Section 4.3.8, we present the results with the time-based algorithms $\mathcal{A}_{\text{mid}}$ and $\mathcal{A}_{\text{avg}}$. Furthermore, we present simulations of settings with a 2–local fault assumption in Section 4.3.9.

4.3.1 Results of the Pulse-based Algorithms in $\mathcal{G}_H$

We compare the original HEX architecture ($\mathcal{G}_H$ with $\mathcal{A}^2_{\text{adj}}$) with $\mathcal{A}^2_{\text{any}}$. While the results in the fault-free case are basically identical, we can see differences in sets with faults as shown in Table 4.3. In these sets, the added firing rules allow certain nodes to fire earlier in $\mathcal{A}^2_{\text{any}}$, thus especially reducing the spread of the skews, as shown in Figure 4.8. This is due to the faster triggering of nodes with skewed inputs, facilitated by the

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

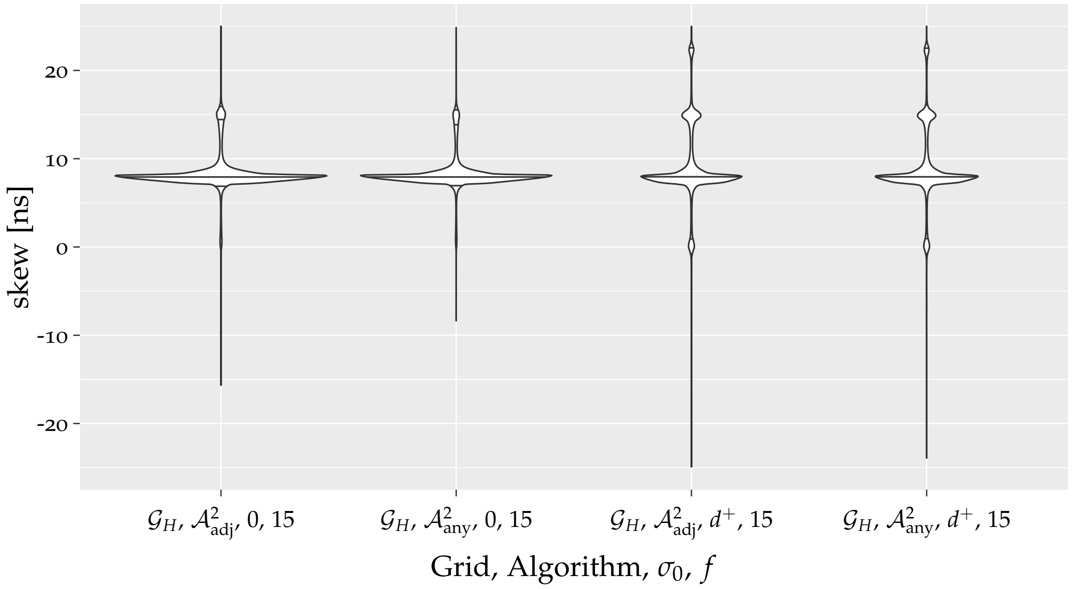


Figure 4.8: Comparison of the inter-layer skews of $\mathcal{G}_H$ under different algorithms and input skews, with $f = 15$ Byzantine faults. The plot has been limited to be within $[-25, 25]$ ns to focus on the majority of the skews.

removal of the adjacent neighbor constraint. While this leads to a visible reduction of the maximal skews, the IQD^1 and IQD^5 only show a minor reduction that is almost invisible in the violin plots given in Appendix A.1 on Page 237pp.

Note that the performance of the $\mathcal{A}^2_{\text{adj}}$ algorithm improves noticeable better, than $\mathcal{A}^2_{\text{any}}$ when excluding the directed neighbors of the faults ($h = 1$). However, $\mathcal{A}^2_{\text{any}}$ still performs better. The generally smaller IQD's of $\mathcal{A}^2_{\text{any}}$ are a result of the increased number of firing rules: They weaken the influence of the faults on the firing times of their neighbors. Hence, the exclusion of the neighborhood of the faults has a smaller effect as the synchronization is already rather good.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.3: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in $\mathcal{G}_H$.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^2_{\text{adj}}$	0	0	0	0.01	0.03	0.41	0.33	1.03	1.47	3.64	4.30	6.84	7.24	7.94	7.94	8.77	9.26	11.91
		15	0	0.01	0.04	1.38	0.49	7.12	7.90	28.31	−15.65	1.19	6.92	8.44	7.97	14.48	15.96	36.11
		15	1	0.01	0.04	1.25	0.48	6.69	7.58	21.05	−7.88	1.72	7.00	8.38	7.97	13.97	15.59	28.25
$\mathcal{A}^2_{\text{any}}$	0	0	0	0.01	0.03	0.41	0.33	1.03	1.47	3.64	4.30	6.84	7.24	7.94	7.94	8.77	9.26	11.91
		15	0	0.01	0.04	1.22	0.47	6.67	7.67	16.17	−8.35	1.73	6.98	8.36	7.96	13.89	15.58	29.50
		15	1	0.01	0.04	1.11	0.45	6.20	7.40	8.75	−4.15	2.39	7.06	8.31	7.96	13.24	15.29	24.02
$\mathcal{A}^2_{\text{adj}}$	ε	0	0	0.01	0.04	0.63	0.46	1.81	2.68	6.36	1.66	6.05	7.02	8.05	7.97	9.54	10.61	15.06
		15	0	0.01	0.05	1.51	0.64	7.14	7.91	29.15	−21.61	1.20	6.56	8.50	7.99	14.47	16.14	43.60
		15	1	0.01	0.05	1.38	0.62	6.70	7.60	18.84	−8.43	1.77	6.66	8.44	7.99	13.88	15.70	26.45
$\mathcal{A}^2_{\text{any}}$	ε	0	0	0.01	0.04	0.63	0.46	1.81	2.68	6.36	1.66	6.05	7.02	8.05	7.97	9.54	10.61	15.06
		15	0	0.01	0.05	1.33	0.60	6.65	7.69	16.70	−8.53	1.83	6.66	8.41	7.99	13.77	15.64	30.21
		15	1	0.01	0.05	1.23	0.58	6.14	7.42	8.56	−4.06	2.56	6.74	8.36	7.99	13.03	15.35	23.91
$\mathcal{A}^2_{\text{adj}}$	2ε	0	0	0.01	0.04	0.88	0.53	2.85	4.02	7.99	−0.25	5.03	6.79	8.18	7.98	10.69	12.46	17.81
		15	0	0.01	0.06	1.71	0.75	7.18	7.92	26.98	−18.42	1.09	6.03	8.60	8.00	14.56	16.31	41.37
		15	1	0.01	0.06	1.58	0.72	6.79	7.62	18.23	−10.95	1.59	6.17	8.54	8.00	14.06	15.81	24.51
$\mathcal{A}^2_{\text{any}}$	2ε	0	0	0.01	0.04	0.88	0.53	2.85	4.02	7.99	−0.25	5.03	6.79	8.18	7.98	10.69	12.46	17.81
		15	0	0.01	0.05	1.54	0.70	6.75	7.70	17.17	−9.27	1.64	6.17	8.51	8.00	13.97	15.72	29.45
		15	1	0.01	0.05	1.43	0.68	6.30	7.46	8.20	−5.09	2.27	6.31	8.47	8.00	13.36	15.44	23.82
$\mathcal{A}^2_{\text{adj}}$	d^+	0	0	0.01	0.05	2.30	0.63	7.73	8.06	8.71	−8.90	−0.44	5.23	8.82	7.97	15.45	22.53	24.18
		15	0	0.01	0.07	3.23	1.24	7.86	8.16	41.32	−33.37	−0.72	0.96	9.20	8.01	16.66	22.63	46.76
		15	1	0.01	0.07	3.12	1.17	7.81	8.11	33.59	−26.11	−0.62	1.15	9.17	8.01	16.36	22.57	38.08
$\mathcal{A}^2_{\text{any}}$	d^+	0	0	0.01	0.05	2.30	0.63	7.73	8.06	8.71	−8.90	−0.44	5.23	8.82	7.97	15.45	22.53	24.18
		15	0	0.01	0.07	3.09	1.12	7.83	8.12	31.72	−23.90	−0.70	0.99	9.10	8.00	16.14	22.58	36.91
		15	1	0.01	0.07	3.03	1.08	7.80	8.09	23.07	−22.68	−0.63	1.18	9.10	8.00	16.09	22.55	27.89

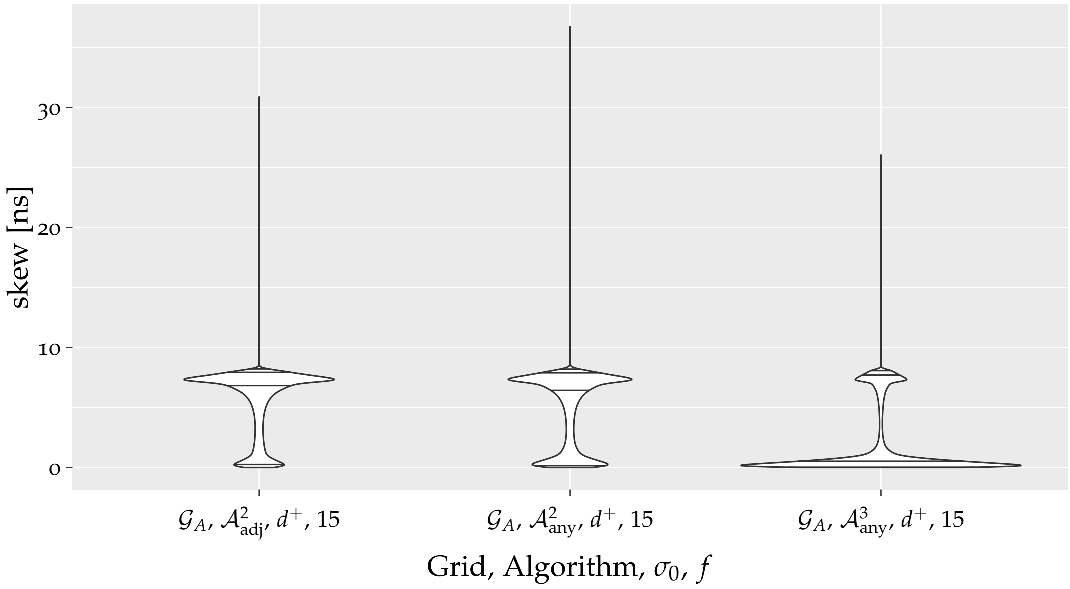


Figure 4.9: Comparison of the intra-layer skews under $\mathcal{G}_A$ with $\sigma_0 = d^+$ and 15 Byzantine faults. Notice that the median under $\mathcal{A}^3_{\text{any}}$ is in the lower cluster, while the other two algorithms have their median in their upper cluster.

4.3.2 Results of the Pulse-based Algorithms in $\mathcal{G}_A$

Comparing the skews in Table 4.3 with Tables 4.4 and 4.5 reveals that both algorithms, $\mathcal{A}^2_{\text{adj}}$ and $\mathcal{A}^2_{\text{any}}$, show a similar behavior in $\mathcal{G}_A$ as they did in $\mathcal{G}_H$, with the exception that $\mathcal{A}^2_{\text{any}}$ also performs better than $\mathcal{A}^2_{\text{adj}}$ in the fault-free case. A comparison of these two algorithms with $\mathcal{A}^3_{\text{any}}$, however, shows that the additional tolerance to faults of $\mathcal{A}^3_{\text{any}}$ leads to a stronger increase in the maximum inter-layer skew with rising input skew. Nonetheless, a closer look reveals that the intra-layer skews under $\mathcal{A}^3_{\text{any}}$ are better than under $\mathcal{A}^2_{\text{any}}$, which is remarkably visible in the high-input skew set with 15 Byzantine faults where the median intra-layer skew is at 0.51 ns cf. Figure 4.9. Furthermore, even though the IQD^0 is larger with $\mathcal{A}^3_{\text{any}}$, the IQD^1 and IQD^5 are both smaller than for the other two algorithms. Further, we notice that $\mathcal{A}^3_{\text{any}}$'s tendency for higher inter-layer skews reduces the number of negative inter-layer skews.

The detailed violin plots of $\mathcal{G}_A$ can be found in Appendix A.2 on Page 242pp.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.4: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_A$ topology with input skews 0 and ε.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^2_{\text{adj}}$	0	0	0	0.01	0.03	0.45	0.35	1.21	1.83	4.79	2.78	6.38	7.01	7.79	7.79	8.60	9.19	12.43
		15	0	0.01	0.04	1.28	0.45	6.43	7.48	13.18	−5.49	0.95	5.02	7.81	7.80	11.06	14.80	22.71
		15	1	0.01	0.04	1.23	0.44	6.22	7.30	8.20	−0.94	1.14	5.37	7.79	7.80	10.14	14.38	16.29
$\mathcal{A}^2_{\text{any}}$	0	0	0	0.01	0.02	0.29	0.25	0.71	0.92	2.24	5.40	6.91	7.18	7.68	7.68	8.18	8.45	9.98
		15	0	0.01	0.02	0.31	0.26	0.75	1.00	2.86	4.89	6.85	7.17	7.67	7.68	8.19	8.49	10.42
		15	1	0.01	0.02	0.30	0.25	0.72	0.95	2.84	4.89	6.89	7.18	7.68	7.68	8.18	8.46	10.10
$\mathcal{A}^3_{\text{any}}$	0	0	0	0.00	0.02	0.30	0.25	0.75	0.99	2.32	7.16	7.29	7.48	8.03	8.02	8.63	8.94	10.67
		15	0	0.01	0.03	1.17	0.37	7.05	7.78	19.61	−4.16	7.22	7.42	8.88	8.08	14.95	16.43	38.95
		15	1	0.01	0.03	1.00	0.35	6.43	7.48	11.41	−0.09	7.24	7.44	8.75	8.08	14.53	15.94	24.20
$\mathcal{A}^2_{\text{adj}}$	ε	0	0	0.01	0.05	0.92	0.66	2.62	3.78	7.88	−0.26	4.73	5.98	7.75	7.77	9.46	10.72	15.65
		15	0	0.01	0.06	1.52	0.73	6.43	7.55	14.73	−6.83	0.85	4.75	7.78	7.79	10.99	14.81	22.78
		15	1	0.01	0.06	1.46	0.72	6.19	7.39	8.20	−0.98	1.02	4.94	7.76	7.79	10.55	14.45	16.27
$\mathcal{A}^2_{\text{any}}$	ε	0	0	0.01	0.05	0.71	0.53	1.99	2.80	6.13	1.71	5.42	6.34	7.68	7.68	9.02	9.94	13.77
		15	0	0.01	0.04	0.72	0.51	2.18	3.42	8.01	−0.36	5.09	6.30	7.67	7.67	9.03	10.23	16.11
		15	1	0.01	0.04	0.72	0.50	2.12	3.31	7.69	−0.12	5.21	6.34	7.68	7.68	9.01	10.14	15.14
$\mathcal{A}^3_{\text{any}}$	ε	0	0	0.01	0.03	0.41	0.30	1.23	2.00	5.36	7.16	7.26	7.45	8.14	8.05	9.22	10.26	13.49
		15	0	0.01	0.04	1.23	0.43	7.03	7.79	21.06	−4.73	7.21	7.40	8.95	8.11	14.94	16.63	38.77
		15	1	0.01	0.03	1.07	0.42	6.36	7.48	12.53	−0.25	7.23	7.42	8.82	8.11	14.49	16.02	24.22

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.5: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_A$ topology with input skews 2ε and d^+.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^2_{\text{adj}}$	2ε	0	0	0.02	0.10	1.76	1.51	4.43	5.88	8.19	−0.93	2.80	4.47	7.72	7.75	10.93	12.61	16.11
		15	0	0.02	0.08	2.04	1.38	6.77	7.64	19.26	−11.89	0.59	3.64	7.76	7.77	12.00	14.93	22.18
		15	1	0.02	0.08	1.99	1.37	6.59	7.51	10.63	−4.59	0.72	3.81	7.74	7.77	11.63	14.68	16.32
$\mathcal{A}^2_{\text{any}}$	2ε	0	0	0.02	0.08	1.46	1.12	3.89	5.15	8.15	−0.79	3.37	4.86	7.68	7.68	10.50	11.99	15.90
		15	0	0.01	0.07	1.47	0.88	4.83	6.91	14.41	−7.88	2.00	4.72	7.66	7.67	10.58	13.29	16.96
		15	1	0.01	0.07	1.46	0.90	4.71	6.77	8.20	−0.92	2.22	4.81	7.67	7.68	10.53	13.12	16.12
$\mathcal{A}^3_{\text{any}}$	2ε	0	0	0.01	0.03	0.53	0.31	2.14	3.15	7.06	7.16	7.25	7.45	8.27	8.06	10.25	12.24	16.31
		15	0	0.01	0.04	1.32	0.46	7.04	7.79	20.83	−5.11	7.21	7.40	9.05	8.12	14.95	16.70	37.79
		15	1	0.01	0.04	1.16	0.44	6.39	7.49	13.21	−0.28	7.23	7.41	8.93	8.12	14.51	16.04	24.14
$\mathcal{A}^2_{\text{adj}}$	d^+	0	0	0.08	0.47	6.33	7.13	7.90	8.11	8.54	−9.07	−0.96	−0.24	7.52	7.60	15.15	15.50	16.33
		15	0	0.04	0.23	5.53	6.85	7.93	8.16	30.87	−22.91	−3.62	−0.29	7.40	7.62	15.12	15.53	34.12
		15	1	0.05	0.25	5.54	6.85	7.91	8.13	22.48	−17.79	−2.14	−0.25	7.43	7.62	15.10	15.48	16.33
$\mathcal{A}^2_{\text{any}}$	d^+	0	0	0.04	0.23	5.83	6.98	7.88	8.10	8.40	−9.07	−0.92	−0.19	7.51	7.58	15.12	15.48	16.33
		15	0	0.03	0.15	4.88	6.42	7.90	8.14	36.75	−29.53	−4.43	−0.25	7.29	7.58	15.03	15.47	32.97
		15	1	0.03	0.16	4.92	6.46	7.88	8.11	29.47	−28.28	−2.45	−0.20	7.37	7.59	15.02	15.44	16.33
$\mathcal{A}^3_{\text{any}}$	d^+	0	0	0.01	0.03	1.26	0.33	7.57	8.01	8.59	−0.98	6.86	7.39	9.03	8.06	15.64	22.66	24.48
		15	0	0.01	0.04	2.07	0.51	7.71	8.08	26.03	−16.01	6.44	7.32	9.75	8.13	17.31	22.75	51.96
		15	1	0.01	0.04	1.93	0.48	7.65	8.04	23.52	−7.69	6.61	7.34	9.67	8.13	17.08	22.70	32.09

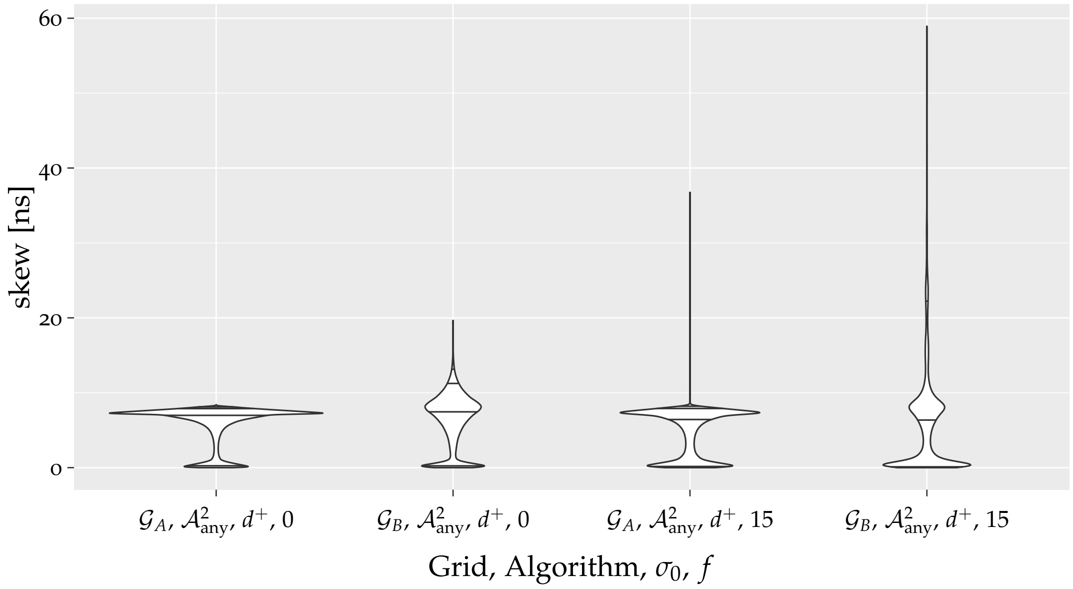


Figure 4.10: Comparison of the intra-layer skews of $\mathcal{G}_A$ and $\mathcal{G}_B$ with $\mathcal{A}^2_{\text{any}}$. We see that the inter-layer skews in $\mathcal{G}_B$ show a strong degradation in the set with $f = 15$ Byzantine faults and $\sigma_0 = d^+$.

4.3.3 Results of the Pulse-based Algorithms in $\mathcal{G}_B$

With $\mathcal{A}^2_{\text{any}}$ being the only algorithm applicable in $\mathcal{G}_B$, we compare $\mathcal{G}_B$ with $\mathcal{G}_A$ under this algorithm. This reveals a similar behavior under well-behaved input skews, as can be seen by comparing Tables 4.4 and 4.5 with Table 4.6. Although the IQD^0 is larger with $\mathcal{G}_B$, the IQD^1 and IQD^5 are similar in both grids. However, with $\sigma_0 = d^+$ we see that the skews are degrading, as shown in Figure 4.10: $\mathcal{G}_B$ lacks the intra-layer links which allow an earlier triggering of a node in a heavily skewed scenario. This missing influence from the direct neighbors significantly increases the impact that faults have on the skews. Compared to $\mathcal{G}_H$ and $\mathcal{G}_A$, even the removal of the direct neighborhood of the faults from the skew calculation ($h = 1$) reveals no changes to the IQD^0 and only slightly improves the other parameters of the distribution.

The detailed violin plots of $\mathcal{G}_B$ can be found in Appendix A.3 on Page 249pp.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.6: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\mathrm{op}}$ (in [ns]) in the $\mathcal{G}_B$ topology.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^2_{\mathrm{any}}$	0	0	0	0.01	0.02	0.29	0.25	0.71	0.92	2.15	5.55	6.91	7.18	7.68	7.68	8.18	8.45	9.77
		15	0	0.01	0.02	0.30	0.25	0.73	0.97	3.11	5.14	6.89	7.18	7.68	7.68	8.18	8.47	10.25
		15	1	0.01	0.02	0.30	0.25	0.72	0.94	2.47	5.19	6.91	7.18	7.68	7.68	8.18	8.45	9.93
$\mathcal{A}^2_{\mathrm{any}}$	ε	0	0	0.01	0.05	0.71	0.52	2.00	2.82	5.83	1.98	5.41	6.33	7.68	7.68	9.02	9.95	13.47
		15	0	0.01	0.05	0.72	0.51	2.13	3.25	7.97	−0.67	5.20	6.33	7.68	7.68	9.03	10.15	15.30
		15	1	0.01	0.05	0.71	0.51	2.10	3.17	7.65	0.24	5.26	6.35	7.68	7.68	9.01	10.09	15.30
$\mathcal{A}^2_{\mathrm{any}}$	2ε	0	0	0.02	0.08	1.46	1.11	3.90	5.17	9.36	−1.70	3.35	4.86	7.68	7.68	10.50	12.01	16.92
		15	0	0.01	0.07	1.47	0.91	4.60	7.03	14.07	−6.28	2.28	4.83	7.68	7.68	10.53	13.07	21.84
		15	1	0.01	0.07	1.46	0.92	4.52	6.91	13.75	−5.77	2.42	4.87	7.68	7.68	10.48	12.93	21.60
$\mathcal{A}^2_{\mathrm{any}}$	d^+	0	0	0.04	0.19	6.57	7.43	11.24	13.13	19.65	−11.82	−4.16	−1.98	7.67	7.68	17.34	19.53	27.30
		15	0	0.02	0.11	6.50	6.10	22.03	30.60	58.92	−50.95	−16.68	−3.00	7.67	7.68	18.36	32.02	66.71
		15	1	0.02	0.12	6.49	6.19	21.73	30.28	58.92	−49.17	−16.39	−2.74	7.67	7.68	18.09	31.75	64.48

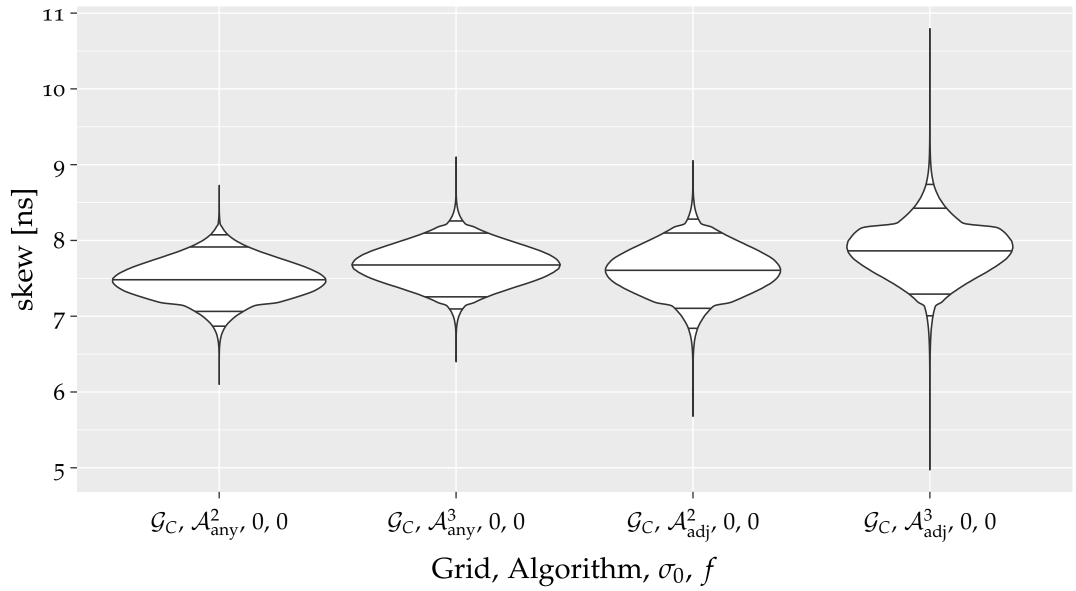


Figure 4.11: The comparison of the algorithms in $\mathcal{G}_C$ under $\sigma_0 = 0$ reveals a similar behavior on the inter-layer skews, with the exception of $\mathcal{A}^3_{\text{adj}}$.

4.3.4 Results of the Pulse-based Algorithms in $\mathcal{G}_C$

Through its increased number of links, $\mathcal{G}_C$ can be used in combination with all algorithms presented above. This multitude of options does not simplify the evaluation, however. Hence, we first compare the four applicable algorithms in a benign setting, shown in Figure 4.11. This immediately reveals that $\mathcal{A}^3_{\text{adj}}$ performs poorly in this grid. However, this does not mean that waiting for three neighbors is ineffective in general, as shown by $\mathcal{A}^3_{\text{any}}$. Even tough $\mathcal{A}^3_{\text{any}}$ preforms not as good as $\mathcal{A}^2_{\text{any}}$ in high input skew settings, the differences are small in well-behaved settings as can be seen in Tables 4.7 and 4.8. Given the increased fault-tolerance of this algorithm, and the restrictions that come with it, this is an impressive result.

Comparing the results of $\mathcal{A}^2_{\text{any}}$ and $\mathcal{A}^2_{\text{adj}}$ with the previously presented grids shows that they deliver slightly better results in $\mathcal{G}_C$ as they did in $\mathcal{G}_H$ and $\mathcal{G}_A$. This is unsurprising, however, as the additional links of this grid increase the number of firing rules and hence limit the effect of skewed inputs.

Comparing the behavior of $\mathcal{G}_A$ and $\mathcal{G}_C$ under $\mathcal{A}^2_{\text{any}}$ and $\mathcal{A}^3_{\text{any}}$, however, reveals an interesting correlation between the pairwise different settings under high input skews, cf. Figure 4.12. The skew distribution of $\mathcal{G}_C$ with $\mathcal{A}^2_{\text{any}}$ exhibits very low inter-layer minimum skews, as the nodes can select from a spatially wide set of neighbors, thus enabling an early triggering. Even though $\mathcal{G}_A$ shows a similar intra-layer skew behavior with $\mathcal{A}^3_{\text{any}}$, this is not true for the inter-layer skew: The reason is that the comparatively low number of firing rules force the node to wait longer, which is clearly visible by the

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

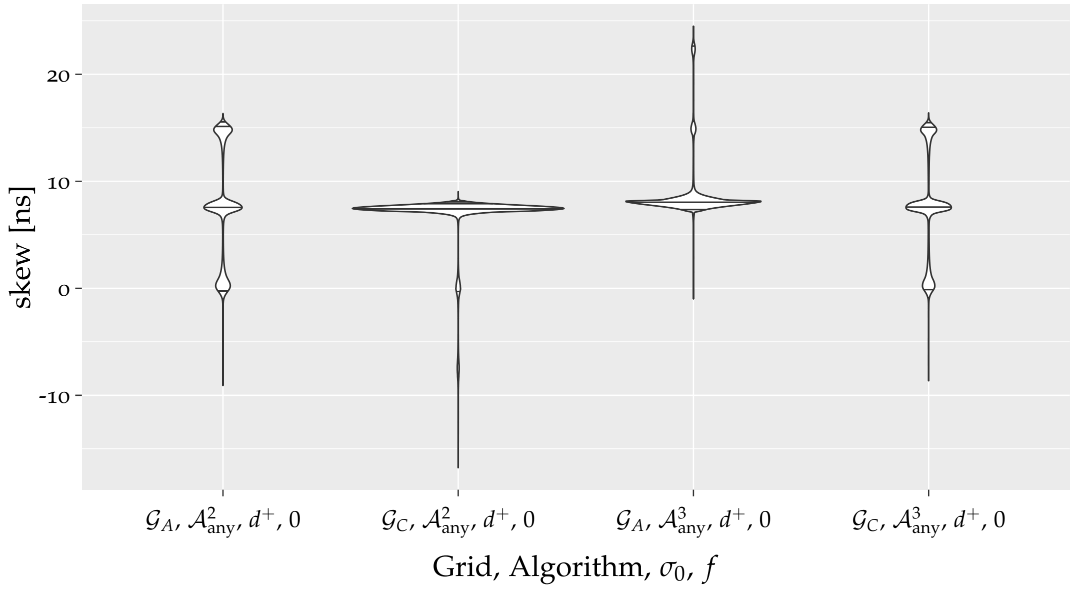


Figure 4.12: Comparison of the inter-layer skew of the $\mathcal{A}^2_{\text{any}}$ and $\mathcal{A}^3_{\text{any}}$ algorithms in $\mathcal{G}_A$ and $\mathcal{G}_C$ under $\sigma_0 = 0$.

minimum inter-layer skew of ≈ -1 ns. The resemblance between $\mathcal{G}_A$ with $\mathcal{A}^2_{\text{any}}$ and $\mathcal{G}_C$ with $\mathcal{A}^3_{\text{any}}$ is striking: The additional link required to fire by $\mathcal{A}^3_{\text{any}}$ effectively reduces the effect of the high input skew in the grid.

It is interesting to note that for $\sigma_0 = d^+$, the minimum inter-layer skews are prominently lower than in $\mathcal{G}_A$. This can be attributed to the high number of links that enable the nodes in the grid to trigger earlier, even when its intra-layer neighbors exhibit large skews: While a part of the neighbors has already been triggered, the others are still waiting for trigger messages to be delivered due to their skewed inputs.

The detailed violin plots of $\mathcal{G}_C$ can be found in Appendix A.4 on Page 252pp.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.7: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology with input skews 0 and ε.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^2_{\text{adj}}$	0	0	0	0.00	0.02	0.27	0.23	0.65	0.85	2.00	5.68	6.84	7.11	7.61	7.61	8.10	8.28	9.05
		15	0	0.00	0.02	0.28	0.24	0.68	0.90	2.45	5.37	6.80	7.08	7.60	7.61	8.11	8.31	10.02
		15	1	0.00	0.02	0.27	0.23	0.66	0.86	2.25	5.38	6.82	7.10	7.60	7.61	8.10	8.29	10.02
$\mathcal{A}^3_{\text{adj}}$	0	0	0	0.01	0.02	0.31	0.25	0.77	1.08	2.85	4.97	7.01	7.30	7.86	7.87	8.43	8.75	10.79
		15	0	0.01	0.03	1.04	0.33	6.10	7.42	13.74	−6.04	1.24	6.34	7.90	7.88	9.96	14.75	22.87
		15	1	0.01	0.03	0.98	0.32	5.78	7.18	8.20	−1.43	1.51	6.57	7.86	7.88	9.10	14.10	16.32
$\mathcal{A}^2_{\text{any}}$	0	0	0	0.00	0.02	0.22	0.19	0.54	0.69	1.35	6.10	6.87	7.07	7.49	7.48	7.92	8.08	8.72
		15	0	0.00	0.02	0.23	0.19	0.56	0.72	1.69	5.57	6.83	7.04	7.48	7.48	7.92	8.09	9.05
		15	1	0.00	0.02	0.22	0.19	0.54	0.70	1.54	5.82	6.85	7.05	7.48	7.48	7.92	8.08	8.86
$\mathcal{A}^3_{\text{any}}$	0	0	0	0.00	0.02	0.22	0.19	0.54	0.68	1.26	6.40	7.10	7.26	7.68	7.68	8.10	8.26	9.10
		15	0	0.00	0.02	0.23	0.19	0.55	0.70	1.54	6.27	7.08	7.25	7.68	7.68	8.11	8.27	9.11
		15	1	0.00	0.02	0.22	0.19	0.54	0.69	1.31	6.27	7.09	7.26	7.68	7.68	8.10	8.26	9.11
$\mathcal{A}^2_{\text{adj}}$	ε	0	0	0.01	0.02	0.35	0.26	0.98	1.78	4.35	2.47	5.36	6.60	7.50	7.58	8.11	8.29	9.14
		15	0	0.01	0.02	0.36	0.27	1.02	1.87	5.91	0.66	5.31	6.58	7.50	7.57	8.11	8.32	12.40
		15	1	0.01	0.02	0.35	0.26	1.00	1.82	5.70	1.33	5.32	6.57	7.50	7.58	8.11	8.30	12.33
$\mathcal{A}^3_{\text{adj}}$	ε	0	0	0.01	0.04	0.72	0.46	2.24	3.32	8.15	−0.47	5.14	6.30	7.80	7.84	9.15	10.33	15.74
		15	0	0.01	0.04	1.26	0.50	6.11	7.50	15.65	−7.61	1.10	5.47	7.86	7.86	10.43	14.79	22.71
		15	1	0.01	0.04	1.19	0.49	5.77	7.31	8.20	−5.13	1.36	5.63	7.82	7.86	9.92	14.21	16.24
$\mathcal{A}^2_{\text{any}}$	ε	0	0	0.00	0.02	0.30	0.22	0.88	1.69	3.94	2.76	5.39	6.58	7.39	7.46	7.94	8.12	9.03
		15	0	0.00	0.02	0.31	0.22	0.89	1.74	6.08	1.35	5.33	6.56	7.39	7.45	7.95	8.13	11.89
		15	1	0.00	0.02	0.30	0.22	0.88	1.71	5.39	1.35	5.36	6.57	7.39	7.45	7.94	8.12	10.01
$\mathcal{A}^3_{\text{any}}$	ε	0	0	0.01	0.03	0.49	0.33	1.58	2.32	4.81	3.10	5.85	6.72	7.68	7.68	8.64	9.51	12.60
		15	0	0.01	0.03	0.50	0.33	1.64	2.69	6.60	1.15	5.69	6.71	7.67	7.68	8.62	9.65	14.33
		15	1	0.01	0.03	0.49	0.33	1.60	2.58	6.60	1.24	5.77	6.74	7.68	7.68	8.61	9.57	14.32

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.8: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology with input skews 2ε and d^+.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^2_{\text{adj}}$	2ε	0	0	0.01	0.02	0.45	0.27	1.95	2.93	5.75	−0.65	3.28	5.52	7.39	7.57	8.10	8.29	9.15
		15	0	0.01	0.03	0.46	0.27	1.95	3.10	8.20	−4.17	3.20	5.54	7.38	7.57	8.11	8.32	15.17
		15	1	0.01	0.03	0.46	0.27	1.95	3.01	8.11	−3.55	3.22	5.52	7.38	7.57	8.11	8.30	15.08
$\mathcal{A}^3_{\text{adj}}$	2ε	0	0	0.01	0.06	1.50	1.10	4.17	5.63	8.20	−0.95	3.06	4.73	7.76	7.82	10.68	12.36	16.02
		15	0	0.01	0.05	1.76	0.88	6.68	7.66	19.48	−11.69	0.66	4.13	7.83	7.84	11.72	14.97	24.30
		15	1	0.01	0.05	1.69	0.86	6.43	7.49	11.43	−8.86	0.84	4.30	7.80	7.84	11.21	14.62	16.79
$\mathcal{A}^2_{\text{any}}$	2ε	0	0	0.00	0.02	0.40	0.23	1.85	2.88	5.92	−0.56	3.34	5.49	7.27	7.45	7.94	8.12	9.06
		15	0	0.00	0.02	0.40	0.23	1.81	3.00	8.20	−4.14	3.24	5.52	7.27	7.45	7.95	8.13	16.08
		15	1	0.00	0.02	0.40	0.23	1.82	2.93	8.11	−3.86	3.27	5.50	7.27	7.45	7.94	8.12	11.28
$\mathcal{A}^3_{\text{any}}$	2ε	0	0	0.01	0.04	1.13	0.63	3.48	4.70	8.17	−0.73	3.80	5.23	7.68	7.68	10.12	11.57	15.87
		15	0	0.01	0.04	1.13	0.53	4.14	6.49	13.02	−5.93	2.64	5.23	7.67	7.67	10.08	12.67	20.45
		15	1	0.01	0.04	1.11	0.53	4.01	6.34	8.20	−0.90	2.86	5.30	7.68	7.68	10.04	12.45	16.00
$\mathcal{A}^2_{\text{adj}}$	d^+	0	0	0.01	0.02	1.02	0.27	7.39	7.94	8.20	−16.65	−8.21	−0.32	6.68	7.56	8.10	8.29	9.15
		15	0	0.01	0.03	1.00	0.28	7.39	7.96	32.93	−27.70	−8.37	−0.29	6.68	7.56	8.11	8.32	16.97
		15	1	0.01	0.03	1.00	0.27	7.39	7.95	25.94	−27.70	−8.41	−0.34	6.66	7.56	8.10	8.30	16.21
$\mathcal{A}^3_{\text{adj}}$	d^+	0	0	0.03	0.14	5.68	6.94	7.87	8.10	8.22	−8.64	−0.87	−0.17	7.58	7.70	15.12	15.50	16.39
		15	0	0.02	0.11	4.93	6.43	7.93	8.54	41.27	−33.38	−3.86	−0.20	7.54	7.73	15.12	15.64	39.34
		15	1	0.02	0.11	4.87	6.41	7.88	8.14	34.02	−27.32	−2.14	−0.16	7.53	7.73	15.07	15.51	31.67
$\mathcal{A}^2_{\text{any}}$	d^+	0	0	0.00	0.02	0.88	0.23	7.35	7.92	8.20	−16.76	−8.19	−0.28	6.57	7.45	7.93	8.11	9.02
		15	0	0.00	0.02	0.87	0.23	7.33	7.95	25.85	−26.05	−8.44	−0.27	6.56	7.44	7.94	8.12	16.01
		15	1	0.00	0.02	0.88	0.23	7.34	7.94	18.36	−25.99	−8.42	−0.31	6.55	7.44	7.93	8.11	14.71
$\mathcal{A}^3_{\text{any}}$	d^+	0	0	0.02	0.10	5.03	6.58	7.83	8.08	8.20	−8.64	−0.82	−0.09	7.54	7.61	15.06	15.47	16.39
		15	0	0.02	0.07	4.03	4.88	7.83	8.13	32.48	−24.78	−3.10	−0.08	7.37	7.62	14.95	15.46	29.68
		15	1	0.02	0.08	4.02	4.92	7.80	8.09	24.43	−24.69	−1.50	−0.03	7.43	7.62	14.93	15.43	22.49

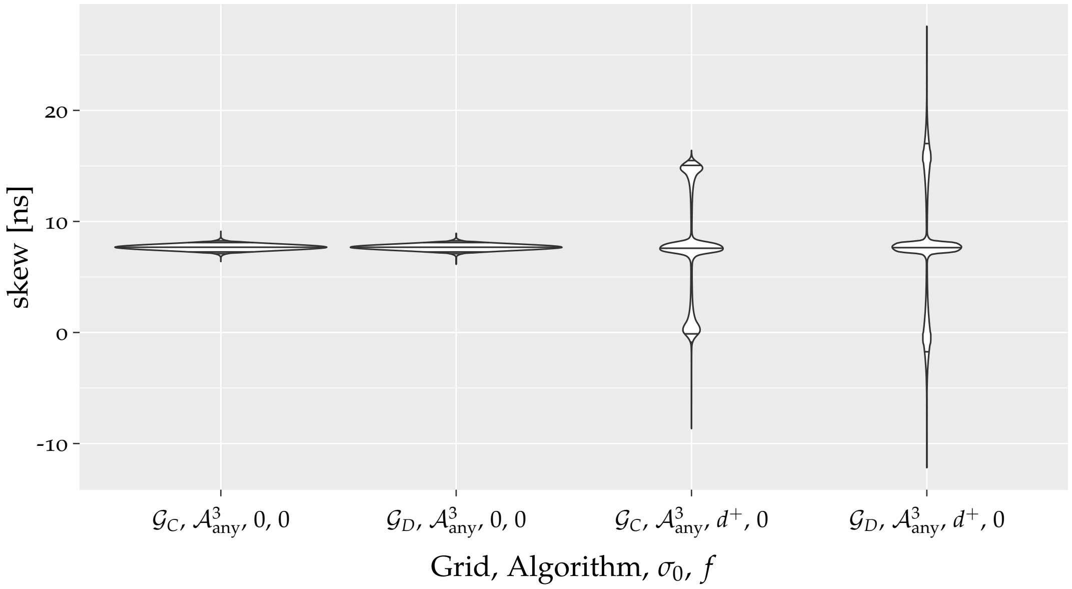


Figure 4.13: A comparison of the inter-layer skews between $\mathcal{G}_C$ and $\mathcal{G}_D$ with $\mathcal{A}^3_{\text{any}}$.

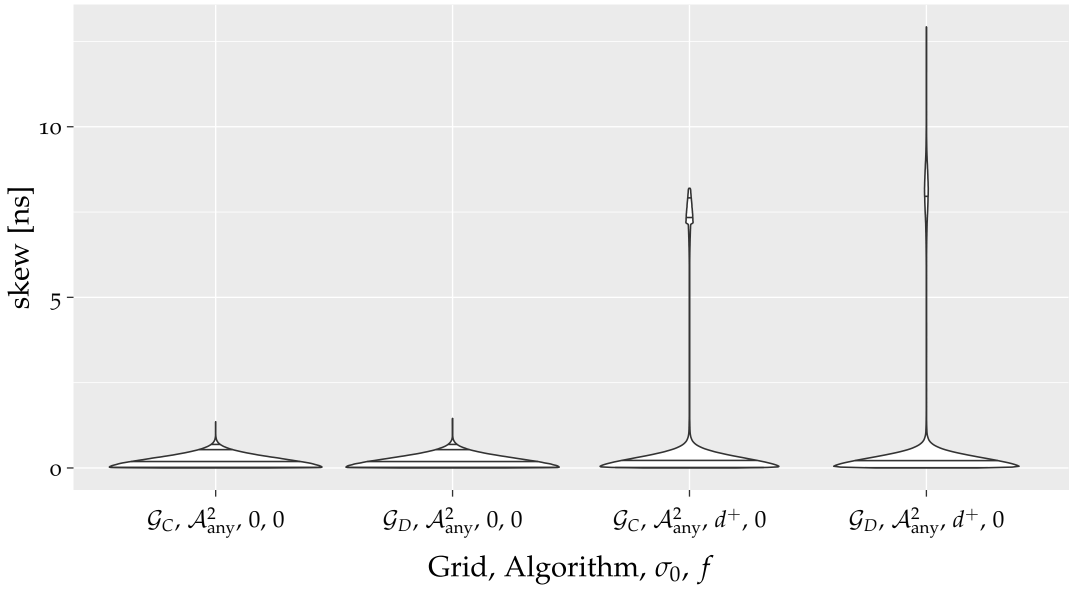
4.3.5 Results of the Pulse-based Algorithms in $\mathcal{G}_D$

We chose to compare $\mathcal{G}_D$ with $\mathcal{G}_C$, as the presence of the intra-layer links is the only difference between them. The comparison of the behavior of $\mathcal{A}^2_{\text{any}}$ shows that, without faults, in the well-behaved input skew sets, only minor differences appear, see Figure 4.14. This comes as no surprise, however, as the intra-layer links are not used in these cases. As soon as we consider high input skews, however, we see an increase in intra-layer skews due to the missing links. Moreover, the inter-layer skews minimum decreased, which, while counter-intuitive, make sense: The higher intra-layer skews indicate a skewed input for the nodes that allows them to trigger way before a fraction of their incoming inter-layer neighbors. The results for $\mathcal{A}^2_{\text{adj}}$ are similar.

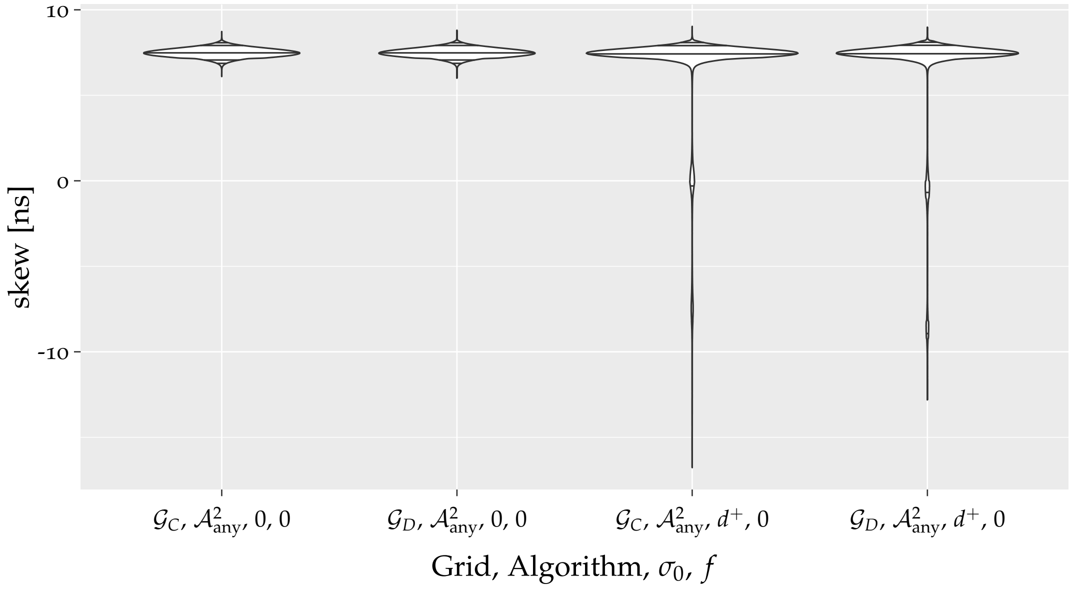
The other interesting algorithm in $\mathcal{G}_D$ is $\mathcal{A}^3_{\text{any}}$. As with $\mathcal{A}^2_{\text{any}}$, the missing intra-layer links show no noticeable impact in well-behaved input skew sets, as can be seen in Figure 4.13 and Table 4.9. As expected, with faults and higher input skews, this dramatically changes due to the reduced number of applicable firing rules. Hence, we see a strong increase in intra- and inter-layer skews that is especially prominent in sets with higher input skews. As we have already seen in $\mathcal{G}_B$, even the removal of the direct neighborhood of a fault in the evaluation $(h = 1)$ has no noticeable impact on the skews, as can be seen in Tables 4.9 and 4.10.

The detailed violin plots of $\mathcal{G}_D$ can be found in Appendix A.8 on Page 274pp.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) intra-layer skews



(b) inter-layer skews

Figure 4.14: A comparison of the skews between $\mathcal{G}_C$ and $\mathcal{G}_D$ with $\mathcal{A}^2_{\text{any}}$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.9: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_D$ topology with input skews 0 and ε.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^2_{\text{adj}}$	0	0	0	0.00	0.02	0.27	0.23	0.65	0.84	1.86	5.57	6.85	7.11	7.61	7.61	8.10	8.28	9.08
		15	0	0.00	0.02	0.28	0.23	0.67	0.88	2.15	5.34	6.82	7.09	7.61	7.61	8.11	8.30	9.43
		15	1	0.00	0.02	0.27	0.23	0.66	0.85	2.15	5.34	6.83	7.10	7.61	7.61	8.10	8.29	9.43
$\mathcal{A}^2_{\text{any}}$	0	0	0	0.00	0.02	0.22	0.19	0.54	0.69	1.45	6.00	6.87	7.07	7.49	7.48	7.92	8.08	8.81
		15	0	0.00	0.02	0.23	0.19	0.55	0.71	1.71	5.93	6.85	7.05	7.48	7.48	7.92	8.08	8.77
		15	1	0.00	0.02	0.22	0.19	0.54	0.70	1.39	6.10	6.86	7.06	7.49	7.48	7.92	8.08	8.77
$\mathcal{A}^3_{\text{any}}$	0	0	0	0.00	0.02	0.22	0.19	0.54	0.68	1.19	6.14	7.10	7.26	7.68	7.68	8.10	8.26	8.94
		15	0	0.00	0.02	0.23	0.19	0.54	0.70	1.41	6.24	7.09	7.25	7.68	7.68	8.10	8.27	9.19
		15	1	0.00	0.02	0.22	0.19	0.54	0.69	1.25	6.45	7.10	7.26	7.68	7.68	8.10	8.26	8.95
$\mathcal{A}^2_{\text{adj}}$	ε	0	0	0.01	0.02	0.35	0.26	0.98	1.78	4.37	2.58	5.36	6.60	7.50	7.58	8.11	8.29	9.11
		15	0	0.01	0.02	0.36	0.26	1.01	1.84	5.65	0.78	5.32	6.58	7.50	7.58	8.11	8.31	11.42
		15	1	0.01	0.02	0.35	0.26	1.00	1.81	5.41	1.22	5.33	6.58	7.50	7.58	8.11	8.30	11.31
$\mathcal{A}^2_{\text{any}}$	ε	0	0	0.00	0.02	0.30	0.22	0.88	1.69	3.93	2.87	5.39	6.58	7.39	7.46	7.94	8.12	9.09
		15	0	0.00	0.02	0.31	0.22	0.89	1.73	6.40	1.07	5.36	6.57	7.39	7.45	7.95	8.12	10.72
		15	1	0.00	0.02	0.30	0.22	0.88	1.70	4.64	1.07	5.37	6.56	7.39	7.46	7.94	8.12	9.35
$\mathcal{A}^3_{\text{any}}$	ε	0	0	0.01	0.03	0.49	0.33	1.58	2.32	4.94	2.96	5.85	6.72	7.68	7.68	8.64	9.51	12.68
		15	0	0.01	0.03	0.50	0.33	1.63	2.59	6.75	1.78	5.74	6.73	7.68	7.68	8.63	9.62	14.05
		15	1	0.01	0.03	0.49	0.33	1.60	2.52	5.75	1.78	5.80	6.74	7.68	7.68	8.62	9.56	13.53

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.10: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_D$ topology with input skews 2ε and d^+.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^2_{\text{adj}}$	2ε	0	0	0.01	0.02	0.45	0.27	1.95	2.93	6.61	0.03	3.28	5.52	7.39	7.57	8.10	8.29	9.16
		15	0	0.01	0.03	0.46	0.27	1.96	3.05	12.25	−6.15	3.21	5.54	7.38	7.57	8.11	8.31	15.61
		15	1	0.01	0.02	0.46	0.27	1.96	3.00	10.83	−3.66	3.22	5.50	7.38	7.57	8.11	8.30	14.26
$\mathcal{A}^2_{\text{any}}$	2ε	0	0	0.00	0.02	0.39	0.23	1.84	2.89	5.50	−0.28	3.33	5.49	7.27	7.45	7.94	8.12	9.03
		15	0	0.00	0.02	0.40	0.23	1.82	2.98	11.41	−3.98	3.26	5.52	7.27	7.45	7.95	8.13	14.63
		15	1	0.00	0.02	0.40	0.23	1.83	2.93	9.92	−3.91	3.28	5.49	7.27	7.45	7.94	8.12	9.83
$\mathcal{A}^3_{\text{any}}$	2ε	0	0	0.01	0.04	1.13	0.62	3.49	4.72	8.82	−1.00	3.78	5.23	7.68	7.68	10.13	11.58	16.53
		15	0	0.01	0.04	1.13	0.54	4.02	6.50	12.75	−4.98	2.85	5.26	7.68	7.68	10.09	12.51	20.87
		15	1	0.01	0.04	1.12	0.54	3.92	6.36	12.15	−4.85	3.01	5.30	7.68	7.68	10.05	12.35	19.95
$\mathcal{A}^2_{\text{adj}}$	d^+	0	0	0.01	0.02	1.09	0.27	8.09	9.08	12.68	−12.73	−8.97	−0.67	6.68	7.57	8.10	8.29	9.21
		15	0	0.01	0.03	1.11	0.28	8.08	9.26	40.80	−41.47	−9.06	−0.64	6.67	7.57	8.11	8.31	39.93
		15	1	0.01	0.03	1.11	0.27	8.09	9.20	40.51	−40.66	−9.04	−0.68	6.66	7.57	8.11	8.30	33.89
$\mathcal{A}^2_{\text{any}}$	d^+	0	0	0.00	0.02	0.93	0.23	7.97	9.03	12.92	−12.81	−8.91	−0.66	6.57	7.45	7.94	8.12	8.98
		15	0	0.00	0.02	0.95	0.23	7.94	9.18	41.64	−41.81	−8.99	−0.64	6.56	7.45	7.95	8.13	33.27
		15	1	0.00	0.02	0.94	0.23	7.95	9.11	41.54	−41.40	−8.97	−0.66	6.56	7.45	7.95	8.12	16.93
$\mathcal{A}^3_{\text{any}}$	d^+	0	0	0.02	0.09	5.66	6.79	10.94	12.87	20.05	−12.17	−3.88	−1.70	7.68	7.68	17.05	19.24	27.57
		15	0	0.01	0.06	5.55	2.34	20.27	31.82	67.93	−60.33	−16.35	−2.56	7.68	7.68	17.91	31.69	75.83
		15	1	0.01	0.06	5.47	2.37	19.64	31.50	67.93	−60.33	−15.94	−2.24	7.68	7.68	17.59	31.30	75.55

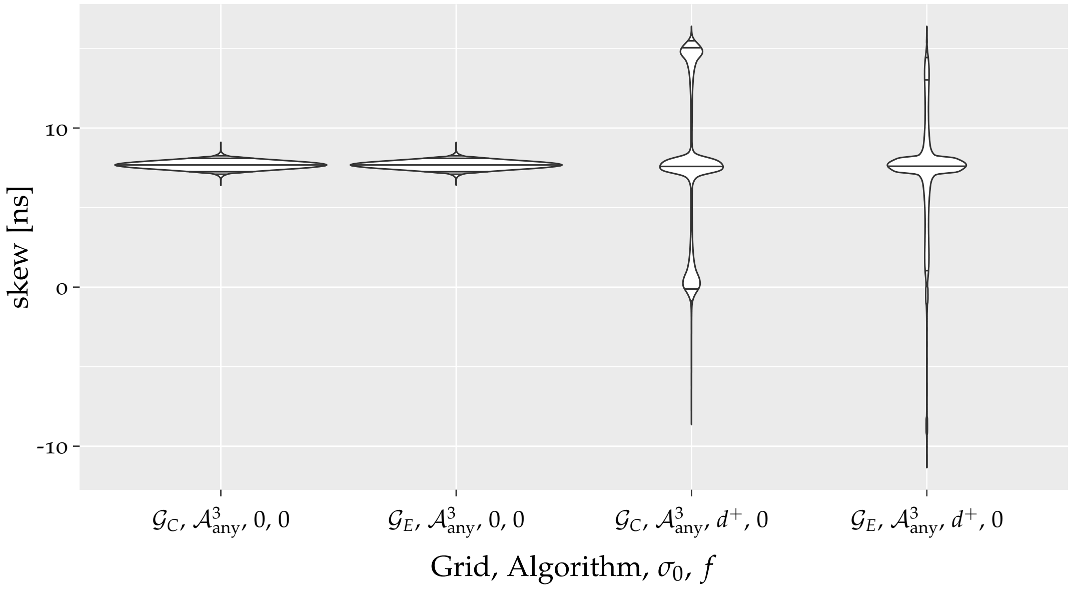


Figure 4.15: A comparison of the inter-layer skews between $\mathcal{G}_C$ and $\mathcal{G}_E$ under $\mathcal{A}^3_{\text{any}}$.

4.3.6 Results of the Pulse-based Algorithms in $\mathcal{G}_E$

The reason behind the construction of $\mathcal{G}_E$ was to investigate the effect of extended intra-layer links on the skew behavior of the grid. With $\mathcal{G}_C$ being structurally the closest grid, a comparison with it will provide an answer to this question.

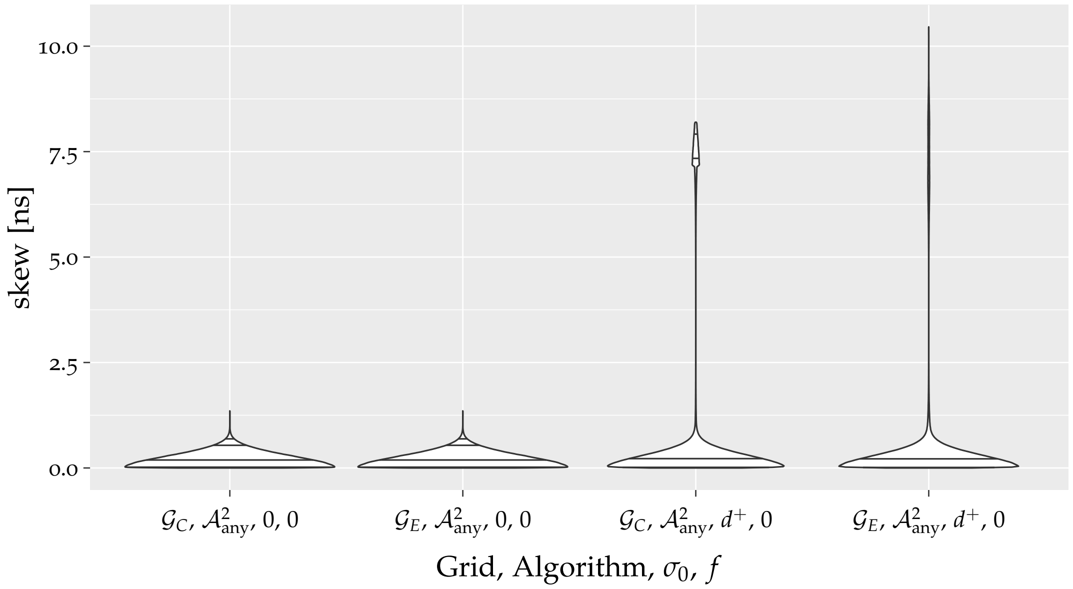
We focus on the behavior under $\mathcal{A}^2_{\text{any}}$ and $\mathcal{A}^3_{\text{any}}$, as the results with $\mathcal{A}^2_{\text{adj}}$ and $\mathcal{A}^3_{\text{adj}}$ are similar. Figure 4.16 shows the plots of the comparison under $\mathcal{A}^2_{\text{any}}$, with the expected behavior that in the low input skew set, there is no noticeable difference. However, in the high input skew set, $\mathcal{G}_E$ shows a clear increase of the maximum intra-layer skew as well as a decrease of the minimum inter-layer skew. This is unsurprising as the wider reach of the nodes in this grid with such a skewed input, with just two neighbors required for firing, will facilitate an earlier triggering time.

The properties of $\mathcal{A}^3_{\text{any}}$ are similar to $\mathcal{A}^2_{\text{any}}$: No noteworthy effect in well-behaved input skew sets, whereas the high input skew set shows a lower IQD^5 that is clearly visible in the absence of the skew clusters at 0 and $2d^+$ for the inter-layer skews, shown in Figure 4.15. However, the minimum inter-layer skew is again lower compared to $\mathcal{G}_C$.

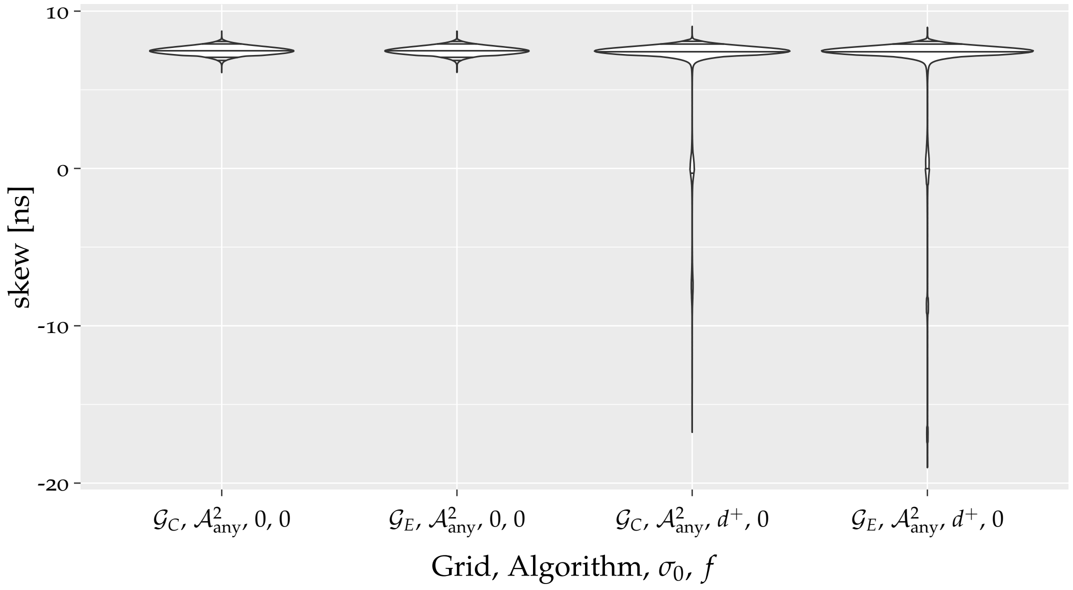
We conclude that the effect of wider reaching intra-layer links increased IQD^0, yet can reduce the higher IQD’s in some sets, as can be seen by comparing Tables 4.7 and 4.8 with Tables 4.11 and 4.12. Hence, $\mathcal{G}_E$ shows that these wider intra-layer links do not provide an general advantage.

The detailed violin plots of $\mathcal{G}_E$ can be found in Appendix A.10 on Page 290pp.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skews



(b) inter-layer skews

Figure 4.16: A comparison of the skews between $\mathcal{G}_C$ and $\mathcal{G}_E$ under $\mathcal{A}^2_{\text{any}}$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.11: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_E$ topology with input skews 0 and ε.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^2_{\text{adj}}$	0	0	0	0.00	0.02	0.27	0.23	0.65	0.85	2.00	5.68	6.84	7.11	7.61	7.61	8.10	8.28	9.05
		15	0	0.00	0.02	0.28	0.23	0.67	0.88	2.17	5.38	6.81	7.09	7.60	7.61	8.11	8.30	9.31
		15	1	0.00	0.02	0.27	0.23	0.66	0.86	2.02	5.38	6.82	7.10	7.60	7.61	8.10	8.29	9.23
$\mathcal{A}^3_{\text{adj}}$	0	0	0	0.01	0.02	0.31	0.25	0.77	1.08	2.85	4.97	7.01	7.30	7.86	7.87	8.43	8.75	10.79
		15	0	0.01	0.03	1.00	0.33	5.92	7.32	14.55	−3.30	1.38	6.50	7.89	7.88	9.64	14.57	22.60
		15	1	0.01	0.03	0.94	0.32	5.61	7.02	8.56	−0.76	1.66	6.69	7.86	7.88	8.97	13.94	16.39
$\mathcal{A}^2_{\text{any}}$	0	0	0	0.00	0.02	0.22	0.19	0.54	0.69	1.35	6.10	6.87	7.07	7.49	7.48	7.92	8.08	8.72
		15	0	0.00	0.02	0.23	0.19	0.56	0.73	1.64	5.79	6.83	7.04	7.48	7.48	7.92	8.09	8.93
		15	1	0.00	0.02	0.22	0.19	0.55	0.70	1.59	5.92	6.85	7.05	7.48	7.48	7.92	8.08	8.93
$\mathcal{A}^3_{\text{any}}$	0	0	0	0.00	0.02	0.22	0.19	0.54	0.68	1.26	6.40	7.10	7.26	7.68	7.68	8.10	8.26	9.10
		15	0	0.00	0.02	0.23	0.19	0.55	0.70	1.43	6.27	7.08	7.25	7.68	7.68	8.11	8.27	9.09
		15	1	0.00	0.02	0.22	0.19	0.54	0.69	1.25	6.37	7.09	7.25	7.68	7.68	8.10	8.27	9.02
$\mathcal{A}^2_{\text{adj}}$	ε	0	0	0.01	0.02	0.35	0.26	0.98	1.78	4.35	2.47	5.36	6.60	7.50	7.58	8.11	8.29	9.14
		15	0	0.01	0.02	0.35	0.26	1.00	1.84	5.57	0.66	5.31	6.59	7.50	7.58	8.11	8.31	10.86
		15	1	0.01	0.02	0.35	0.26	0.99	1.82	5.40	1.70	5.32	6.58	7.50	7.58	8.11	8.30	10.52
$\mathcal{A}^3_{\text{adj}}$	ε	0	0	0.01	0.04	0.72	0.46	2.24	3.32	7.59	0.38	5.14	6.30	7.80	7.84	9.15	10.33	15.43
		15	0	0.01	0.04	1.16	0.50	5.49	7.18	14.52	−6.52	1.65	5.64	7.85	7.86	10.15	14.24	22.39
		15	1	0.01	0.04	1.11	0.49	5.18	6.90	9.25	−1.20	1.90	5.76	7.82	7.86	9.77	13.64	16.94
$\mathcal{A}^2_{\text{any}}$	ε	0	0	0.00	0.02	0.30	0.22	0.88	1.69	3.94	2.76	5.39	6.58	7.39	7.46	7.94	8.12	9.03
		15	0	0.00	0.02	0.31	0.22	0.89	1.75	5.83	1.07	5.33	6.57	7.39	7.45	7.95	8.13	11.93
		15	1	0.00	0.02	0.30	0.22	0.87	1.71	5.51	1.07	5.35	6.57	7.39	7.45	7.94	8.12	9.38
$\mathcal{A}^3_{\text{any}}$	ε	0	0	0.01	0.03	0.49	0.33	1.58	2.32	4.81	3.10	5.85	6.72	7.68	7.68	8.64	9.51	12.60
		15	0	0.01	0.03	0.50	0.33	1.65	2.67	6.67	0.93	5.67	6.71	7.67	7.68	8.63	9.65	14.44
		15	1	0.01	0.03	0.49	0.33	1.62	2.58	6.61	0.93	5.74	6.72	7.67	7.68	8.61	9.59	14.44

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.12: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_E$ topology with input skews 2ε and d^+.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^2_{\text{adj}}$	2ε	0	0	0.01	0.02	0.45	0.27	1.95	2.93	5.75	−0.22	3.28	5.52	7.39	7.57	8.10	8.29	9.15
		15	0	0.01	0.03	0.46	0.27	1.94	3.10	10.56	−2.87	3.22	5.53	7.38	7.57	8.11	8.31	13.49
		15	1	0.01	0.02	0.45	0.27	1.95	3.05	8.74	−2.77	3.22	5.50	7.38	7.57	8.11	8.30	12.98
$\mathcal{A}^3_{\text{adj}}$	2ε	0	0	0.01	0.06	1.48	1.11	4.09	5.36	8.18	−0.47	3.18	4.74	7.76	7.82	10.66	12.23	15.81
		15	0	0.01	0.05	1.62	0.90	5.63	7.18	15.10	−7.26	1.54	4.40	7.81	7.84	11.26	14.19	22.46
		15	1	0.01	0.05	1.58	0.90	5.41	6.94	12.66	−5.22	1.75	4.50	7.78	7.83	10.97	13.74	20.40
$\mathcal{A}^2_{\text{any}}$	2ε	0	0	0.00	0.02	0.40	0.23	1.85	2.88	5.92	−0.17	3.34	5.49	7.27	7.45	7.94	8.12	9.06
		15	0	0.00	0.02	0.40	0.23	1.80	3.05	9.81	−5.53	3.24	5.52	7.27	7.45	7.95	8.13	14.07
		15	1	0.00	0.02	0.40	0.23	1.81	2.96	7.97	−3.37	3.26	5.50	7.27	7.45	7.94	8.12	11.16
$\mathcal{A}^3_{\text{any}}$	2ε	0	0	0.01	0.04	1.13	0.63	3.47	4.66	7.60	0.23	3.81	5.23	7.68	7.68	10.12	11.55	15.22
		15	0	0.01	0.04	1.12	0.54	4.03	5.93	8.96	−2.65	2.92	5.20	7.66	7.67	10.09	12.35	16.49
		15	1	0.01	0.04	1.10	0.54	3.91	5.79	8.31	−1.01	3.08	5.25	7.67	7.67	10.05	12.23	15.89
$\mathcal{A}^2_{\text{adj}}$	d^+	0	0	0.01	0.02	0.79	0.27	6.60	8.36	10.48	−19.01	−11.20	0.03	6.68	7.56	8.10	8.29	9.12
		15	0	0.01	0.02	0.79	0.27	6.62	8.40	34.18	−42.70	−11.53	−0.00	6.68	7.56	8.11	8.31	25.49
		15	1	0.01	0.02	0.79	0.27	6.64	8.39	34.18	−42.70	−15.83	−0.06	6.65	7.56	8.10	8.30	25.26
$\mathcal{A}^3_{\text{adj}}$	d^+	0	0	0.02	0.10	2.98	2.35	7.22	8.55	11.07	−11.45	−8.34	1.00	7.31	7.72	13.34	14.58	16.63
		15	0	0.02	0.07	2.75	1.58	7.47	8.76	42.99	−34.92	−8.40	0.77	7.33	7.75	13.58	15.00	41.45
		15	1	0.02	0.07	2.71	1.58	7.38	8.60	35.01	−34.77	−8.41	0.81	7.31	7.75	13.44	14.81	40.34
$\mathcal{A}^2_{\text{any}}$	d^+	0	0	0.00	0.02	0.68	0.22	6.23	8.30	10.45	−19.01	−9.96	0.01	6.57	7.44	7.93	8.11	8.96
		15	0	0.00	0.02	0.69	0.23	6.20	8.33	32.55	−42.61	−10.31	−0.02	6.56	7.44	7.93	8.12	15.56
		15	1	0.00	0.02	0.68	0.23	6.21	8.31	27.68	−35.84	−10.61	−0.07	6.54	7.44	7.93	8.11	9.57
$\mathcal{A}^3_{\text{any}}$	d^+	0	0	0.01	0.07	2.53	1.57	6.98	8.48	10.85	−11.35	−8.22	1.06	7.26	7.63	13.05	14.46	16.39
		15	0	0.01	0.06	2.27	0.93	7.15	8.55	34.03	−34.78	−8.30	0.82	7.20	7.62	12.98	14.60	26.26
		15	1	0.01	0.06	2.24	0.93	7.08	8.51	32.98	−34.78	−8.32	0.86	7.21	7.62	12.95	14.56	26.26

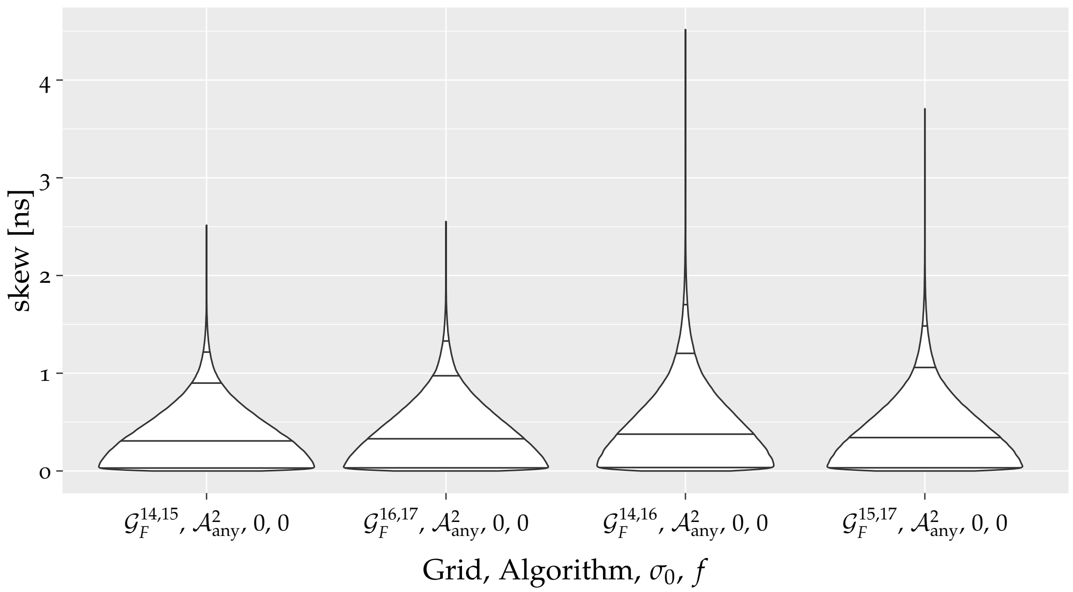
4.3.7 Results of the Pulse-based Algorithms in $\mathcal{G}_F$

Recall that for $\mathcal{G}_F$ we use $\mathcal{G}_H$ as the neighborhood-defining grid in the skew evaluation, so that only neighbors in the direct vicinity are included. To start this evaluation, we first consider the impact of the newly added parameter: The delay of long inter-layer links. Figures 4.17 and 4.18 show the behavior of the skew for different parameterizations of the long link delay. We see that changes to the delay of the long link either speed up the propagation, have no effect on it or hinders it. Hence, not only has the variance of the long link delay an effect on the skew, but also the relation to $2d^+$ respectively $2d^-$. Oddly enough, in high input skew sets, we see that the skews have a tendency to cluster for a long link delay uncertainty of 1, whereas for a uncertainty of 2, the inter-layer skews resemble a gyroscope. The introduction of faults into the grid, as shown in Figures 4.19 and 4.20 only has a minor impact on the distribution of the skews in general: However, the IQD^o is heavily increased.

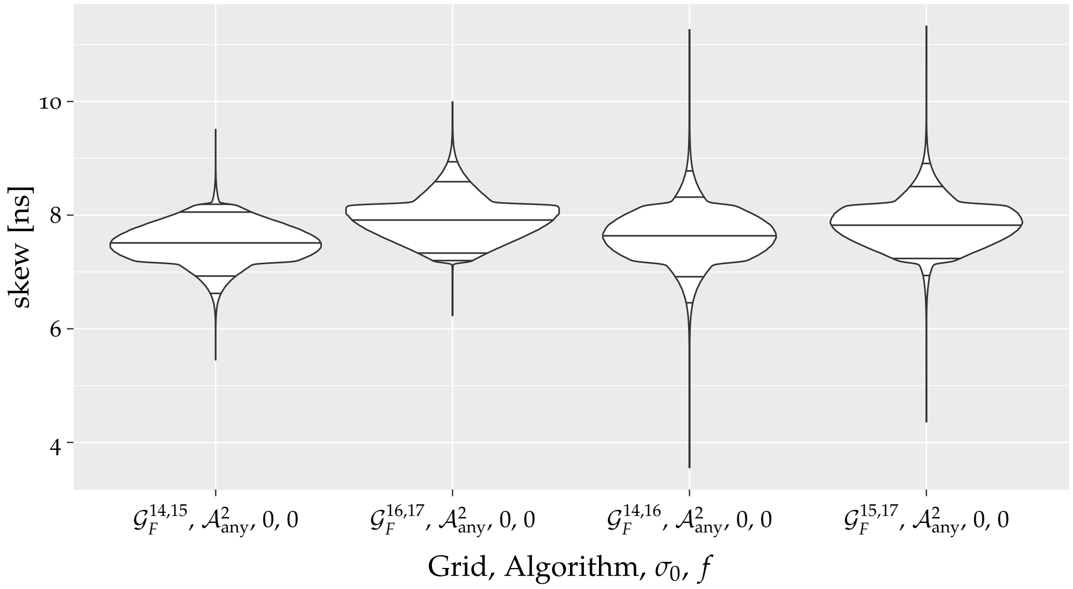
A comparison with $\mathcal{G}_A$ under $\mathcal{A}^2_{\text{any}}$ reveals that $\mathcal{G}_F$ performs, independently of the input skew set, only slightly worse in the fault-free case as shown in Figures 4.21 and 4.22. However, in the presence of faults, the IQD^o increases.

A complete overview of all conducted simulations in $\mathcal{G}_F$ can be found in Appendix A.11 on Page 299pp, with the corresponding tables in Appendix A.12 on Page 316pp. In conclusion, we argue against $\mathcal{G}_F$ due to the engineering effort required for the long inter-layer link which, from this investigation, does not seems to delivery any substantial benefit in the observed settings.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



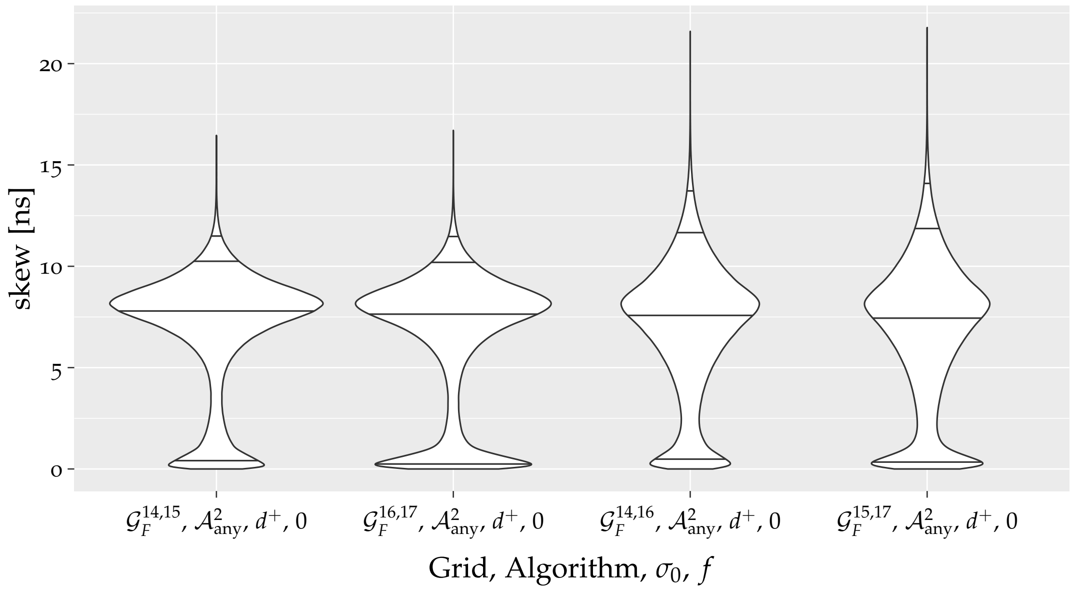
(a) intra-layer skews



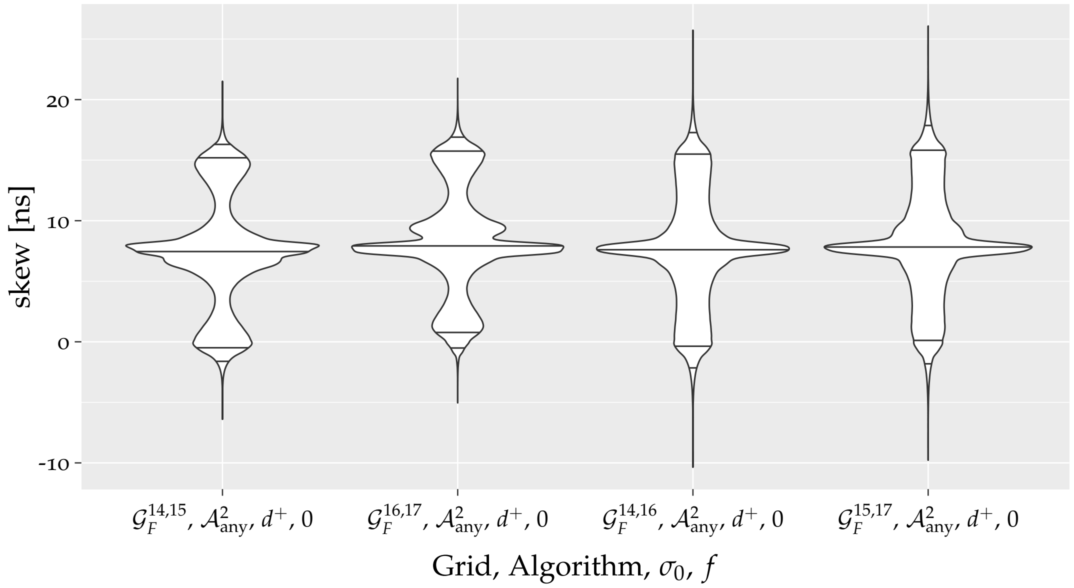
(b) inter-layer skews

Figure 4.17: Comparison of the effect of the long inter-link delay in $\mathcal{G}_F$ on the skews under $\sigma_0 = 0$ and no faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skews

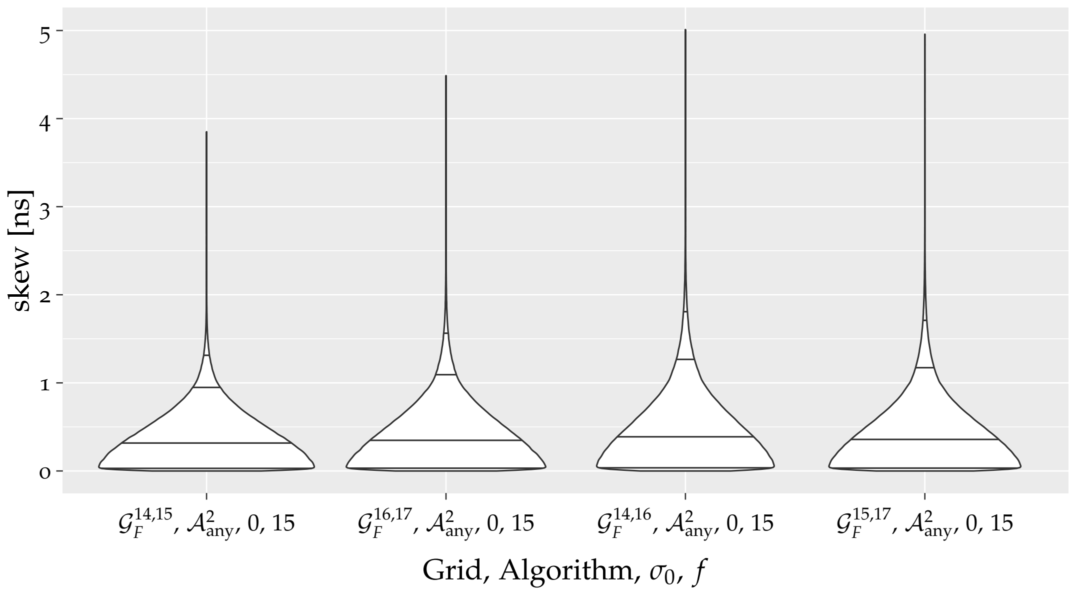


(b) inter-layer skews

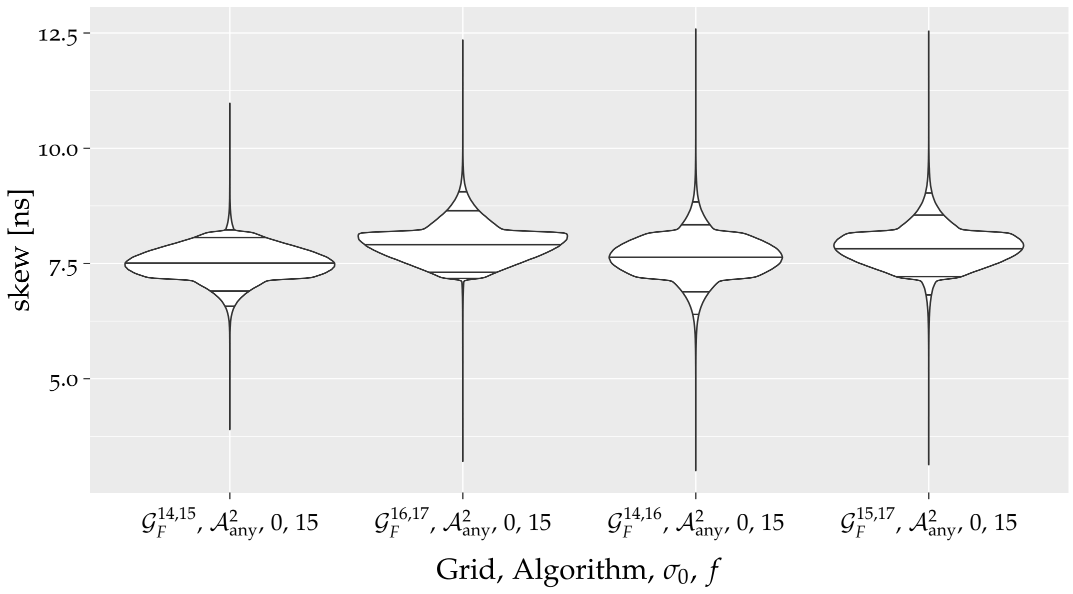
Figure 4.18: Comparison of the effect of the long inter-link delay in $\mathcal{G}_F$ on the skews under $\sigma_0 = d^+$ and no faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

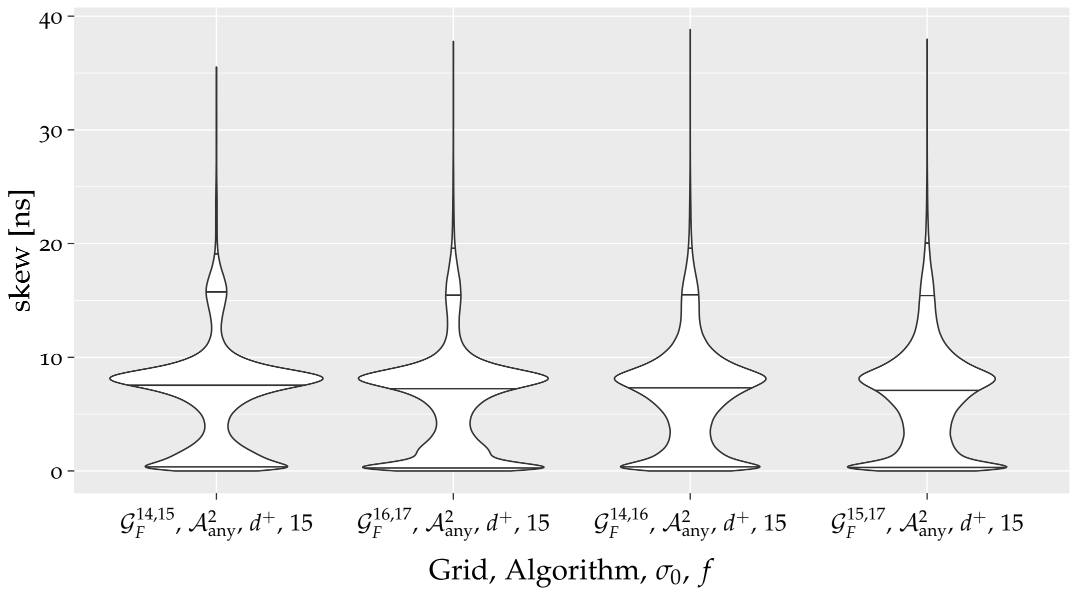


(a) intra-layer skews

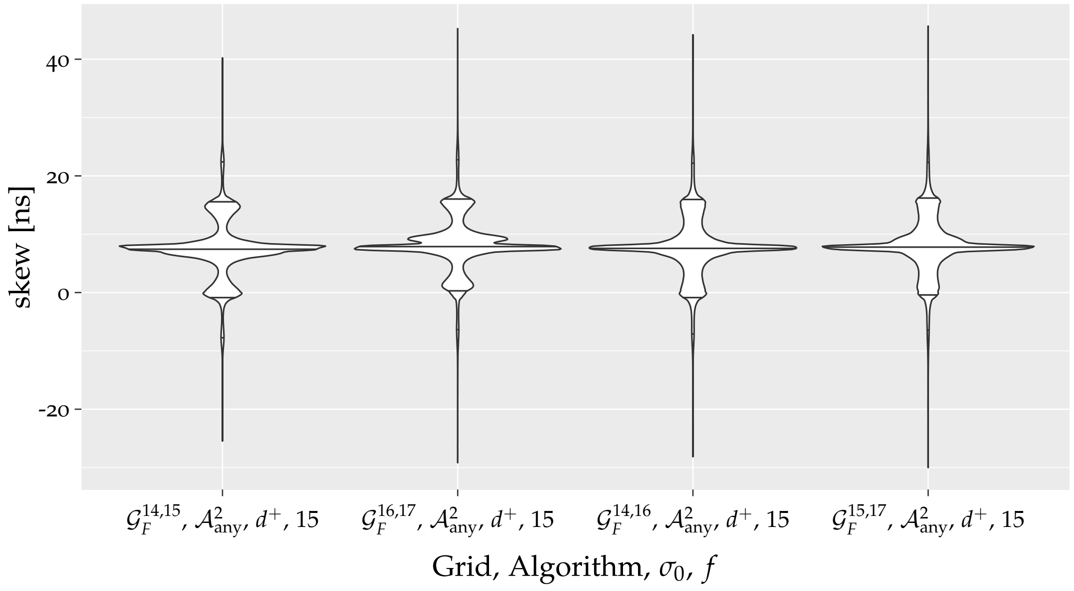


(b) inter-layer skews

Figure 4.19: Comparison of the effect of the long inter-link delay in $\mathcal{G}_F$ on the skews under $\sigma_0 = 0$ with 15 Byzantine faults.



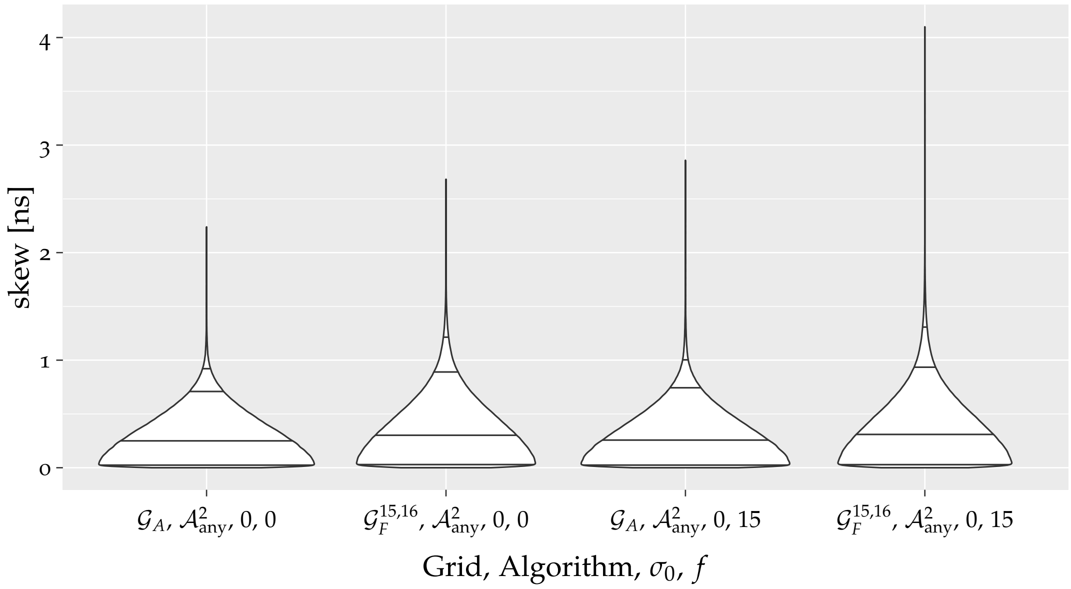
(a) intra-layer skews



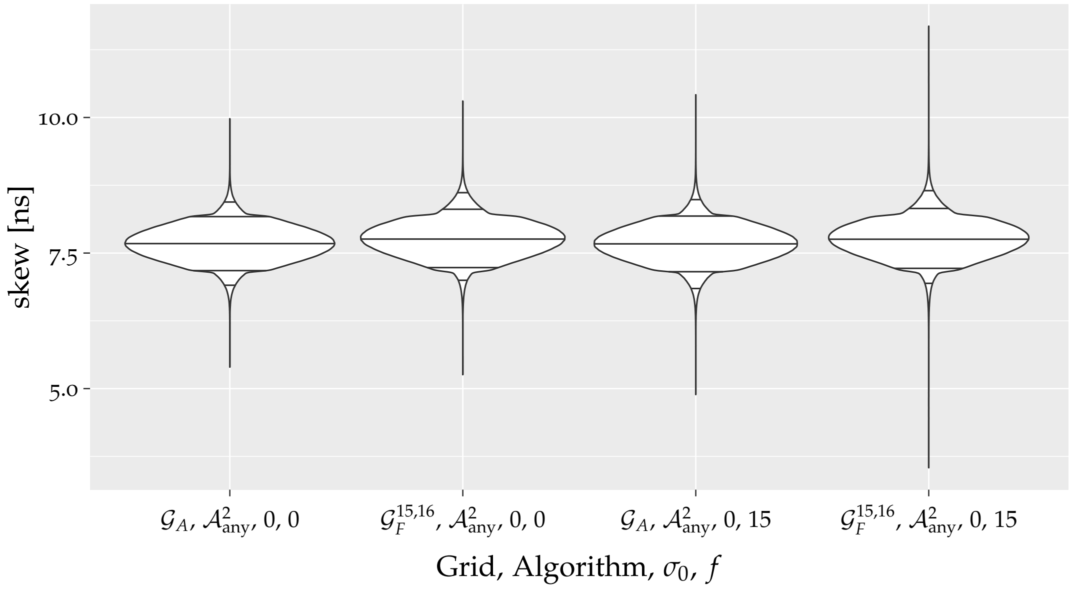
(b) inter-layer skews

Figure 4.20: Comparison of the effect of the long inter-link delay in $\mathcal{G}_F$ on the skews under $\sigma_0 = d^+$ with 15 Byzantine faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



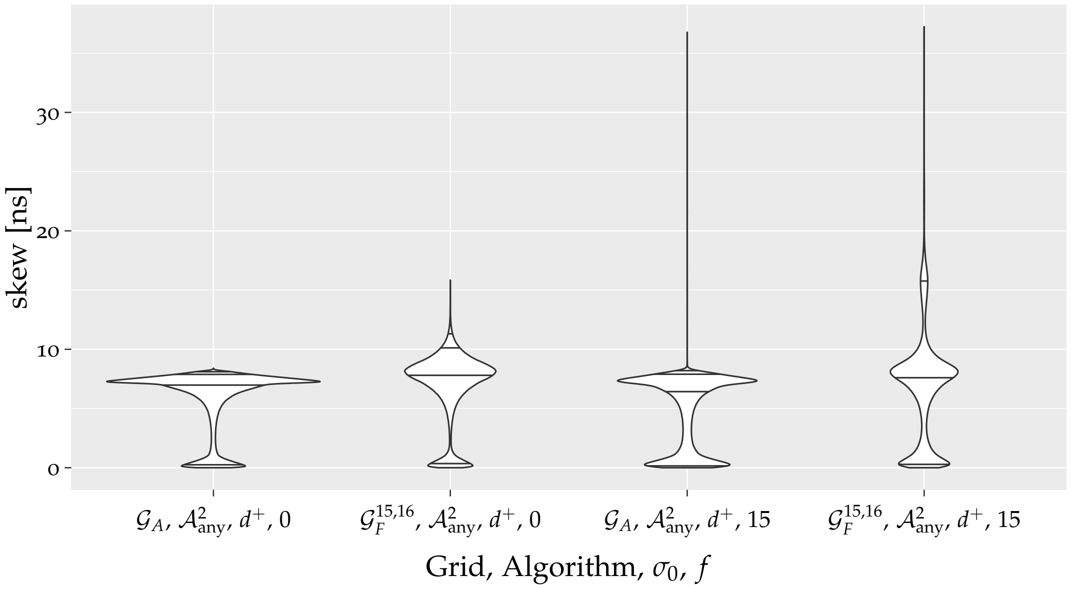
(a) intra-layer skews



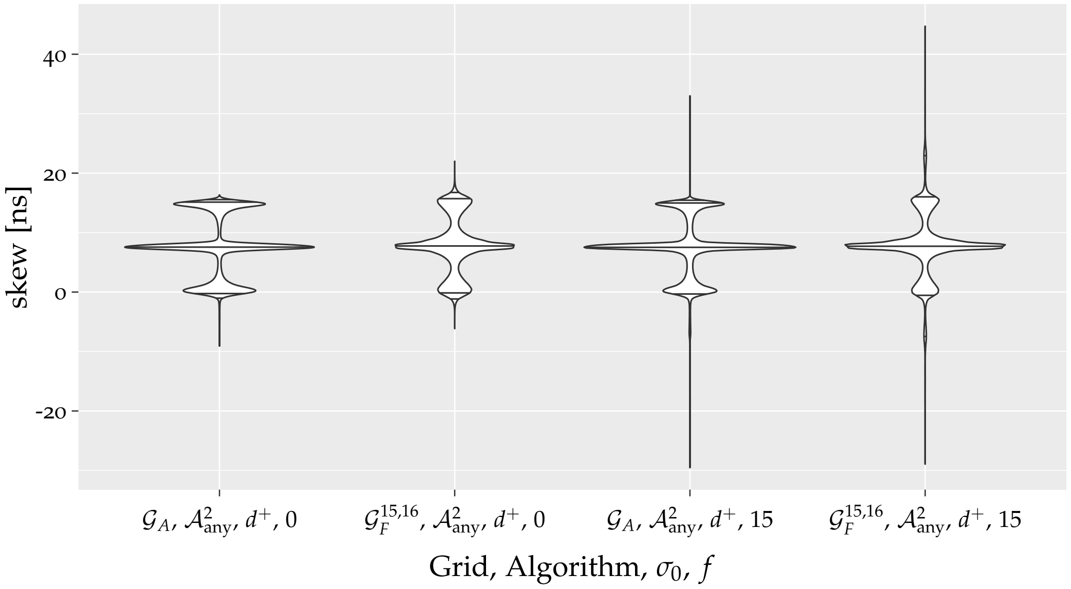
(b) inter-layer skews

Figure 4.21: Comparison of $\mathcal{G}_A$ with $\mathcal{G}_F^{15,16}$ under $\sigma_0 = 0$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) intra-layer skews



(b) inter-layer skews

Figure 4.22: Comparison of $\mathcal{G}_A$ with $\mathcal{G}_F^{15,16}$ under $\sigma_0 = d^+$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

4.3.8 Evaluation of the Time-based Algorithms

Following the pulse-based algorithms considered before, we now focus on the time-based algorithms. Compared to the pulse-based algorithms, they require a parameterization of their offset value, cf. Algorithms 4.3 and 4.4. To determine a suitable value for the offset, preliminary simulations were conducted which determined an offset of $o = 25\,\text{ns}$ to be a fitting value. The correctness of this offset has been verified by recording the maximum differences between calculated average respectively midpoint and the last received triggering relevant trigger message: This difference is a lower bound for the offset. The recorded differences for the fault-free case and the case with 15 faults are given in Table 4.13. Note that the bound is only violated for $\mathcal{G}_C$ with $d = 1$; this has already been discussed in Section 4.1.2 and is of no consequence for the other combinations.

The detailed violin plots of the time-based algorithms in $\mathcal{G}_C$ and $\mathcal{G}_D$ can be found in Appendices A.5 and A.6 on Page 261pp and in Appendix A.9 on Page 281pp. Note that we compensated the calculated inter-layer skews of the time-based algorithms in the figures and tables by the offset o to gain easily comparable results with respect to the other grids.

4.3.8.1 The Time-based Algorithms $\mathcal{A}^2_{\text{avg}}$ and $\mathcal{A}^2_{\text{mid}}$ in $\mathcal{G}_C$

Recall that the time-based algorithms $\mathcal{A}_{\text{avg}}$ and $\mathcal{A}_{\text{mid}}$ in $\mathcal{G}_C$ require $d = 2$ for correct operation. We see, by comparing the violin plots in Figure 4.23, that the time-based algorithms have a tighter distribution than the pulse-based algorithms in the low input skew set. This continues for the well-behaved input skew, shown in Figure 4.24, where both time-based algorithms show a similar control of the IQD^o even under faults.

Table 4.13: Recorded time differences in 1500 simulation runs between the arrival of the pulse that started the trigger time calculation and the computed trigger time before adding the offset o. The fault-type causing the maximum varied between the combinations.

		fault-free [ps]	15 faults [ps]	fault-type
$\mathcal{G}_C$	$\mathcal{A}^1_{\text{mid}}$	13 269	71 484	fail-silent
	$\mathcal{A}^1_{\text{avg}}$	13 517	104 180	fail-silent
$\mathcal{G}_C$	$\mathcal{A}^2_{\text{mid}}$	9178	20 778	Byzantine
	$\mathcal{A}^2_{\text{avg}}$	9241	21 131	Byzantine
$\mathcal{G}_D$	$\mathcal{A}^0_{\text{mid}}$	9154	19 794	Byzantine
	$\mathcal{A}^0_{\text{avg}}$	9247	22 705	Byzantine
$\mathcal{G}_D$	$\mathcal{A}^1_{\text{mid}}$	5164	16 739	Byzantine
	$\mathcal{A}^1_{\text{avg}}$	5164	16 739	Byzantine

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Furthermore, while the median of the skews is similar to $\mathcal{A}^3_{\text{any}}$, their IQD^1 and IQD^5 are closer in value to $\mathcal{A}^2_{\text{any}}$, cf. Table 4.14. For the high input skew sets, the similarity shifts to $\mathcal{A}^3_{\text{any}}$, however, as shown in Figures 4.25 and 4.26.

4.3.8.2 Time-based Algorithms in $\mathcal{G}_D$

Comparing Table 4.14 with Table 4.15 shows that the time-based algorithms $\mathcal{A}^0_{\text{avg}}$ and $\mathcal{A}^0_{\text{mid}}$ in $\mathcal{G}_D$ behave similar to the time-based algorithms with $d = 2$ in $\mathcal{G}_C$. We also see that in the presence of faults $\mathcal{G}_D$ shows only a slightly larger IQD^0. This is contrary to the behavior of the pulse-based algorithms, which degrade in the absence of intra-layer links in the high input skew set. We can attribute this behavior to the averaging methods, which limit the effect of outliers on a node's firing time, compared to the pulse-based algorithms that are dominated by the outliers. The advantage over $\mathcal{G}_C$ in this perspective lies also in the grid itself: Due to the higher number of neighbors, a node in $\mathcal{G}_C$, unlike in $\mathcal{G}_D$, has to wait for an intra-layer neighbor if one of their inter-layer neighbors is faulty.

Further, we compare the time-based algorithms $\mathcal{A}^1_{\text{avg}}$ and $\mathcal{A}^1_{\text{mid}}$ in $\mathcal{G}_D$ with $\mathcal{G}_C$ under $\mathcal{A}^2_{\text{any}}$ and $\mathcal{A}^3_{\text{any}}$. We see a slightly better behavior for a low input skew, cf. Figure 4.27 and well-behaved input skews. For high input skews, however, $\mathcal{G}_D$ shows the typical tendency of a grid without intra-layer links to degrade, see Figure 4.28. Note that the behavior of $\mathcal{G}_D$ with the time-based algorithms parameterized with $d = 0$ and $d = 1$ differ, cf. Appendix A.9 on Page 281pp, in the sense that the $d = 1$ parameterization has, in most cases, lower IQD's. As with the pulse-based algorithms, the removal of a neighbor from the decision process allows better behavior in well-behaved settings, while causing degradation with higher input skew sets.

4.3.8.3 The Time-based Algorithms $\mathcal{A}^1_{\text{avg}}$ and $\mathcal{A}^1_{\text{mid}}$ in $\mathcal{G}_C$

Despite the issue with the offset determination, as elaborated in Section 4.1.2, we have simulated the behavior of the algorithms $\mathcal{A}^1_{\text{avg}}$ and $\mathcal{A}^1_{\text{mid}}$ in $\mathcal{G}_C$. Recall that in this combination acausal triggering times are calculated in certain cases by the time-based algorithms, hence this setting is excluded from general consideration due to the unknown effect on the performance in corner cases. On inspection of the plots given in Appendix A.6 on Page 266pp, we see for low input skews in the fault-free case a resemblance with the time-based algorithms $\mathcal{A}^1_{\text{avg}}$ and $\mathcal{A}^1_{\text{mid}}$ in $\mathcal{G}_D$. Yet, with high input skews or faults, this algorithm degrades heavily.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

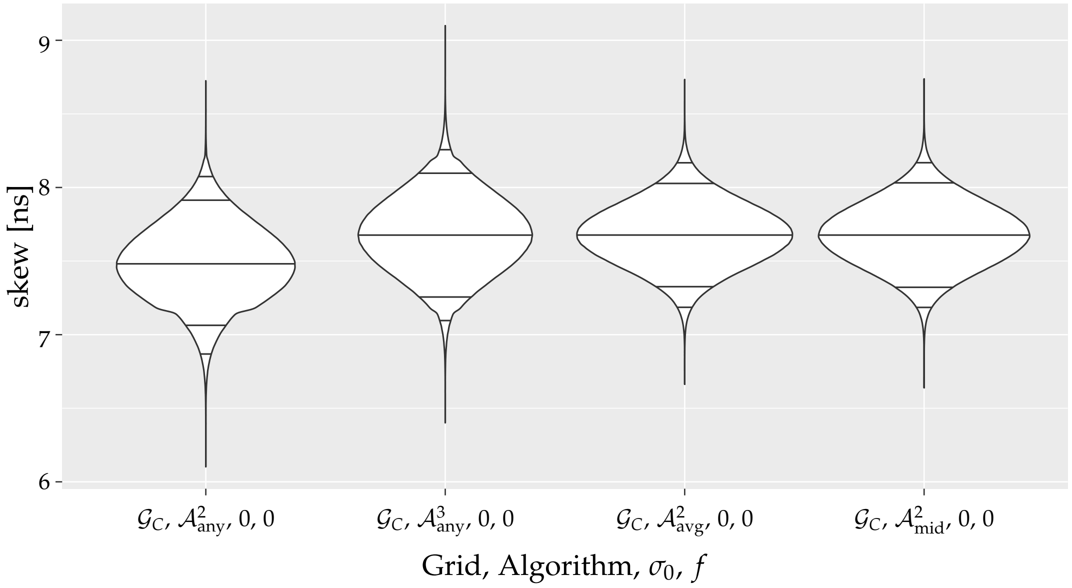


Figure 4.23: Comparison of the inter-layer skews of the pulse-based algorithms and the time-based algorithms, with $d = 2$, in $\mathcal{G}_C$ with $\sigma_0 = 0$ and no faults.

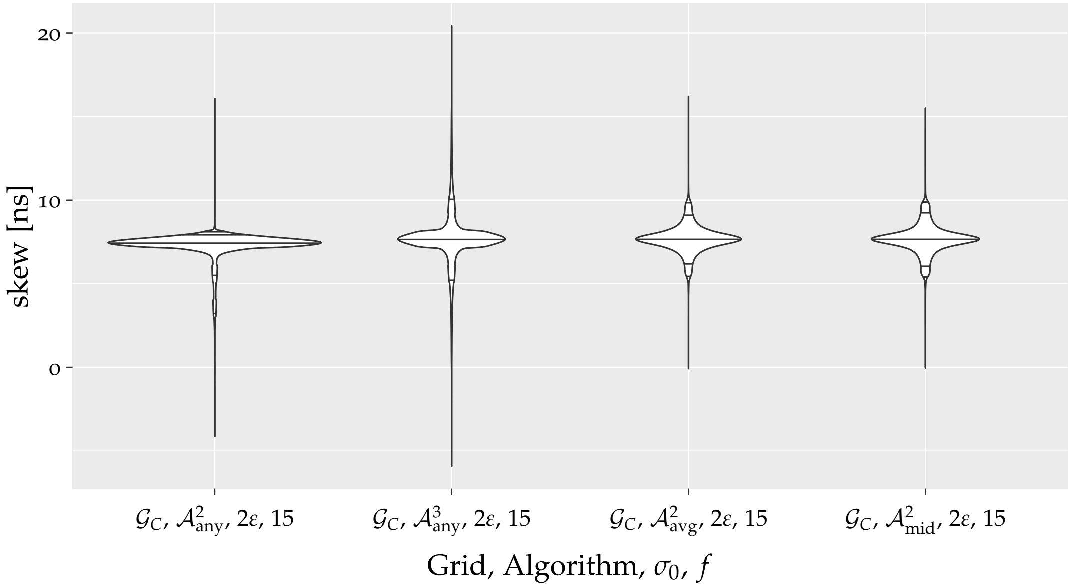


Figure 4.24: Comparison of the inter-layer skews of the pulse-based algorithms and time-based algorithms, with $d = 2$, in $\mathcal{G}_C$ with $\sigma_0 = 2\varepsilon$ and 15 Byzantine faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

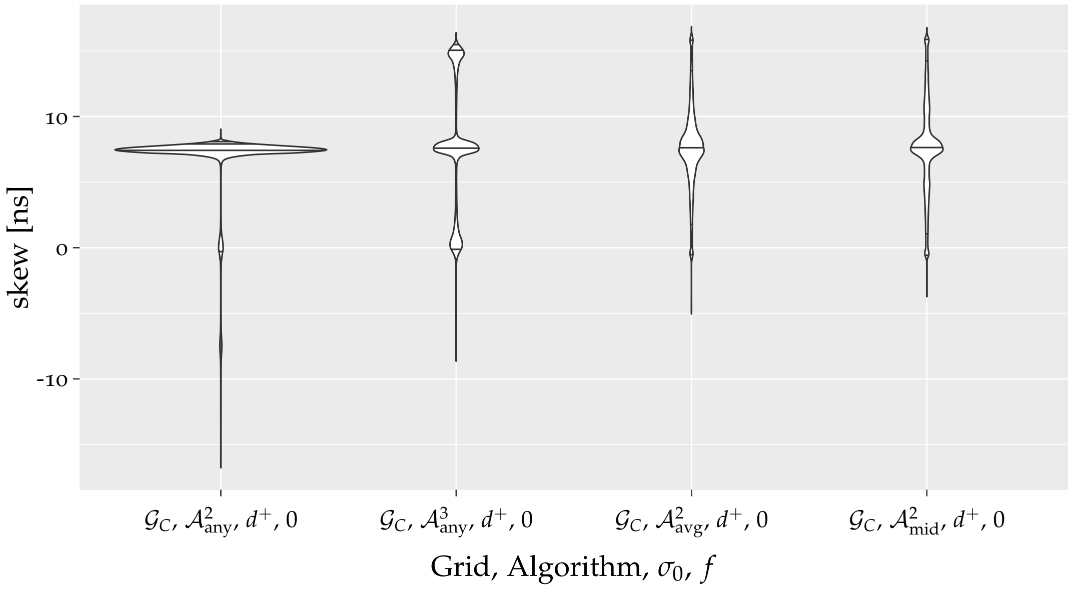


Figure 4.25: Comparison of the inter-layer skews of the pulse-based algorithms and time-based algorithms, with $d = 2$, in $\mathcal{G}_C$ with $\sigma_0 = d^+$ and no faults.

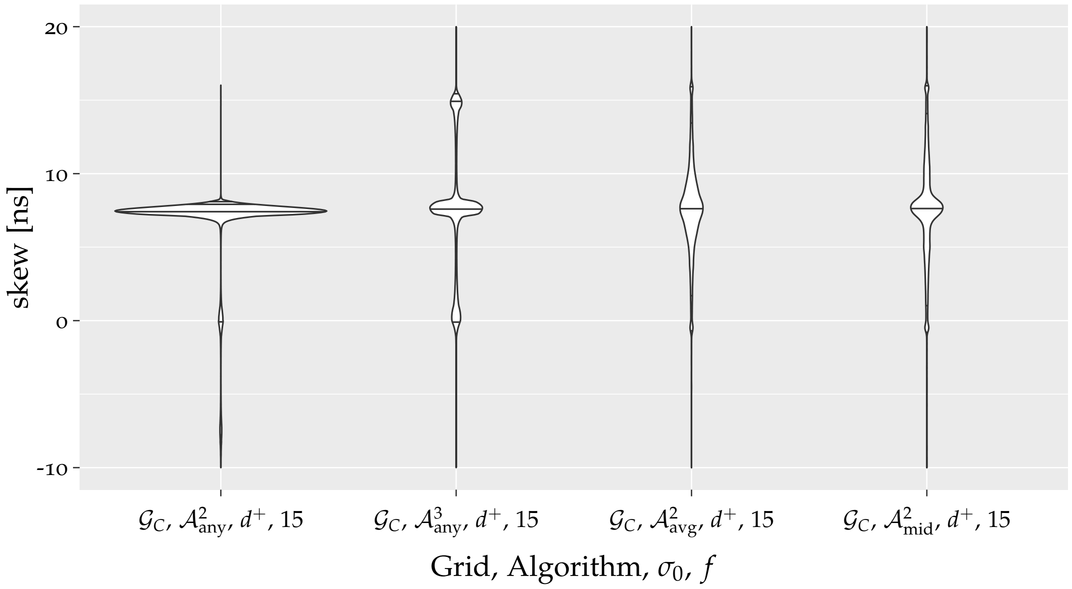


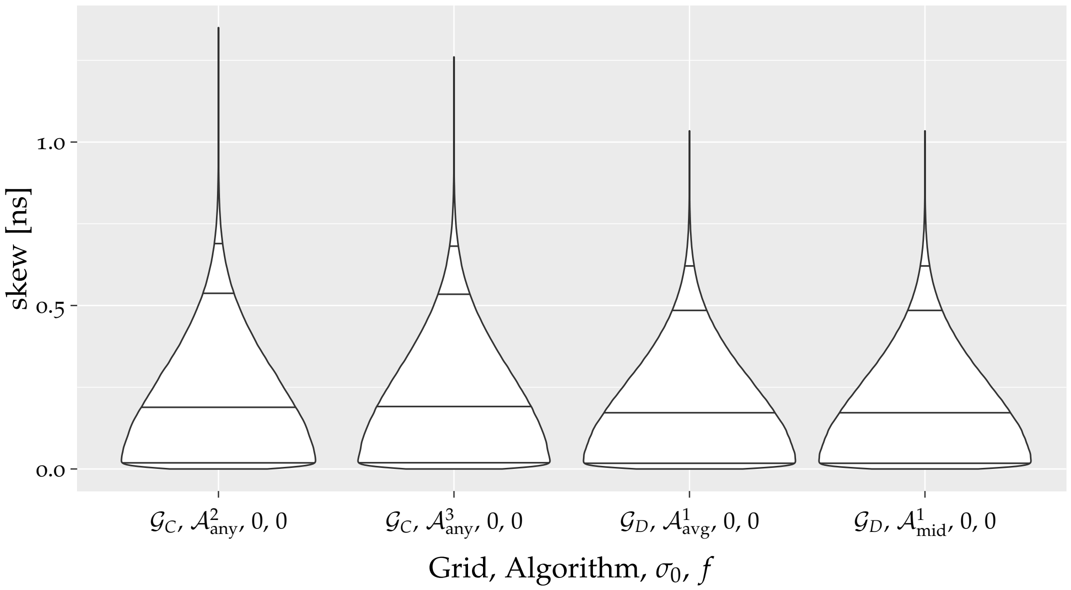
Figure 4.26: Comparison of pulse-based and time-based algorithms, with $d = 2$, in $\mathcal{G}_C$ with $\sigma_0 = d^+$ and 15 Byzantine faults. The plot has been limited to be within $[-10, 20]$ ns to improve the visibility of the majority of the skews.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

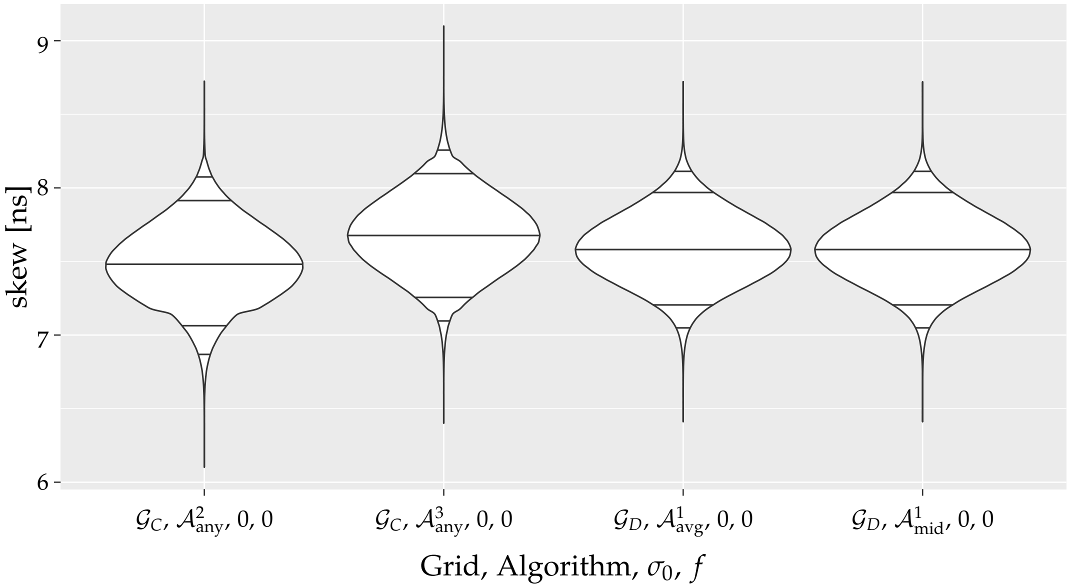
TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.14: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\mathrm{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology with time-based algorithms parameterized with $d = 2$. The inter-layer skews have been compensated by 25 ns.

				intra-layer							inter-layer						
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^2_{\mathrm{mid},\,25\,\mathrm{ns}}$	0	0	0	0.00	0.02	0.18	0.16	0.45	0.57	1.02	6.66	7.19	7.33	7.68	7.68	8.03	8.17	8.73
		15	0	0.00	0.02	0.19	0.16	0.46	0.60	1.20	6.54	7.17	7.32	7.68	7.68	8.04	8.18	8.92
		15	1	0.00	0.02	0.18	0.16	0.45	0.58	0.97	6.65	7.18	7.32	7.68	7.68	8.03	8.17	8.76
$\mathcal{A}^2_{\mathrm{avg},\,25\,\mathrm{ns}}$	0	0	0	0.00	0.02	0.19	0.16	0.45	0.58	1.01	6.64	7.19	7.32	7.68	7.68	8.03	8.17	8.74
		15	0	0.00	0.02	0.19	0.16	0.46	0.60	1.12	6.48	7.17	7.31	7.68	7.68	8.04	8.18	8.84
		15	1	0.00	0.02	0.19	0.16	0.45	0.58	0.99	6.63	7.18	7.32	7.68	7.68	8.04	8.17	8.72
$\mathcal{A}^2_{\mathrm{mid},\,25\,\mathrm{ns}}$	ε	0	0	0.01	0.02	0.36	0.27	1.04	1.43	2.43	5.58	6.53	6.96	7.68	7.68	8.40	8.82	9.66
		15	0	0.01	0.02	0.36	0.27	1.06	1.51	3.80	3.55	6.48	6.94	7.67	7.68	8.41	8.85	12.13
		15	1	0.01	0.02	0.36	0.26	1.04	1.45	2.78	4.77	6.51	6.96	7.68	7.68	8.40	8.83	10.25
$\mathcal{A}^2_{\mathrm{avg},\,25\,\mathrm{ns}}$	ε	0	0	0.01	0.03	0.38	0.29	1.04	1.33	2.03	5.71	6.52	6.92	7.68	7.68	8.44	8.83	9.64
		15	0	0.01	0.03	0.38	0.29	1.07	1.41	3.86	3.56	6.48	6.91	7.67	7.68	8.44	8.86	11.68
		15	1	0.01	0.03	0.37	0.28	1.04	1.34	2.40	4.66	6.51	6.92	7.68	7.68	8.43	8.83	10.09
$\mathcal{A}^2_{\mathrm{mid},\,25\,\mathrm{ns}}$	2ε	0	0	0.01	0.04	0.68	0.47	2.03	2.52	3.81	4.09	5.53	6.24	7.68	7.68	9.11	9.81	10.73
		15	0	0.01	0.04	0.68	0.45	2.09	2.74	6.90	−0.08	5.46	6.21	7.67	7.67	9.12	9.87	16.21
		15	1	0.01	0.04	0.66	0.44	2.05	2.61	5.78	1.94	5.51	6.24	7.67	7.68	9.09	9.83	12.52
$\mathcal{A}^2_{\mathrm{avg},\,25\,\mathrm{ns}}$	2ε	0	0	0.01	0.05	0.77	0.58	2.03	2.35	3.07	4.43	5.51	6.07	7.68	7.68	9.28	9.84	10.66
		15	0	0.01	0.05	0.76	0.55	2.07	2.54	7.02	−0.03	5.42	6.06	7.67	7.67	9.26	9.90	15.50
		15	1	0.01	0.05	0.74	0.55	2.04	2.40	4.19	1.99	5.48	6.09	7.67	7.68	9.25	9.86	12.38
$\mathcal{A}^2_{\mathrm{mid},\,25\,\mathrm{ns}}$	d^+	0	0	0.03	0.13	2.85	2.24	7.75	8.49	9.95	−5.05	−0.50	1.80	7.66	7.64	13.52	15.84	16.84
		15	0	0.02	0.10	2.77	1.92	8.12	10.48	25.72	−23.64	−0.66	1.73	7.64	7.65	13.51	15.96	38.42
		15	1	0.02	0.11	2.73	1.92	8.04	9.46	20.65	−14.75	−0.58	1.85	7.65	7.66	13.42	15.89	27.38
$\mathcal{A}^2_{\mathrm{avg},\,25\,\mathrm{ns}}$	d^+	0	0	0.02	0.12	3.37	3.04	7.96	8.41	9.14	−3.73	−0.55	1.08	7.67	7.65	14.25	15.89	16.77
		15	0	0.02	0.11	3.27	2.78	8.11	9.69	27.19	−25.24	−0.74	1.05	7.64	7.66	14.19	16.04	39.09
		15	1	0.02	0.12	3.23	2.79	8.03	8.63	15.40	−13.60	−0.63	1.19	7.66	7.67	14.09	15.96	25.09



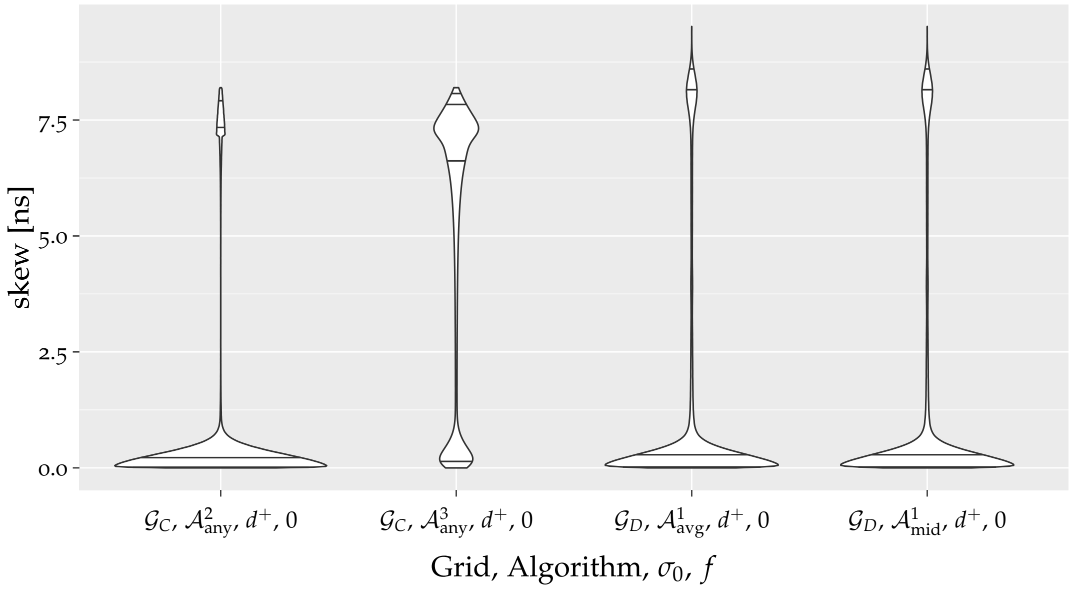
(a) intra-layer skews



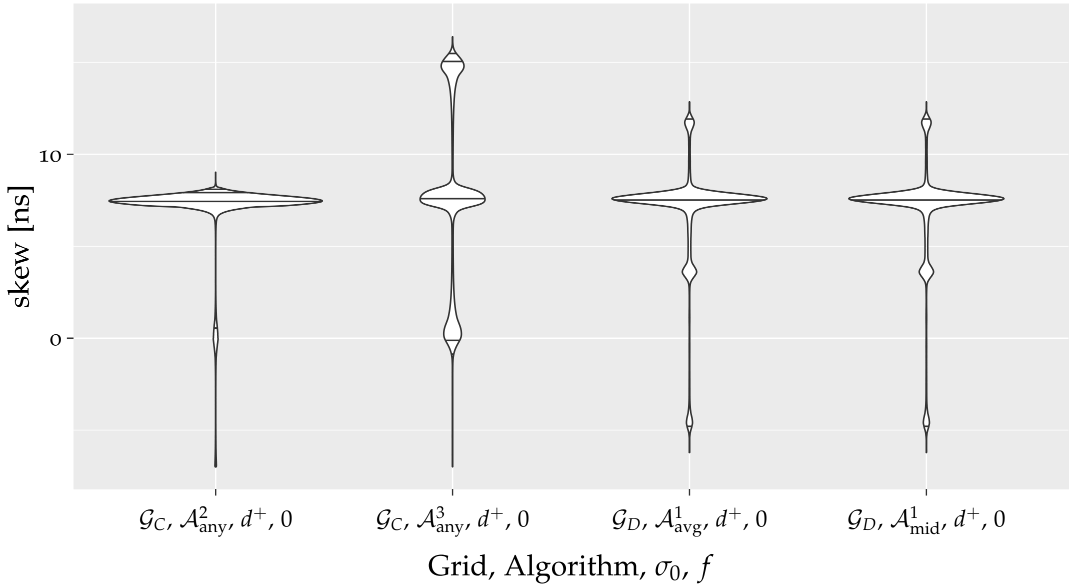
(b) inter-layer skews

Figure 4.27: Comparison between the pulse-based algorithms in $\mathcal{G}_C$ and the time-based algorithms with $d = 1$ in $\mathcal{G}_D$ with $\sigma_0 = 0$ and no faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



(a) intra-layer skews



(b) inter-layer skews

Figure 4.28: Comparison between the pulse-based algorithms in $\mathcal{G}_C$ and the time-based algorithms with $d = 1$ in $\mathcal{G}_D$ with $\sigma_0 = d^+$ and no faults. The plot (b) has been cropped at -7 ns to improve the visibility of the majority of the skews.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.15: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\mathrm{op}}$ (in [ns]) in the $\mathcal{G}_D$ topology with time-based algorithms parameterized with $d = 0$. The inter-layer skews have been compensated by 25 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^0_{\mathrm{mid},25\,\mathrm{ns}}$	0	0	0	0.00	0.02	0.18	0.16	0.45	0.57	1.06	6.60	7.19	7.33	7.68	7.68	8.03	8.17	8.70
		15	0	0.00	0.02	0.19	0.16	0.46	0.59	1.09	6.59	7.18	7.32	7.68	7.68	8.04	8.18	8.74
		15	1	0.00	0.02	0.18	0.16	0.45	0.58	1.08	6.59	7.18	7.33	7.68	7.68	8.03	8.17	8.72
$\mathcal{A}^0_{\mathrm{avg},25\,\mathrm{ns}}$	0	0	0	0.00	0.02	0.19	0.16	0.45	0.58	1.03	6.64	7.19	7.32	7.68	7.68	8.03	8.17	8.70
		15	0	0.00	0.02	0.19	0.16	0.46	0.59	1.13	6.65	7.18	7.32	7.68	7.68	8.04	8.18	8.74
		15	1	0.00	0.02	0.19	0.16	0.45	0.58	0.96	6.65	7.18	7.32	7.68	7.68	8.04	8.17	8.69
$\mathcal{A}^0_{\mathrm{mid},25\,\mathrm{ns}}$	ε	0	0	0.01	0.02	0.36	0.27	1.04	1.43	2.45	5.53	6.53	6.96	7.68	7.68	8.40	8.82	9.68
		15	0	0.01	0.02	0.36	0.27	1.06	1.49	3.53	3.99	6.50	6.95	7.68	7.68	8.41	8.84	11.63
		15	1	0.01	0.02	0.36	0.26	1.04	1.44	2.93	4.98	6.52	6.96	7.68	7.68	8.40	8.83	10.40
$\mathcal{A}^0_{\mathrm{avg},25\,\mathrm{ns}}$	ε	0	0	0.01	0.03	0.38	0.29	1.04	1.33	1.99	5.74	6.52	6.92	7.68	7.68	8.44	8.83	9.64
		15	0	0.01	0.03	0.38	0.29	1.06	1.38	3.72	3.29	6.49	6.92	7.68	7.68	8.44	8.85	11.90
		15	1	0.01	0.03	0.37	0.28	1.04	1.34	2.24	4.65	6.51	6.92	7.68	7.68	8.43	8.83	10.49
$\mathcal{A}^0_{\mathrm{mid},25\,\mathrm{ns}}$	2ε	0	0	0.01	0.04	0.68	0.47	2.03	2.52	3.85	4.10	5.53	6.24	7.68	7.68	9.11	9.81	10.75
		15	0	0.01	0.04	0.67	0.45	2.08	2.69	6.13	0.24	5.49	6.23	7.68	7.68	9.12	9.85	14.76
		15	1	0.01	0.04	0.67	0.45	2.05	2.59	5.27	3.15	5.52	6.25	7.68	7.68	9.10	9.82	12.16
$\mathcal{A}^0_{\mathrm{avg},25\,\mathrm{ns}}$	2ε	0	0	0.01	0.05	0.77	0.58	2.03	2.35	3.03	4.44	5.51	6.07	7.68	7.68	9.28	9.84	10.67
		15	0	0.01	0.05	0.76	0.56	2.06	2.48	7.48	−0.51	5.45	6.08	7.68	7.68	9.27	9.89	15.30
		15	1	0.01	0.05	0.75	0.55	2.03	2.39	3.81	2.15	5.49	6.09	7.68	7.68	9.26	9.86	12.42
$\mathcal{A}^0_{\mathrm{mid},25\,\mathrm{ns}}$	d^+	0	0	0.02	0.13	2.85	2.25	7.75	8.49	9.90	−5.08	−0.50	1.80	7.66	7.64	13.52	15.84	16.86
		15	0	0.02	0.11	2.78	1.96	8.08	10.04	23.93	−19.28	−0.61	1.80	7.66	7.66	13.52	15.95	35.12
		15	1	0.02	0.11	2.75	1.95	8.02	9.28	20.74	−9.54	−0.56	1.86	7.66	7.66	13.46	15.89	24.12
$\mathcal{A}^0_{\mathrm{avg},25\,\mathrm{ns}}$	d^+	0	0	0.02	0.12	3.37	3.04	7.96	8.41	9.25	−3.75	−0.54	1.08	7.67	7.65	14.25	15.89	16.76
		15	0	0.02	0.11	3.29	2.83	8.08	9.13	28.00	−22.29	−0.68	1.10	7.67	7.67	14.22	16.01	36.46
		15	1	0.02	0.12	3.25	2.83	8.02	8.59	14.68	−11.05	−0.61	1.18	7.67	7.67	14.14	15.95	26.22

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 4.16: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_D$ topology with time-based algorithms parameterized with $d = 1$. The inter-layer skews have been compensated by 25 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^1_{\text{mid},25\,\text{ns}}$	0	0	0	0.00	0.02	0.20	0.17	0.49	0.62	1.03	6.41	7.05	7.21	7.58	7.58	7.97	8.11	8.72
		15	0	0.00	0.02	0.20	0.17	0.50	0.64	1.30	6.43	7.04	7.20	7.58	7.58	7.97	8.12	8.78
		15	1	0.00	0.02	0.20	0.17	0.49	0.63	1.07	6.45	7.05	7.20	7.58	7.58	7.97	8.12	8.68
$\mathcal{A}^1_{\text{avg},25\,\text{ns}}$	0	0	0	0.00	0.02	0.20	0.17	0.49	0.62	1.03	6.41	7.05	7.21	7.58	7.58	7.97	8.11	8.72
		15	0	0.00	0.02	0.20	0.17	0.50	0.64	1.30	6.43	7.04	7.20	7.58	7.58	7.97	8.12	8.78
		15	1	0.00	0.02	0.20	0.17	0.49	0.63	1.07	6.45	7.05	7.20	7.58	7.58	7.97	8.12	8.68
$\mathcal{A}^1_{\text{mid},25\,\text{ns}}$	ε	0	0	0.00	0.02	0.33	0.23	1.06	1.45	2.39	4.45	5.94	6.73	7.51	7.56	8.09	8.39	9.16
		15	0	0.00	0.02	0.33	0.23	1.06	1.50	5.03	2.27	5.92	6.73	7.51	7.56	8.09	8.41	10.98
		15	1	0.00	0.02	0.33	0.23	1.06	1.47	2.89	3.38	5.93	6.73	7.51	7.56	8.09	8.40	9.77
$\mathcal{A}^1_{\text{avg},25\,\text{ns}}$	ε	0	0	0.00	0.02	0.33	0.23	1.06	1.45	2.39	4.45	5.94	6.73	7.51	7.56	8.09	8.39	9.16
		15	0	0.00	0.02	0.33	0.23	1.06	1.50	5.03	2.27	5.92	6.73	7.51	7.56	8.09	8.41	10.98
		15	1	0.00	0.02	0.33	0.23	1.06	1.47	2.89	3.38	5.93	6.73	7.51	7.56	8.09	8.40	9.77
$\mathcal{A}^1_{\text{mid},25\,\text{ns}}$	2ε	0	0	0.00	0.02	0.50	0.25	2.08	2.50	3.53	2.79	4.38	5.76	7.41	7.55	8.34	8.90	9.80
		15	0	0.00	0.02	0.51	0.25	2.09	2.64	9.02	−2.01	4.33	5.77	7.41	7.54	8.32	8.93	14.15
		15	1	0.00	0.02	0.50	0.25	2.08	2.57	5.14	−0.64	4.36	5.79	7.41	7.55	8.33	8.91	10.96
$\mathcal{A}^1_{\text{avg},25\,\text{ns}}$	2ε	0	0	0.00	0.02	0.50	0.25	2.08	2.50	3.53	2.79	4.38	5.76	7.41	7.55	8.34	8.90	9.80
		15	0	0.00	0.02	0.51	0.25	2.09	2.64	9.02	−2.01	4.33	5.77	7.41	7.54	8.32	8.93	14.15
		15	1	0.00	0.02	0.50	0.25	2.08	2.57	5.14	−0.64	4.36	5.79	7.41	7.55	8.33	8.91	10.96
$\mathcal{A}^1_{\text{mid},25\,\text{ns}}$	d^+	0	0	0.01	0.03	1.57	0.29	8.16	8.60	9.51	−6.22	−4.78	−0.92	6.78	7.53	10.94	11.93	12.85
		15	0	0.01	0.03	1.57	0.29	8.19	9.30	33.38	−30.29	−4.88	−0.43	6.78	7.53	10.52	12.01	33.22
		15	1	0.01	0.03	1.55	0.29	8.18	8.91	21.34	−26.18	−4.84	−0.41	6.78	7.53	10.56	11.98	20.67
$\mathcal{A}^1_{\text{avg},25\,\text{ns}}$	d^+	0	0	0.01	0.03	1.57	0.29	8.16	8.60	9.51	−6.22	−4.78	−0.92	6.78	7.53	10.94	11.93	12.85
		15	0	0.01	0.03	1.57	0.29	8.19	9.30	33.38	−30.29	−4.88	−0.43	6.78	7.53	10.52	12.01	33.22
		15	1	0.01	0.03	1.55	0.29	8.18	8.91	21.34	−26.18	−4.84	−0.41	6.78	7.53	10.56	11.98	20.67

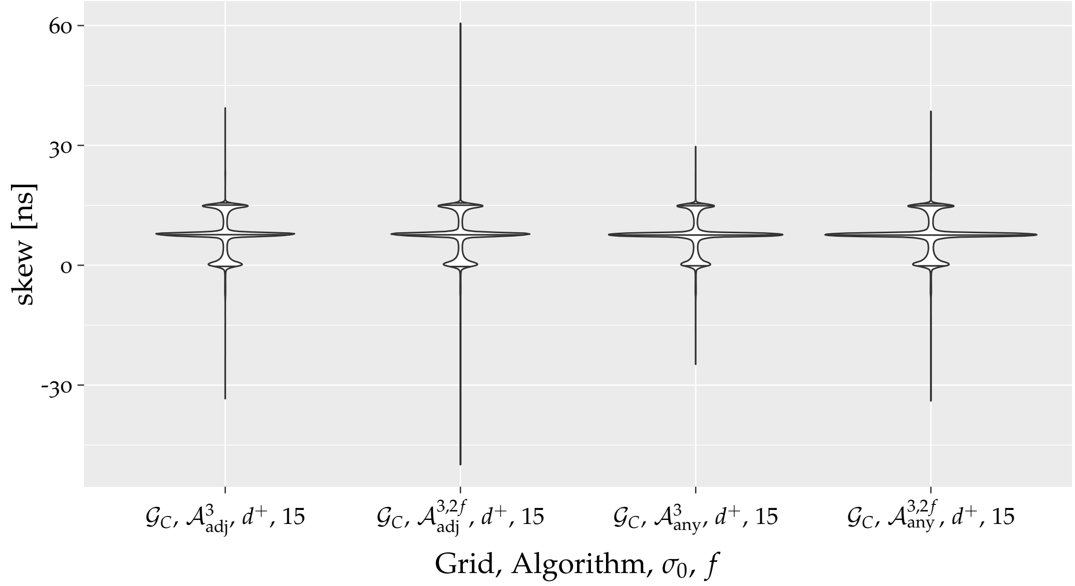


Figure 4.29: Comparison of the inter-layer skews of the pulse-based algorithms in the 1–local fault and 2–local fault model with an input skew of $\sigma_0 = d^+$ and 15 Byzantine faults.

4.3.9 Algorithms under the 2–local fault Model

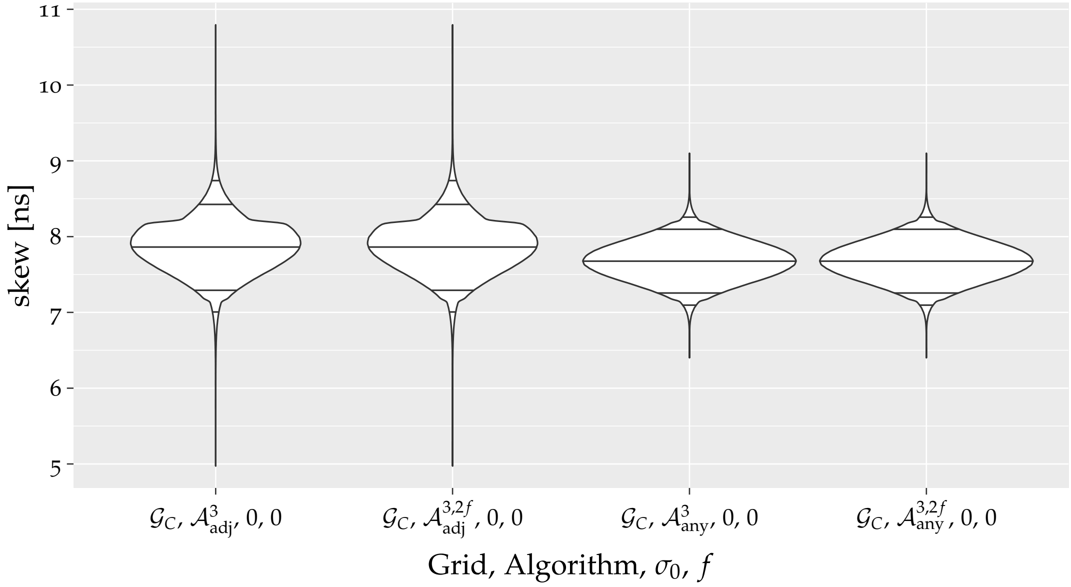
The increased number of links in $\mathcal{G}_C$ and $\mathcal{G}_D$ allows a suitable algorithm to withstand two faults on its incoming edges, as opposed to the single fault handleable by, e.g., $\mathcal{G}_A$. We, hence, also evaluate the effect that such a failure model has on the skews in a grid. Note that, we compare algorithms using the 2–local fault model with algorithms using the 1–local fault model, however.

First, note that the algorithms in $\mathcal{G}_C$ require $d \geq 1$ under 2–local faults. Consider two directly adjacent faults in a setting with $d = 0$: Each node requires 5 out of its 7 neighbors to send a pulse to trigger. The two nodes directly above these faults would hence require a pulse from each other to trigger. Hence, a cyclic dependency arises which cannot be resolved.

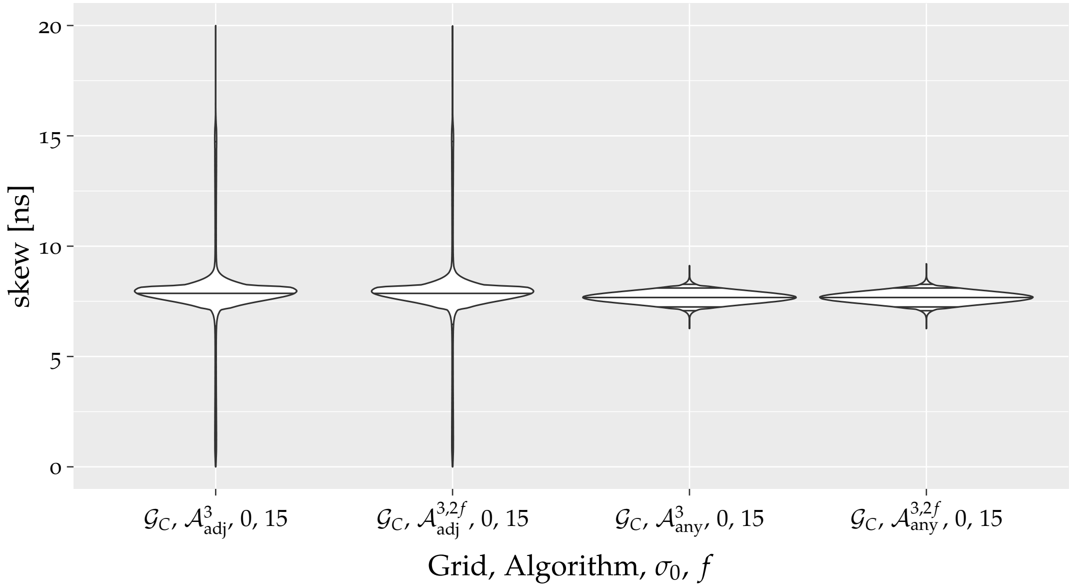
We now consider the effect of the 2–local fault model on $\mathcal{G}_C$, with the offset $o = 25\,\text{ns}$ as used in the 1–local fault model. Due to the additional resilience, we cannot expect the algorithm to behave identical, even in low input skew sets. Nonetheless, as can be seen in Figures 4.30a and 4.30b, the $f = 2$ algorithm shows similar skews as the $f = 1$ versions in the low skew set. However, in higher input skew sets, especially in the presence of faults, as shown in Figure 4.29, the $f = 2$ algorithms show a worse control of their IQD°. This is rather unsurprising due to the adverse fault-model applied.

As the recorded values suggested a rather tight margin between the minimal value required for the offset and the chosen value, we conducted simulations with an increased offset value. In Appendix A.14, we provided the results for $\mathcal{G}_C$ with $\mathcal{A}^{1,2f}_{\text{avg},\,80\,\text{ns}}$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



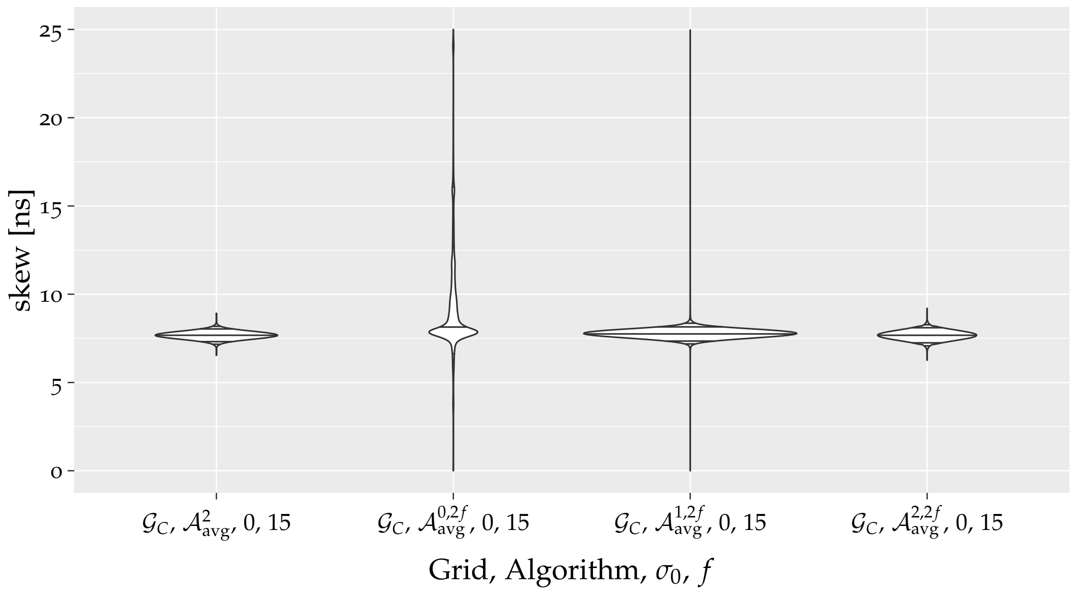
(a) fault-free



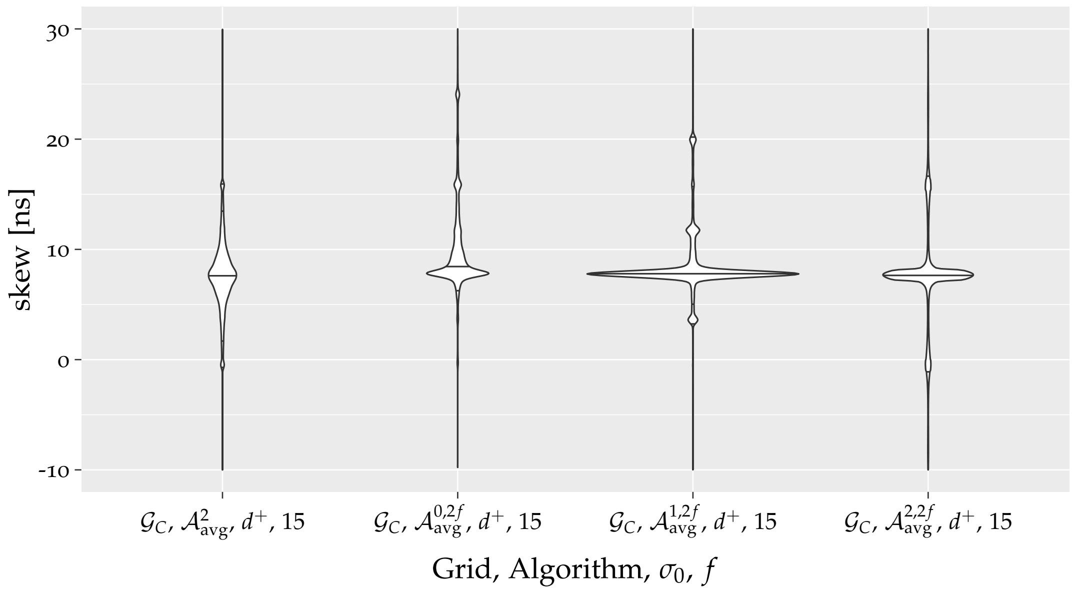
(b) 15 Byzantine faults

Figure 4.30: Comparison of the pulse-based algorithms in the 1–local fault and 2–local fault model in $\mathcal{G}_C$ with an input skew of $\sigma_0 = 0$ and (a) no faults faults, respectively (b) 15 Byzantine faults. The plot (b) has been limited to be within $[0, 20]$ ns to improve the visibility of the majority of the skews.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub



(a) σ_0= 0 input skew



(b) σ_0= d^+ input skew

Figure 4.31: Comparison of the time-based algorithms in the 1–local fault and 2–local fault model in $\mathcal{G}_C$ with an input skew of (a) $\sigma_0 = 0$ respectively (b) $\sigma_0 = d^+$, and 15 Byzantine faults. The plot have been limited to improve the visibility of the majority of the skews: (a) has been cropped at 25 ns, and (b) has been limited to be within $[-10, 30]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

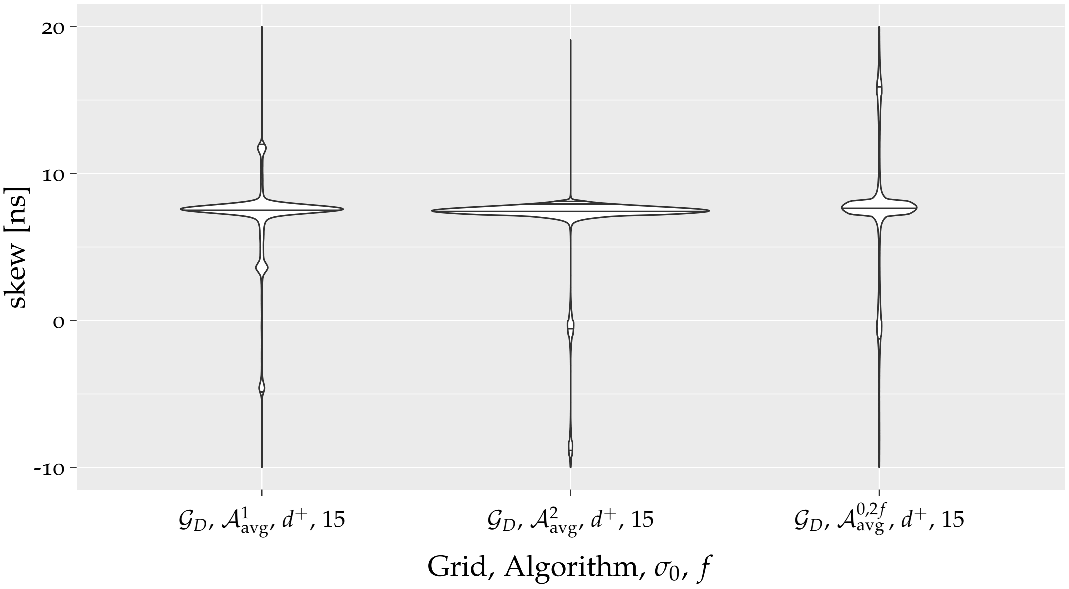


Figure 4.32: Comparison of the inter-layer skew of the time-based algorithms in the 1–local fault and 2–local fault model in $\mathcal{G}_D$ with an input skew of $\sigma_0 = d^+$ and 15 Byzantine faults. The plot has been limited to be within $[-10, 20]$ ns to improve the visibility of the majority of the skews.

and $\mathcal{A}^{1,2f}_{\text{mid},\,80\,\text{ns}}$ (Tables A.17 and A.18). A comparison with $\mathcal{A}^{1,2f}_{\text{avg},\,25\,\text{ns}}$ and $\mathcal{A}^{1,2f}_{\text{mid},\,25\,\text{ns}}$ (Tables A.15 and A.16, Page 347pp) shows that the fault-free case is unaffected by the increased offset value, after compensation of the inter-layer skews by the respective offset value. However, in the presence of faults IQD^0 increases, while the other IQD's show no clear tendency. Nonetheless, the safety margin of the offset value is still as tight as with 25 ns.

Yet, for the time-based algorithms, a different picture has to be drawn. In the low input skew set, as shown Figure 4.31a, $\mathcal{A}^{2,2f}_{\text{avg}}$ shows a remarkable performance. However, with rising input skews, see Figure 4.31b, their performance dramatically deteriorates.

The behavior of the grid under $\mathcal{A}^{d,2f}_{\text{mid}}$ is by and large the same as for $\mathcal{A}^{d,2f}_{\text{avg}}$ and hence omitted; we refer to Appendix A.13 on Page 325pp for the plots and to Appendix A.14 on Page 342pp for the tables of the simulation results.

Note that, even though $\mathcal{G}_D$ has 5 neighbors, its suitability for the 2–local fault model is limited, as in such a case only one neighbor will determines the firing time of a node. Nonetheless, the grid delivers results similar to $\mathcal{G}_C$ in both fault models. The exception are the time-based algorithms in the high input skew set with faults, see Figure 4.32, where we can see an increase of the IQD's at all considered levels.

We refer to Appendix A.15 on Page 353pp for the plots and to Appendix A.16 on Page 360pp for the tables of the simulation results.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

4.4 Conclusion

We have conducted a suite of exploratory simulations to find alternative grid topologies and firing algorithms. The results were presented and reviewed to highlight the specific advantages and shortcomings of each combination.

The extensive results show diverse optimality conditions. Hence, the hope to find a single replacement for HEX for all settings was not fulfilled. However, given the goal for improving the handling of faults in a well-behaved input skew scenario, algorithm $\mathcal{A}^2_{\text{any}}$ has shown a clear advantage over $\mathcal{A}^2_{\text{adj}}$ used by HEX. Given the desire for low hardware costs, the low-connectivity $\mathcal{G}_B$ topology provides good results in combination with $\mathcal{A}^2_{\text{any}}$. Hence, we consider this combination a suitable fit for the requirements posed.

A shift of the requirements towards fault-tolerance and high input skews reveals that the combination of $\mathcal{G}_C$ and $\mathcal{A}^2_{\text{any}}$ achieves the tightest inter-layer skews.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

CHAPTER 5

TRIX – A TRANSISTOR-LEVEL IMPLEMENTATION

> Any intelligent fool can make things bigger,
> more complex, and more violent.
> It takes a touch of genius—and a lot of
> courage—to move in the opposite direction.
>
> — ERNST FRIEDRICH SCHUMACHER

GIVEN THE ROOM for improvement that HEX left as revealed in Chapter 3, the previous chapter investigated alternative grid topology and alternative implementation of the nodes. Our choice of a successor for HEX relies on the assumption that a high-precision clock synchronization exists that provides a well-synchronized input to the initial layer of the grid. As we also strive for an implementation with low overhead, in particular small area for each node, the simple $\mathcal{G}_B$ topology (Figure 5.1) with $\mathcal{A}^2_{\text{any}}$ (Algorithm 5.1) provides the best solution. Despite the compellingly simple interconnection topology, *TRIX*, for triangular grid, depicted in Figure 5.1, significantly outperforms HEX in comparable settings.

In this chapter, we will first introduce, in Section 5.1, the structure of the grid and the algorithmic behavior of the nodes. In Section 5.2, we will provide a transistor-level cell implementation of a TRIX node. Section 5.3 provides the setup and results of the simulation-based evaluation of the TRIX cell, which resulted in an accurate delay model. Furthermore, we present grid-level simulations, using the acquired delay model, and discuss their results, which allow to reason about the behavior of TRIX as a CDN on a chip. The results of this chapter are summarized in Section 5.4.

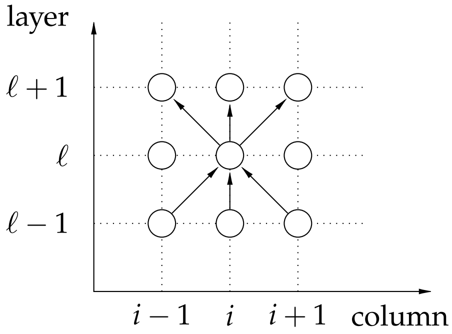


Figure 5.1: Overview of node (ℓ, i) and its links in the $\mathcal{G}_B$ topology. Column coordinates are modulo W and layer coordinates (rows) are between 0 and L.

Algorithm 5.1: TRIX forwarding algorithm for nodes in layer $\ell > 0$.

```
upon having active pulses from at least two neighbors do
    produce pulse;
    sleep for some time;
```

5.1 Topology & Pulse Forwarding

We again consider nodes executing a pulse generation and forwarding algorithm, which communicates by sending pulses over the links of the grid. Letting $L \in \mathbb{N}$ denote its length and $W \in \mathbb{N}$ its width, the set of nodes V is the set of tuples $\{(\ell, i) \in [L+1] \times [W]\}$. Here, $[L+1] := \{0, \ldots, L\}$ denotes the row index set, referred to as *layers*, and $[W] := \{0, \ldots, W-1\}$ the column index set of the nodes in the grid. For each node $(\ell, i) \in V$, $\ell \in [L+1] \setminus \{0\}$, $i \in [W]$, the edges in E define the directed links to other nodes in the grid. Every pulse transmitted over such a link is affected by a signal propagation delay. We call incoming and outgoing links to neighboring nodes of the same layer, e.g., from (ℓ, i) to $(\ell, i-1 \mod W)$, intra-layer links. Note that the considered topology has no intra-layer links. Each node (ℓ, i) on layer $\ell < L$ has outgoing links to the three intra-layer neighbor nodes $(\ell+1, i-1)$, $(\ell+1, i)$, and $(\ell+1, i+1)$, where we take all column indices modulo W. Figure 5.1 shows the resulting topology from the perspective of a node.[1] Nodes on layer 0 represent the clock sources of the system.

For each clock pulse, each (non-faulty) node in layer $\ell \neq 0$ executes a simple algorithm whose triggering rule is as follows: As soon as the node received pulses from any two of its layer $\ell-1$ neighbors, it broadcasts a pulse to its neighbors on layer $\ell+1$. This algorithm, Algorithm 5.1, is the concretization of $\mathcal{A}^2_{\text{any}}$ (Algorithm 4.2 on Page 118)

---

[1] In practice, layers would be concentric rings around the clock source, resulting in the "width" of the grid scaling with the layer number. This will not have a negative impact on the synchronization quality, however, as this means that the skew from previous layers is distributed over more nodes, and we are interested in minimizing the skew between adjacent nodes.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

for the specific requirements arising from targeting a transistor-level implementation. We will show in Section 5.2 that our specific implementation of this triggering rule guarantees that even if one in-neighbor of a node is faulty, the node will trigger only due to the pulse(s) provided by its correct neighbor(s). This resilience provided by the algorithm ensures that the grid can tolerate up to one in-neighbor of a node failing in an arbitrary manner, including the sending of deteriorated pulses.

After a node has triggered, i.e., generated a pulse, the node temporarily disables its triggering mechanism, to avoid triggering multiple times for a single pulse. After a sufficiently long time, such that the received pulses from the correct in-neighbors have passed, a node re-enables its triggering mechanism again. Note that this does not require accurate timing, unless one aims for a very high pulse frequency.

Note that local clock multiplication, as has been discussed for HEX in Section 3.5 on Page 98, can be applied to TRIX as well. We did not analyze the required timing constraints, however, as for the clock distribution itself the assumption that the clock generation on layer 0 waits sufficiently long before generating the next pulse is sufficient. Nonetheless, these timing constraints are relevant for self-stabilization of the system, i.e., recovery from arbitrary transient faults. In Section 5.2, we argue why the combination of our implementation of the triggering rule and the simple topology guarantees self-stabilization given fitting timing constraints.

Following the previous chapters, we measure the quality of the synchronization in terms of the skew between spatially near-by nodes, that is, the difference in their triggering time for the same pulse. The rationale behind this is that near-by nodes typically have to communicate with each other, which imposes the most stringent timing requirements in a system. Accordingly, we do not only take into account neighboring nodes in terms of the connectivity defined by the grid, but also the skew to the nodes (ℓ, i) and $(\ell, i+1)$, i.e., the intra-layer neighbors. Once the system has stabilized, each pulse causes each (non-faulty) node to trigger exactly once, meaning that we can determine the skew for an individual pulse, by comparing the triggering times of the nodes as discussed above.

Definition 5.1 (Skew):
For a given pulse, we denote by $t_{\ell,i}$ the triggering time of the non-faulty node (ℓ, i). For $\ell \in [L+1]$, the *intra-layer skew* of layer ℓ is defined as

$$\sigma_\ell := \max_{i \in [W]} \{ |t_{\ell,i} - t_{\ell,i+1}| \}. \tag{5.1}$$

For $\ell \in [L+1] \setminus \{0\}$, the *inter-layer skew* of layer ℓ is defined as

$$\hat{\sigma}_\ell := \max_{i \in [W]} \{ t_{\ell,i} - t_{\ell-1,i-1}, t_{\ell,i} - t_{\ell-1,i}, t_{\ell,i} - t_{\ell-1,i+1} \}. \tag{5.2}$$

□

Note that the inter-layer skew is biased due to the propagation delay of the pulses over the wires. A clock locally attached to a node is not heavily impacted by this: Given that the spread of the inter-layer skews is kept small, a reasonably accurate prediction of the skew can be made and compensated accordingly. In contrast, the intra-layer skew is unbiased and symmetrical.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

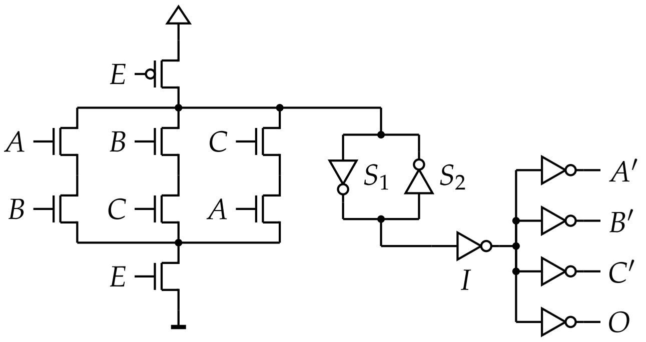


Figure 5.2: The transistor implementation of the triggering circuit. A, B, and C are the inputs, A', B', C', and O the outputs, and E is a locally generated enable signal shaping the pulse. I is a high-threshold inverter, masking metastability of the storage loop consisting of S_1 and S_2, which may be caused by degenerate input from a faulty node.

5.2 Firing Logic Implementation

The implementation of the triggering logic has to achieve the following design goals:

(i) Minimizing the delay variation (as it is the principal source of skew within the cell)

(ii) Repeatedly producing well-shaped pulses.

(iii) Being insensitive to erroneous input from a faulty node.

(iv) Allowing for the recovery from transient faults.

Achieving (i), (ii), and (iv) in conjunction with (iii) would be difficult at the implementation-level that VHDL provides. Hence, a customized transistor-level implementation of the logic in the critical path of the cell has been developed that achieves all these goals, using standard components only. Our solution is depicted in Figure 5.2, where

- A, B, and C are the pulse inputs from the in-neighbors,
- A', B', and C' are the outputs to the out-neighbors,
- O is the output driving the node's local clock tree, and
- E is the enable signal, which is set to low after a pulse was forwarded.

The enable signal E, which is controlled by logic components that are not within the critical path, is primarily responsible for ensuring the correct duration of the generated pulse while it is low. After a sufficiently long timeout, E is set back to high. This timeout can be easily implemented by means of a watchdog timer.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

The circuit implements $\mathcal{A}^2_{\text{any}}$, hence it assumes that two of the inputs are provided by non-faulty nodes. A non-faulty node will generate "clean" pulses that are received by each of its out-neighbors with a duration that is long enough for their circuit to generate a pulse and short enough such that the timeout of the enable signal avoids generation of multiple pulses.

Before analyzing the behavior of the circuit under faults, we provide the sequence of a typical pulse cycle in the fault-free case: The system is initialized by setting E to low. This forces the storage loop, constructed by S_1 and S_2, high and thus accordingly the outputs A', B', C', and O to low. Hence, the corresponding neighbors will be provided with a stable low signal. As all nodes in the grid are initialized in the same way, the inputs A, B and C are also low after the signals from the corresponding neighbors have propagated through the grid.

The pulse cycle starts by setting E high. As the circuit contains an incomplete P-stack[2] its output, at the input of I, would be floating, without the storage loop. Observe that, since E is high, the three parallel paths controlled by A, B, and C connect the internal node to ground if at least two of the three signals are high. As soon as any two of the inputs go high, the storage loop is forced low and the outputs go high. Note that once the storage loop has stabilized, it does not matter if any of the input signals returns to low again. Thus, having output pulses with (roughly) fixed duration can be achieved by setting E to low for a fixed time after O went high. As the width of the pulses from the previous layer is fixed too, the input signals return to low before or shortly after E forced the node's output to low. Hence, after E was low for some (short) fixed amount of time, it is safe to return it to high again. Thus, the cycle is complete and the node is ready for the next pulse.

5.2.1 Fault-tolerance

Above, we assumed that all inputs are connected to non-faulty nodes. However, a minor tweak is sufficient to completely mask a faulty node at one of the inputs. First, assume that the initial state of the circuit is as above, except that one of the inputs, say A, provides an arbitrary (faulty) signal. Clearly, this does not close a path that could pull the internal node down, as B and C are still low. Hence, the storage loop is unaffected and no pulse is generated before a pulse from either B or C (or both) arrives.

However, once a signal from B or C arrives, the storage loop can "capture" the pulse. That is, the outputs of S_1 and S_2 will stabilize to low and high, respectively, and the storage loop's state will become independent of the inputs A, B, and C. However, the storage loop may fail to complete the "capturing" if the signal at A is not well-shaped or goes to low again before the third input goes high. Note that this may result in the storage loop to become metastable. To prevent this we utilize the fact that if the output of S_2 moves notably above the threshold voltage of S_1, the storage loop is

---

[2] Recall that the P-stack is the positive part of a CMOS cell, connecting its output to VCC if active. The complementary part is the N-stack. A CMOS cell's output should always be connected to either VCC or GND. A stack is called incomplete if it is the cause for a cell to have a floating output, due to missing paths for certain input signal patterns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

already stabilizing to a low output at S_1, which is then maintained independently of a connection to ground in the N-stack (until E switches to low, of course). Thus, by choosing a higher threshold for the inverter I than for S_2, we mask any such effects at the output: When the output voltage of S_1 is sufficiently high for I to start to switch, the storage loop will already be driven into the stable state, thus ensuring a clean pulse at the outputs A', B', C', and O. We conclude that the pulse is correctly forwarded, both with respect to timing and signal shape. In particular, the logic of the node omitted in Figure 5.2 is driven by a clean signal, implying that the node's logic is not affected by a bad input signal from a faulty node.

5.2.2 Self-stabilization

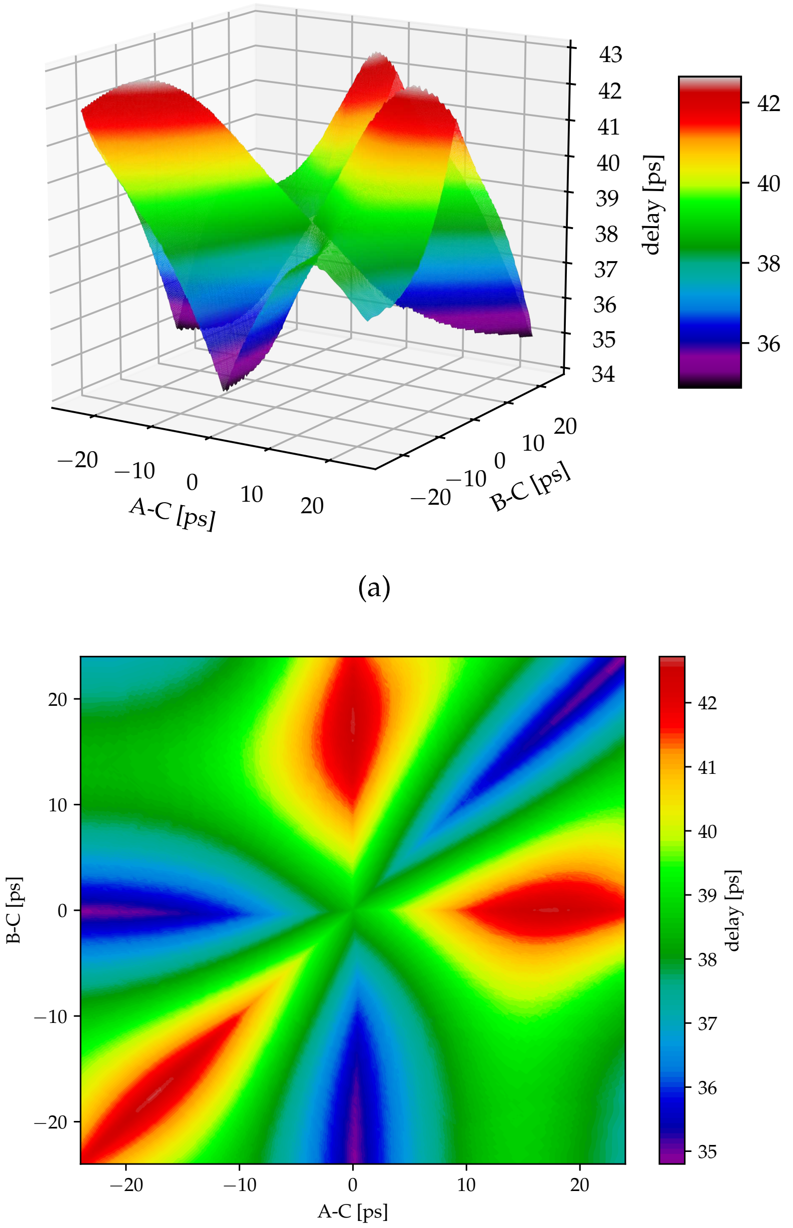
The high-level perspective of the self-stabilization property of TRIX is analog to HEX (cf. Section 3.2.4 on Page 67): Starting with the clock generation in layer 0, the grid will stabilize layer by layer, as every node in each layer stabilizes. An important point to note is that after transient faults, the internal logic of the node may be corrupted and even suffer from metastability. However, metastability is, by definition, a transient state that is extremely likely to subside within a few nanoseconds. Moreover, the logic of the cell relevant for the pulse generation is very simple, meaning that standard design practices can easily ensure that the state-holding components are forced out of any invalid state. Thus, the node will end up in an internal state consistent with some (arbitrary) part of the standard pulse cycle as described above.

From there, the argument of the high-level self-stabilization applies and the node will be triggered by pulses from its correct in-neighbors. This triggering will cause that, after the respective timeouts, E is set to low, re-initializing the node. Following the re-initialization, at the latest in the following pulse the node will trigger in correct phase relation with its preceding layer. As this applies to all correct nodes on layer ℓ, we can argue that within less than three "clean" pulses from layer $\ell - 1$, layer ℓ stabilizes. This guarantees that the TRIX node recovers from transient faults within a number of pulses proportional to the height of the grid L, as does HEX. We remark that, in practice, stabilization typically occurs within a few pulses.

5.3 Simulations

The above presented transistor-cell model of the critical path of the TRIX node allows to determine accurate delay information of the node's switching behavior, depending on the arrival times of the pulses from the in-neighbors. This in turn enables a grid-wide simulation with accurate node switching delays, which allows to evaluate the expected skews in the whole CDN. Furthermore, this approach facilitates a reasonably fair comparison with the performance of clock trees. With an accurate wire delay and port delay model, it would even be possible to acquire results for multi-die or even larger composed systems.

TU Bibliothek, WIEN Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a)

(b)

Figure 5.3: Delay of the circuit from Figure 5.2, with the components in their typical parameter corner, as function of the relative times of the input transitions. It increases if two pulses together arrive early and the third one is late. It decreases if one pulse is early and the other two together arrive late.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

5.3.1 Circuit Delay Evaluation

In order to evaluate the behavior of a grid with TRIX nodes accurately, a comprehensive delay model of a TRIX node had to be established. Therefore, the circuit shown in Figure 5.2 has been modeled in SPICE using transistor models and parasitics for an UMC 65 nm process [UMC05]. The inverter S_2 is modeled using a minimum sized inverter, while the others were chosen to be CLKINV20 to have enough drive strength to ensure fast rise/fall times. The input transistor sizing has been slightly optimized for minimum delay variations under different relative arrival times of the inputs A, B and C. As the node triggers after the second received signal, the delay is measured from the 50% point of the second arriving pulse to the 50% point of the output. Due to the symmetry of the circuit, the delays depend on the relative timing of the pulses only. Hence, one input was set to arrive at a fixed point in time, while the other two signal arrival times were varied within a certain interval. The granularity of the time steps was varied to acquire higher resolution in relevant parts of the parameter space.

The simulation results show a high symmetry between the three input pulse arrival times. For the typical PVT corner (see Figure 5.3), we measured a minimum delay of 35 ps and a maximum delay of 43 ps. In the fast corner, the delays vary within 28 ps and 36 ps, while the slow corner delivered delays within 43 ps and 53 ps. Note that the difference between maximum and minimum delay of the circuit is essentially unaffected by the PVT corner assumed. Plots of the delay for the slow and fast corner can be found in Appendix B. The delay decreases when two pulses are received early and in close proximity, while the third is late. Conversely, the delay increases when one pulse is received early and the other two go high late and in close proximity. This behavior is due to the parasitics capacitances between the components, which need to be charged when a transistor switches and thereby closes an uncharged connection. It is interesting to note that this behavior causes the delay minimum not to be in the center of the plots, where all three pulses arrive at the same time. Instead, there are three distinct minima.

The effects of degenerated inputs have not been studied in SPICE, as finding a pulse form that increases or decreases the delay in a worst-case manner is hard. However, only faulty nodes produce degenerate signals, and these are assumed to be few. Furthermore, benign faults such as stuck-at faults and early or late transitions never produce degenerate signals. Hence, while one may expect that a degenerated signal may affect minimum and maximum delay by a few picoseconds, this “additional” source of skew is very rare. Moreover, as we can see from the grid-level simulations (cf. Figure 5.4), large skews at the input are reduced with increasing layer number, strongly suggests that the grid will automatically mitigate the effect of degenerated signals on the skew.

5.3.2 Simulation of the TRIX Grid

Given the delay characterization of the TRIX node, grid-level simulations could be performed. To facilitate these simulations, the simulation tool developed in the course

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

of Chapter 4 was extended in multiple ways:

1. The absence of intra-layer links allows to calculate pulse propagation very efficiently, as each layer only depends on its direct predecessor layer. Thus, given the state of layer $\ell - 1$, the state of layer ℓ can be immediately computed. Consequently, the grid triggering times in the grid can be computed by a simple iteration over the layers.
2. The node/wire delay model used in Chapters 3 and 4 was based on a uniform distribution. Here, however, we assume normally distributed wire delays, whereas the node delays were, as stated above, extracted from the SPICE simulations.

For our simulation results, 2500 single pulse propagations were performed. The grid size was set to $W = 25$ columns and $L = 50$ layers. We also considered faults and varied the initial skew at layer 0. Faults were placed uniformly at random under the constraint that 1–local fault model holds (cf. Section 3.2.2 on Page 61), i.e., each node has at most one fault on its incoming links. As faulty nodes may behave arbitrarily, they were not considered for skew evaluation, i.e., skews are measured only between connected correct nodes. However, in our simulations we only considered stuck-at 0 (fail-silent) faults.

Concerning the initial skews in layer 0, we choose the nodes' triggering times independently and uniformly distributed between 0 and Δ, for $\Delta \in \{0\text{ps}, 20\text{ps}, 100\text{ps}\}$, resulting in $\sigma_0 \approx \Delta$. Here, $\Delta = 0$ps permits to study the skew exclusively caused during forwarding of the pulse, $\Delta = 20$ps represents inputs that are reasonably well-synchronized, and $\Delta = 100$ps showcases how the grid behaves under bad inputs.

Simulation Framework As $\mathcal{G}_B$ does not include intra-layer links, the simulation can be conducted efficiently on per-layer basis. The resulting state of the nodes in layer ℓ only depends on the state of the nodes in the layers up to $\ell - 1$, which are already computed at this point. By Algorithm 5.1, the triggering time of node (ℓ, i) in layer $\ell > 0$ depends on the arrival time the pulses from the nodes $(\ell - 1, i - 1)$, $(\ell - 1, i)$, and $(\ell - 1, i + 1)$. To calculate the arrival times of the pulses, which is based on their respective triggering time, an independent normally distributed signal propagation delay is added to each triggering time. Based on the interval between the arrival times of the pulses, a lookup in the delay information of the SPICE simulation is conducted. Pulse arrival pattern for which no delay information exist are bilinearly interpolated from existing data points. Finally, another independent normally distributed random value is applied to model additional noise in the system.

The randomness in these delays is assumed to be due to PVT variations. Note that only the effects that differ between the nodes affect the skew, while globally acting influences will affect all nodes in the same way.

The following results were obtained by means of SPICE simulations with all transistors in their typical corner. Accordingly, the signal propagation delays are modeled with 1 ns mean and 1 ps standard deviation, while the noise is considered to have zero mean and a standard deviation of $S = 1$ ps.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

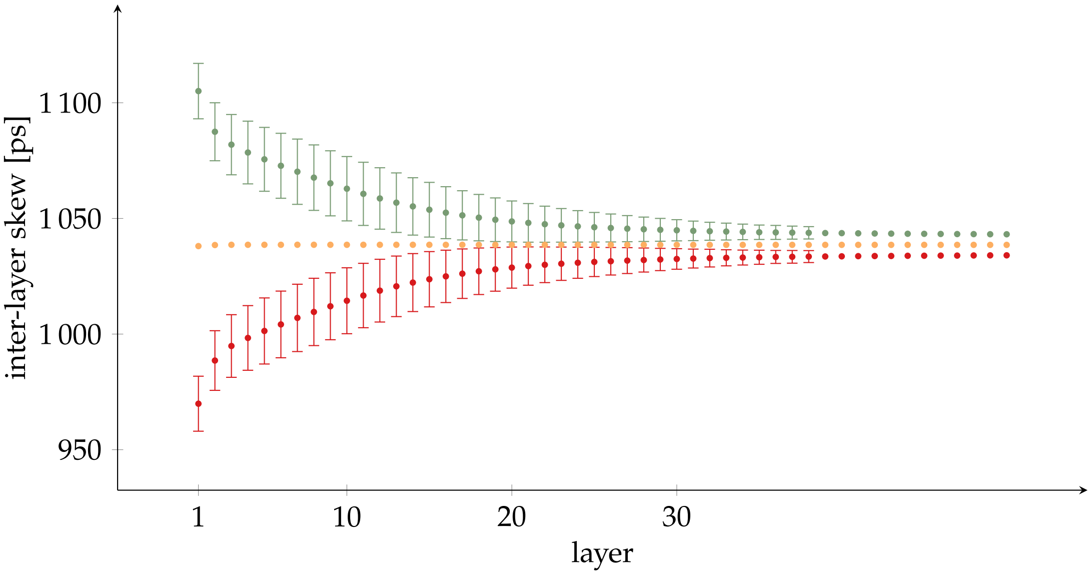


Figure 5.4: Visualization of the inter-layer skew of the data set with $\Delta = 100$, $S = 1\,\text{ps}$, and no faults. Averages and standard deviations over 500 simulation runs are plotted on a per-layer basis. To avoid clutter, the standard deviation was dropped if it was below 2.5 ps. The top data series shows the maximum, the middle the average, and the lower the minimum.

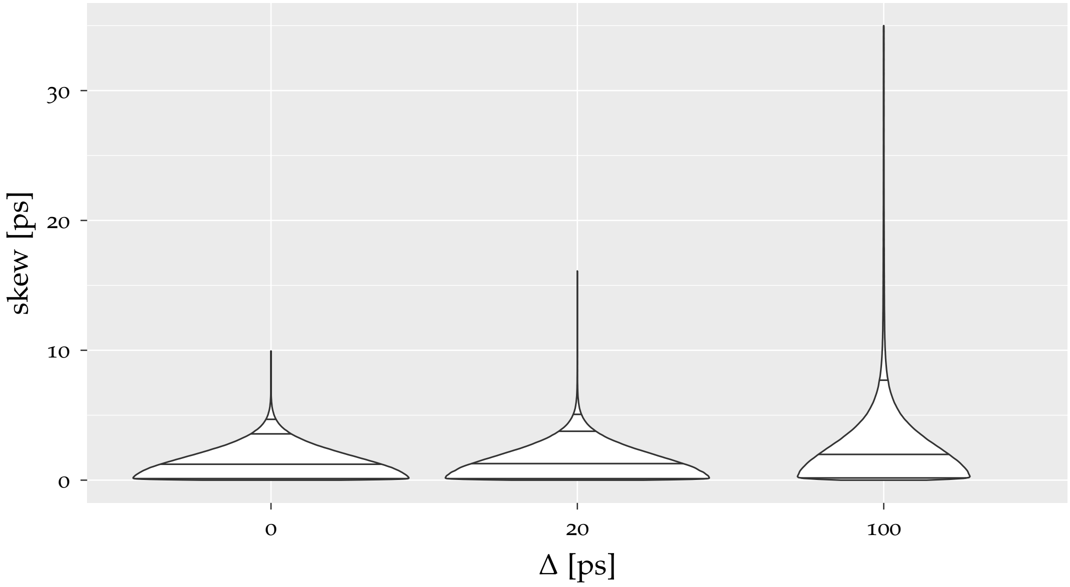
Simulation Results The results of the simulations are visually represented with violin plots, which are basically mirrored kernel density plots.[3] These violin plots were created with R [RC018] and allow an visual comparison of the distributions of the skews between multiple data sets in an aesthetically pleasing way. The violins have horizontal marks to indicate the 1%, 5%, 50%, 95%, and 99% quantile. Due to the structure of the data, for the intra-layer skews, the 1% quantile is not shown. The area of the violins is the same within each plot, hence the violins appearance is independent of the number of samples. This is relevant, as skews from/to faulty nodes are not considered for the skew calculations and hence the number skew can differ between the sets.

Figure 5.5 depicts the calculated skews for the fault-free case. The corresponding statistical data, i.e., the minimum, 1%-quantile, 5%-quantile, average, median, 95%-quantile, 99%-quantile, and maximum, are given in Tables 5.1 to 5.3 under $S = 1\,\text{ps}$, $f = 0$. As we see from the results with $\Delta = 0$, the grid itself maintains small skews under perfect input. The extreme values, acquired in 2500 simulation runs, of the inter-layer skew are less than 18 ps apart. As being approximately twice the skew of the triggering logic, this is an expected result in such a benign setting. It further shows that skews do not accumulate over the layers of the grid. For $\Delta = 20$ and $\Delta = 100$, we see

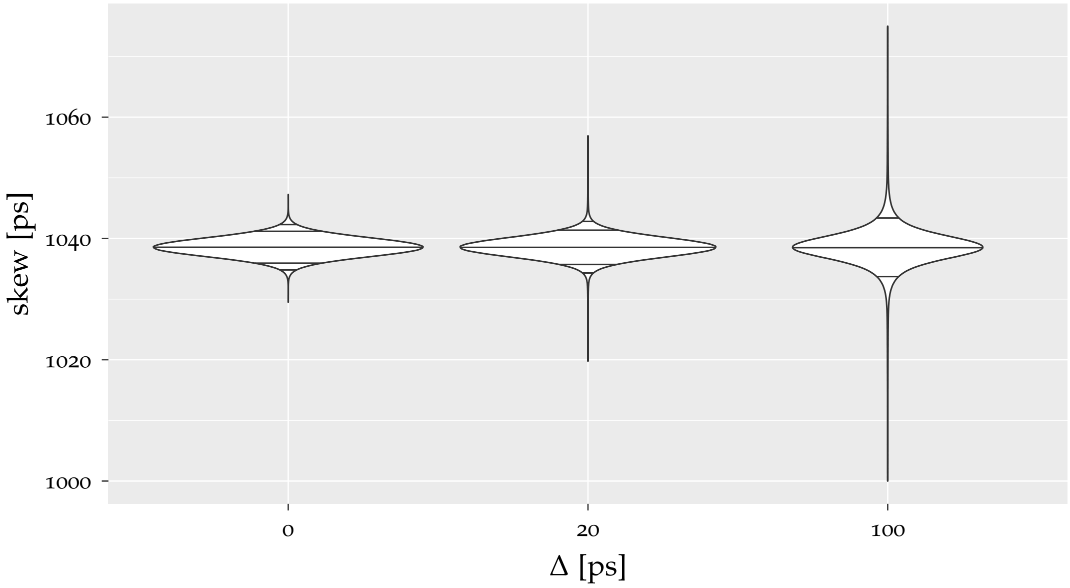
---

[3] Recall that a kernel density estimates the probability density function. That is, the plot represents a continuous and smooth version of the estimated histogram of the provided data set.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



(a) intra-layer skews



(b) inter-layer skews

Figure 5.5: The development of the skews with increasing Δ with $S = 1\,\text{ps}$. Note that, for better readability, we cropped the plot with $\Delta = 100$. Hence the actual values in case of $\Delta = 100$ are in (a) a maximum skew of 93 ps, and in (b) a minimum respectively maximum of 939 ps respectively 1135 ps.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

that the extreme values all change by roughly Δ, while the majority of the skews still remains small. This indicates a "garbage in, garbage out" situation. Indeed, this can be seen when plotting the average maximum, average minimum, and average inter-layer skews, calculated over the 2500 simulation runs, as function of the layer index for $\Delta = 100$: Figure 5.4 shows that the skews drop until the spread between minimum and maximum of the skews is in the range of 10 ps.

Nonetheless, one should take these results with a grain of salt. The discussed simulations represent noise sources, manufacturing differences, etc. by the sum of two normal distributions with standard deviation of 1 ps each. However, we believe that our results demonstrate that, in terms of skew, a well-engineered TRIX grid will perform very similar to more conventional clock distribution topologies. Furthermore, we performed simulations with varying standard deviation S to ensure that the influence is limited as expected (cf. Figures 5.6 to 5.8 and Tables 5.1 to 5.3): The increase of the standard deviation affects even the IQD[5], the difference between q_{95} and q_5, which was to be expected.

Faults in the Grid Complementary to the effects of the initial skew caused by the clock generation system, we also considered the influence of faults on the skews. Considering the results presented in Tables 5.1 to 5.3, we see that even an error rate of close to 1% (10 faulty nodes out of 1250 TRIX nodes) affects the maxima of the skews by roughly 5 ps under perfectly synchronized inputs ($\Delta = 0$). The other statistical parameters vary within a level that could be considered noise. The missing monotonicity of the maxima has to be attributed to the simulations, which are typically imperfect in hitting minimum and maximum values. With larger skews at the input, even this effect disappears, as the influence of the input skews dominates the extreme values. We argue that this data predicts excellent scalability of our clock distribution method.

5.4 Conclusion

We presented TRIX, a fault-tolerant clock distribution scheme. TRIX succeeds at achieving low skews under well-synchronized clock input while having a small hardware footprint.

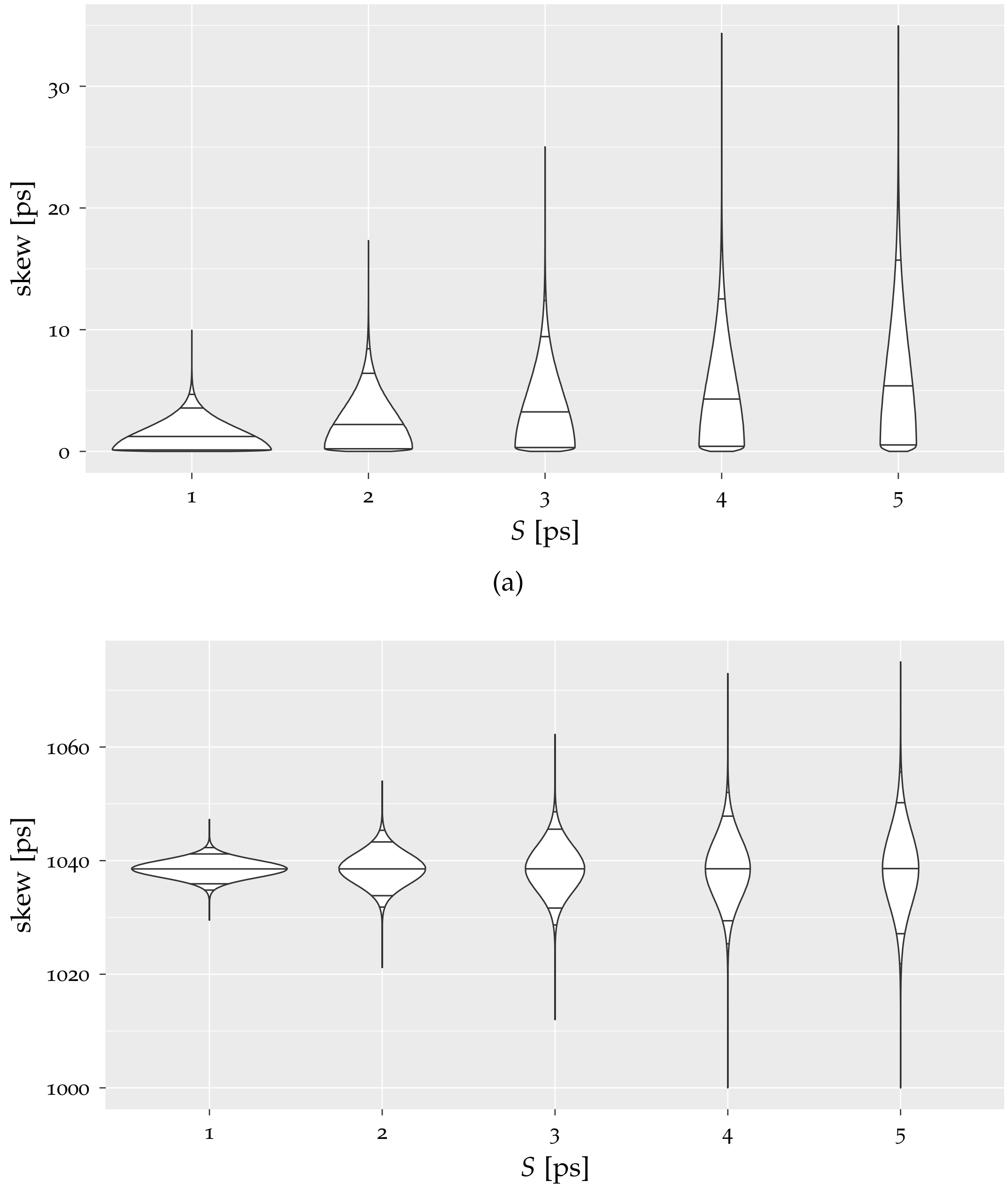
Recall that HEX had node switching delays within $[161, 197]$ ps, excluding wire delays, with an UMC 90 nm process. The constructed TRIX cell shows delays in $[35, 43]$ ps including wire delays with an UMC 65 nm process. A reduction of the delay was expected due to the newer process, however, this large improvement, especially of the variance, is primarily attributable to the efficient design of the transistor cell. Furthermore, the hardware footprint of TRIX is definitely smaller than HEX: Unlike HEX, TRIX requires only 1 timer instead of 4, which are typically large in size.

CDNs used in commercial processors have a skew in the range of $[7, 52]$ ps, as surveyed in [Tam09]. Hence, we are confident that with proper engineering effort, TRIX could be the basis for a fault-tolerant CDN that experiences skews that are only one order of magnitude higher than in non-fault-tolerant CDNs.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

To summarize, we provided an efficient and small transistor-level implementation of the central component of a clock distribution node, resulting in skews comparable to traditional clocking schemes. However, the need for employing small clock trees to locally distribute a node's clock signal is likely to result in slightly larger skews. Nonetheless, due to its high degree of fault-tolerance, TRIX allows to avoid the single point of failure constituted by traditional CDNs.

TU Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

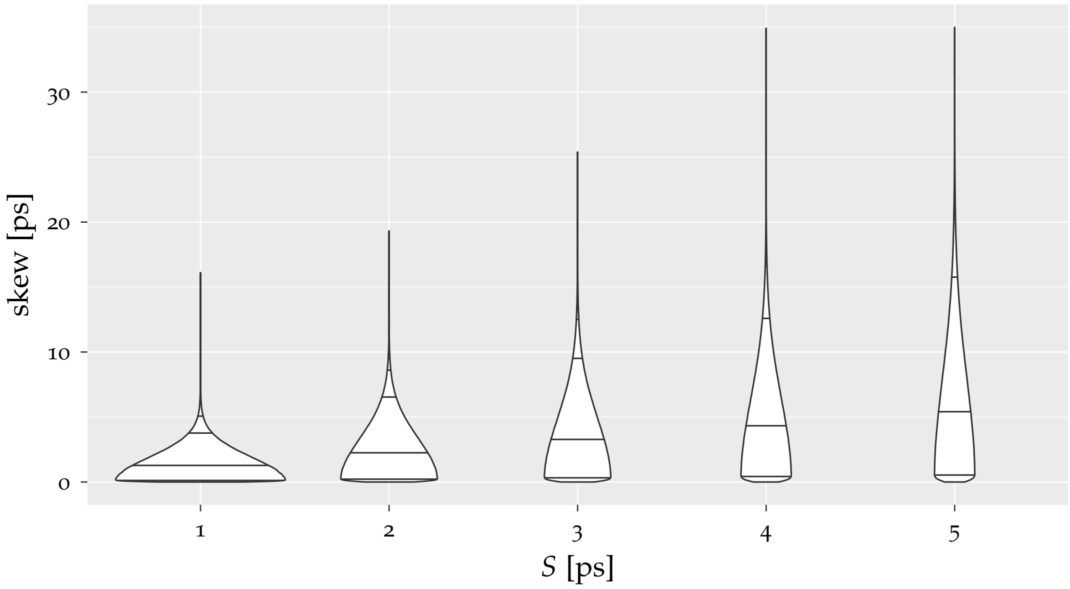


(a)

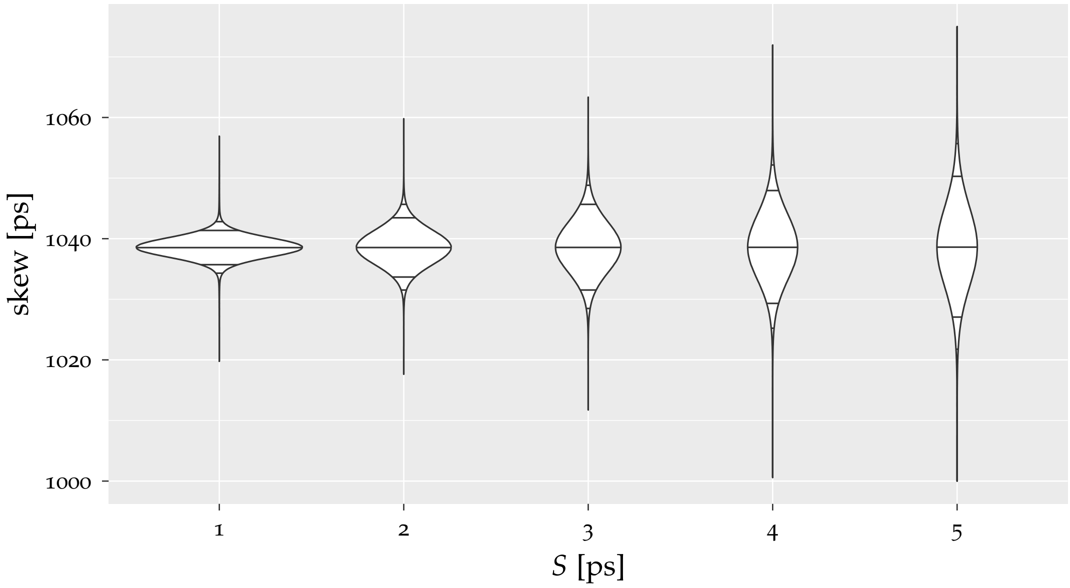
(b)

Figure 5.6: Violin plots of the inter-layer resp. intra-layer skew with an initial skew of $\sigma_0 \approx 0\,\text{ps}$, in the setting of Figure 5.5. The x-axis denotes the standard deviation used for the noise distribution, which has a mean of $0\,\text{ps}$. The intra-layer skews have been cropped at $35\,\text{ps}$, while the inter-layer skews are cropped to be within $[1000, 1075]\,\text{ps}$. We refer to Table 5.1 for a lookup of the actual values.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub



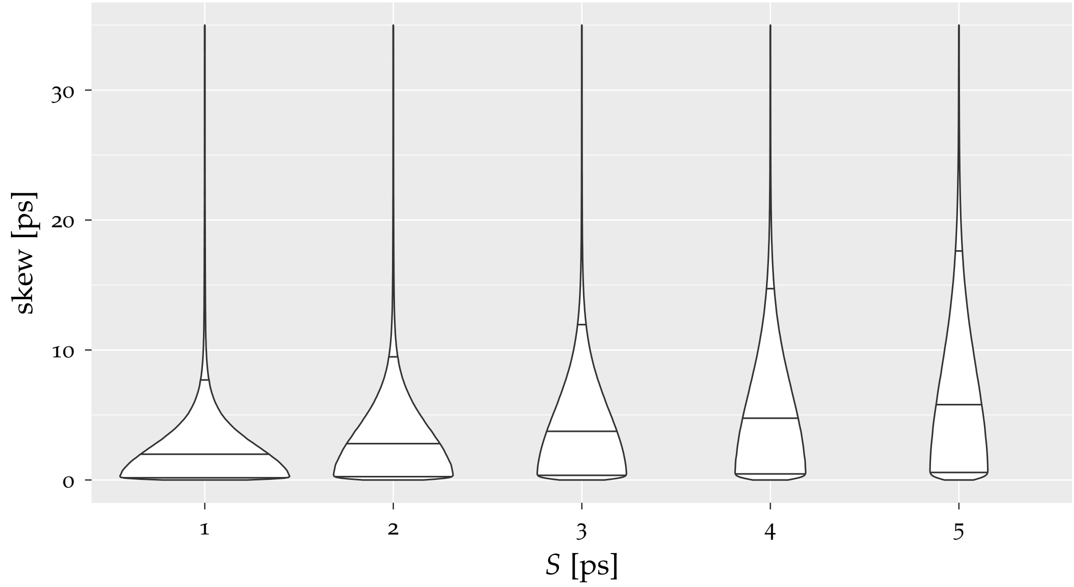
(a)



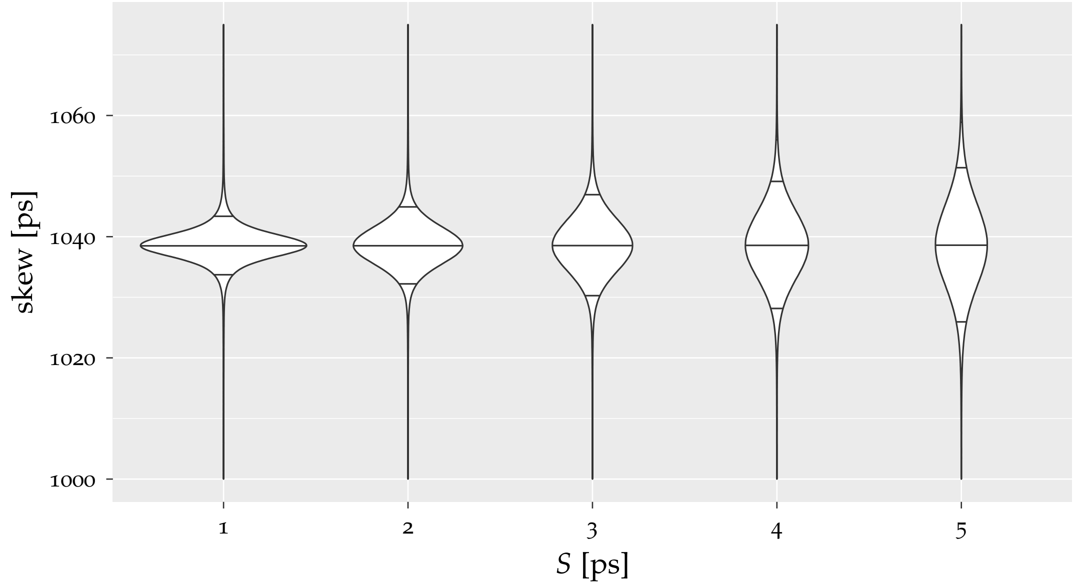
(b)

Figure 5.7: Violin plots of the inter-layer resp. intra-layer skew with an initial skew of $\sigma_0 \approx 20$ ps, in the setting of Figure 5.5. The x-axis denotes the standard deviation used for the used noise distribution, which has a mean of 0 ps. The intra-layer skews have been cropped at 35 ps, while the inter-layer skews are cropped to be within $[1000, 1075]$ ps. We refer to Table 5.2 for a lookup of the actual values.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a)



(b)

Figure 5.8: Violin plots of the inter-layer resp. intra-layer skew with an initial skew of $\sigma_0 \approx 100\,\text{ps}$, in the setting of Figure 5.5. The x-axis denotes the standard deviation used for the used noise distribution, which has a mean of 0 ps. The intra-layer skews have been cropped at 35 ps, while the inter-layer skews are cropped to be within $[1000, 1075]$ ps. We refer to Table 5.3 for a lookup of the actual values.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 5.1: The skew development under an increasing number f of stuck-at-0 faults with $\Delta = 0$ and a noise standard deviation of S. We denote with σ^{op} the intra-layer skews, while $\hat{\sigma}^{\text{op}}$ denotes the corresponding inter-layer skews (both in [ps]).

		intra-layer							inter-layer							
S	f	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
1	0	0.02	0.11	1.45	1.23	3.57	4.69	9.94	1029.53	1034.84	1035.94	1038.56	1038.56	1041.19	1042.31	1047.24
	1	0.02	0.12	1.46	1.23	3.58	4.71	9.86	1026.95	1034.81	1035.93	1038.56	1038.56	1041.20	1042.32	1047.24
	5	0.02	0.12	1.48	1.25	3.64	4.82	9.97	1027.75	1034.67	1035.86	1038.54	1038.55	1041.22	1042.37	1048.72
	10	0.02	0.12	1.51	1.27	3.73	4.95	11.43	1025.97	1034.51	1035.78	1038.52	1038.53	1041.25	1042.43	1049.56
2	0	0.04	0.21	2.61	2.21	6.42	8.44	17.31	1021.18	1031.87	1033.87	1038.58	1038.58	1043.32	1045.37	1053.98
	1	0.04	0.21	2.62	2.21	6.43	8.46	17.23	1018.62	1031.85	1033.86	1038.58	1038.58	1043.33	1045.38	1054.20
	5	0.04	0.21	2.64	2.23	6.48	8.54	17.09	1019.55	1031.75	1033.80	1038.56	1038.56	1043.34	1045.41	1055.90
	10	0.04	0.21	2.66	2.25	6.55	8.64	19.41	1019.72	1031.62	1033.72	1038.54	1038.54	1043.36	1045.45	1057.92
3	0	0.06	0.30	3.84	3.24	9.44	12.41	25.02	1012.02	1028.73	1031.71	1038.62	1038.61	1045.58	1048.65	1062.22
	1	0.06	0.30	3.84	3.25	9.45	12.43	25.15	1009.97	1028.71	1031.69	1038.61	1038.60	1045.59	1048.66	1062.18
	5	0.06	0.30	3.86	3.26	9.50	12.52	26.74	1011.31	1028.60	1031.63	1038.59	1038.58	1045.60	1048.70	1064.47
	10	0.06	0.31	3.89	3.28	9.57	12.62	28.17	1010.54	1028.46	1031.56	1038.57	1038.56	1045.61	1048.75	1067.28
4	0	0.08	0.40	5.09	4.30	12.54	16.53	34.34	1000.03	1025.44	1029.50	1038.66	1038.64	1047.90	1052.09	1072.93
	1	0.08	0.40	5.10	4.30	12.55	16.55	35.91	999.48	1025.40	1029.48	1038.66	1038.64	1047.91	1052.11	1077.56
	5	0.08	0.40	5.12	4.32	12.61	16.66	38.28	997.06	1025.27	1029.41	1038.63	1038.61	1047.92	1052.15	1078.61
	10	0.08	0.40	5.15	4.34	12.69	16.79	40.06	998.82	1025.11	1029.32	1038.60	1038.58	1047.93	1052.22	1081.05
5	0	0.10	0.50	6.38	5.37	15.72	20.82	47.52	988.69	1021.94	1027.22	1038.70	1038.68	1050.27	1055.67	1084.15
	1	0.10	0.50	6.38	5.38	15.74	20.85	47.95	985.43	1021.90	1027.21	1038.69	1038.67	1050.28	1055.69	1089.28
	5	0.10	0.50	6.41	5.40	15.81	20.98	51.05	982.26	1021.75	1027.13	1038.66	1038.65	1050.29	1055.75	1089.89
	10	0.10	0.51	6.44	5.42	15.91	21.16	53.74	986.63	1021.55	1027.03	1038.63	1038.61	1050.30	1055.83	1095.13

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 5.2: The skew development under an increasing number f of stuck-at-0 faults with $\Delta = 20$ and a noise standard deviation of S. We denote with σ^{op} the intra-layer skews, while $\hat{\sigma}^{\text{op}}$ denotes the corresponding inter-layer skews (both in [ps]).

		intra-layer							inter-layer							
S	f	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
1	0	0.02	0.12	1.53	1.28	3.78	5.08	16.10	1019.77	1034.35	1035.75	1038.57	1038.56	1041.39	1042.84	1056.90
	1	0.02	0.12	1.53	1.29	3.79	5.10	16.10	1019.77	1034.32	1035.74	1038.57	1038.56	1041.40	1042.86	1056.90
	5	0.02	0.12	1.56	1.30	3.85	5.20	16.10	1019.77	1034.18	1035.67	1038.55	1038.55	1041.42	1042.90	1056.90
	10	0.02	0.12	1.58	1.32	3.94	5.33	16.10	1019.77	1034.01	1035.58	1038.53	1038.53	1041.46	1042.97	1056.90
2	0	0.04	0.21	2.66	2.25	6.54	8.61	19.31	1017.65	1031.57	1033.72	1038.59	1038.58	1043.48	1045.70	1059.77
	1	0.04	0.21	2.67	2.25	6.55	8.63	19.31	1017.65	1031.55	1033.71	1038.59	1038.58	1043.49	1045.71	1059.77
	5	0.04	0.21	2.68	2.27	6.61	8.71	19.31	1017.65	1031.45	1033.65	1038.57	1038.56	1043.50	1045.75	1059.77
	10	0.04	0.21	2.71	2.28	6.68	8.81	19.37	1017.65	1031.32	1033.57	1038.54	1038.54	1043.53	1045.80	1059.77
3	0	0.06	0.31	3.87	3.27	9.52	12.52	25.38	1011.75	1028.54	1031.59	1038.62	1038.61	1045.71	1048.86	1063.33
	1	0.06	0.31	3.88	3.28	9.53	12.54	25.06	1010.31	1028.52	1031.57	1038.62	1038.60	1045.72	1048.87	1063.33
	5	0.06	0.31	3.90	3.29	9.58	12.63	27.67	1010.91	1028.41	1031.52	1038.60	1038.58	1045.73	1048.91	1064.95
	10	0.06	0.31	3.92	3.31	9.66	12.73	28.02	1010.06	1028.27	1031.43	1038.57	1038.56	1045.75	1048.96	1067.88
4	0	0.08	0.40	5.12	4.33	12.60	16.61	34.90	999.15	1025.31	1029.40	1038.67	1038.64	1048.01	1052.23	1071.94
	1	0.08	0.40	5.13	4.33	12.62	16.63	34.90	999.15	1025.28	1029.39	1038.66	1038.64	1048.01	1052.24	1075.84
	5	0.08	0.40	5.15	4.34	12.67	16.74	40.11	997.08	1025.15	1029.32	1038.63	1038.62	1048.02	1052.29	1079.59
	10	0.08	0.41	5.18	4.37	12.76	16.88	40.95	998.65	1024.99	1029.23	1038.60	1038.59	1048.04	1052.35	1080.58
5	0	0.10	0.50	6.40	5.39	15.77	20.88	49.98	987.48	1021.85	1027.15	1038.70	1038.68	1050.36	1055.76	1083.74
	1	0.10	0.50	6.41	5.40	15.79	20.91	49.98	985.83	1021.81	1027.13	1038.70	1038.67	1050.36	1055.78	1088.43
	5	0.10	0.50	6.43	5.41	15.87	21.05	51.12	982.40	1021.65	1027.05	1038.67	1038.65	1050.37	1055.84	1090.79
	10	0.10	0.51	6.47	5.44	15.96	21.23	53.81	986.97	1021.45	1026.95	1038.63	1038.61	1050.38	1055.93	1094.62

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 5.3: The skew development under an increasing number f of stuck-at-0 faults with $\Delta = 100$ and a noise standard deviation of S. We denote with σ^{op} the intra-layer skews, while $\hat{\sigma}^{\mathrm{op}}$ denotes the corresponding inter-layer skews (both in [ps]).

		intra-layer							inter-layer							
S	f	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
1	0	0.04	0.18	3.28	2.02	8.51	33.85	96.47	941.04	1020.09	1033.65	1038.57	1038.56	1043.63	1056.87	1135.59
	1	0.04	0.19	3.29	2.03	8.53	33.86	96.47	941.04	1020.03	1033.62	1038.57	1038.56	1043.63	1056.87	1135.59
	5	0.04	0.19	3.32	2.06	8.64	34.09	96.47	941.04	1019.73	1033.52	1038.54	1038.54	1043.67	1056.94	1135.59
	10	0.04	0.19	3.37	2.08	8.78	34.24	96.47	941.04	1019.36	1033.40	1038.52	1038.53	1043.71	1057.01	1135.59
2	0	0.05	0.26	4.06	2.84	10.16	32.81	99.68	940.75	1020.58	1032.13	1038.59	1038.58	1045.18	1056.48	1137.39
	1	0.05	0.26	4.07	2.84	10.18	32.81	99.68	940.75	1020.51	1032.11	1038.59	1038.58	1045.18	1056.49	1137.39
	5	0.05	0.26	4.11	2.87	10.28	33.06	99.68	940.75	1020.21	1032.02	1038.56	1038.56	1045.21	1056.58	1137.39
	10	0.05	0.27	4.14	2.89	10.43	33.19	99.68	940.75	1019.84	1031.91	1038.53	1038.53	1045.25	1056.66	1137.39
3	0	0.07	0.35	5.05	3.78	12.59	31.71	102.90	940.25	1020.30	1030.16	1038.62	1038.61	1047.22	1056.85	1139.19
	1	0.07	0.35	5.06	3.79	12.62	31.71	102.90	940.25	1020.23	1030.14	1038.62	1038.60	1047.22	1056.87	1139.19
	5	0.07	0.35	5.09	3.81	12.71	31.98	102.90	940.25	1019.94	1030.05	1038.59	1038.58	1047.25	1056.99	1139.19
	10	0.07	0.35	5.13	3.83	12.86	32.15	102.90	940.25	1019.53	1029.94	1038.56	1038.55	1047.28	1057.14	1139.19
4	0	0.09	0.44	6.14	4.79	15.37	31.00	106.12	938.97	1018.68	1028.02	1038.66	1038.64	1049.44	1058.61	1140.99
	1	0.09	0.44	6.15	4.79	15.40	31.03	106.12	938.97	1018.59	1028.00	1038.65	1038.64	1049.45	1058.63	1140.99
	5	0.09	0.44	6.18	4.81	15.49	31.29	106.12	938.70	1018.30	1027.90	1038.62	1038.61	1049.47	1058.75	1140.99
	10	0.09	0.45	6.23	4.84	15.65	31.54	106.12	938.97	1017.90	1027.77	1038.58	1038.58	1049.50	1058.91	1140.99
5	0	0.11	0.54	7.31	5.83	18.35	31.51	109.34	936.69	1016.00	1025.75	1038.69	1038.67	1051.75	1061.32	1142.78
	1	0.11	0.54	7.32	5.83	18.37	31.59	109.34	936.69	1015.94	1025.73	1038.68	1038.67	1051.76	1061.35	1142.78
	5	0.11	0.54	7.35	5.86	18.48	31.90	109.34	934.95	1015.63	1025.63	1038.65	1038.64	1051.78	1061.48	1142.78
	10	0.11	0.54	7.40	5.88	18.65	32.27	109.34	936.69	1015.23	1025.48	1038.60	1038.60	1051.81	1061.62	1142.78

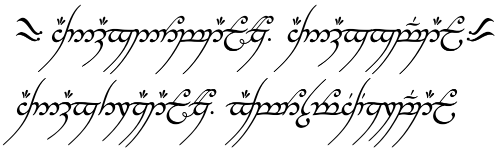
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek, Your knowledge hub, WIEN

CHAPTER 6

POINT-TO-POINT COMMUNICATION IN MULTI-SYNCHRONOUS ARCHITECTURES

> Everything not saved will be lost.
>
> — NINTENDO „QUIT SCREEN“ MESSAGE



> — THE LORD OF THE RINGS, J. R. R. Tolkien

BUILDING UPON the results of the previous chapters, we consider the problem of point-to-point communication in self-stabilizing and possibly Byzantine fault-tolerant multi-synchronous systems, e.g., GALS systems and hardware architectures in general. As elaborated in Section 2.6, various solutions for communication in GALS have been proposed. However, none of the existing solutions can be used if self-stabilization is added to the picture. For example, in the case where the inevitable read/write pointers for the FIFO buffer become corrupted, even the complete recovery of the clocking system is not enough to eventually resume correct communication between reader and writer later on. Actively recovering the state of the read/write pointers is difficult, however, as (i) reader and writer are not always aware of a corruption and (ii) communication must not be stalled deliberately for recovery purposes during normal operation in many applications. Moreover, a communication

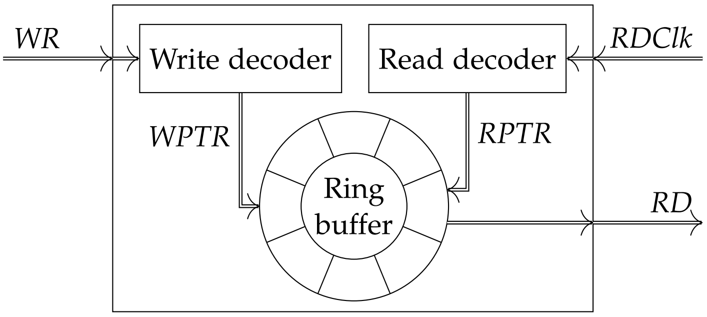


Figure 6.1: An implementation of a FIFO ring buffer with size M. The writer provides a clock and data bundle via WR. The write decoder generates a memory address to write the data, over $WPTR$, into the ring buffer. The reader provides a clock signal $RDClk$, which the read decoder converts into a memory address. This memory address is applied to the ring buffer via $RPTR$, and results in the corresponding memory cell content being outputted at RD.

primitive in the setting of HEX/TRIX should be able to utilize the skew guarantees provided by them in order to optimize the performance of the system.

In this chapter we will introduce a communication primitive, which facilitates metastability-free self-stabilizing communication between correct components in different clock domains. Relying on a fast clock (Section 3.5), it allows to communicate at a much higher frequency as could be derived from a notion of time derived from HEX/TRIX directly.

Our solution relies on a ring buffer with resetable read- and write-pointers as shown in Figure 6.1. Compared to related work for GALS, we refrain from using empty/full-detectors, which are a prime source for metastability issues. Instead, we provide detailed formulas for computing the buffer size and a suitable initial pointer offset for a given set of clock parameters. By periodically reinitializing the positions of the pointers, we ensure self-stabilization.

First, the specific parameters of the underlying clocking system are defined in Section 6.1. The design of the solution, and its properties, are presented in Section 6.2. Sections 6.3 to 6.6 provide the results of the analytical evaluation of the required size of the ring buffer. In Section 6.7, we present a VHDL implementation of our approach and some simulation results, which demonstrate its feasibility. In Section 6.8, we conclude this chapter.

6.1 Clock Model

We consider a circuit consisting of clock domains $1, \ldots, n$, where clock domain i is driven by a dedicated *slow clock* $C_i \colon \mathbb{R} \to \mathbb{N}$ that maps (continuous) real-time $t \geq 0$ to (discrete) clock time $T = C_i(t)$. Correct clocks generate (right-continuous) step

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

functions, i.e., assume all values $0, 1, 2, \dots$ in strict succession, where t_i^k with $t_i^0 = 0$ denotes the smallest real-time with $C_i(t_i^k) = k$. We say that C_i generates tick k at time t_i^k.

Our solution will depend on the synchronization properties guaranteed by the clocking system. Correct clocks will be characterized by (i) their *clock skew*, a time bound within which they produce the same tick k, and (ii) their *tick separation time* between any two consecutive ticks. Formally, these parameters are defined as follows:

Definition 6.1 (Slow clock parameters):
Given a fixed pair of correct clock domains i, j driven by correct clocks, they are characterized by:

1. *Skew bounds* $\sigma'_{i,j}$, defined as $\sigma'_{i,j} := \left[\sigma'^{-}_{i,j}, \sigma'^{+}_{i,j}\right]$ with real $\sigma'^{-}_{i,j} \leq t_j^k - t_i^k \leq \sigma'^{+}_{i,j}$ for all $k \geq 0$.
2. *Tick separation time lower bound* Γ'_i, defined as $0 < \Gamma'_i \leq t_i^{k+1} - t_i^k$ for all $k \geq 0$. □

Since synchronized clocks generated by approaches like [DFL+14; DFL+16] have typically a fairly low-frequency, i.e., have a relatively large Γ'_i, we assume that every clock domain i employs local clock multiplication for speeding-up C_i. Following [PSSL13], we assume that every clock domain i implements *fast clocks* $F_i^k\colon \mathbb{R} \to \{0, \dots, B-1\}$ that map (continuous) real-time $t \geq 0$ to (discrete) clock time $T = F_i^k(t)$. We assume that every tick k of the slow clock C_i starts, after a possibly varying slack time within $\left[\epsilon_i^-, \epsilon_i^+\right]$, a dedicated fast clock F_i^k. This fast clock generates exactly B ticks $0, \dots, B-1$ in strict succession, a so called *burst*, plus a final *virtual* tick B. This virtual tick is only used conceptually for cleanly separating consecutive fast clocks F_i^k and F_i^{k+1}. Analogously to the slow clocks, we say that a fast clock F_i^k generates tick b at time $t_i^{k,b}$. Incorporating some slack in our model allows to cover circuit delays, setup/hold times and other small timing uncertainties, see Section 6.2.

In addition, we assume that a dedicated *period signal* p_i is issued along with the slow clock tick at time t_i^k, which can be sampled at time $t_i^{k,0}$ in the clock domain i. We require that this signal is issued if and only if $k = 0 \bmod I$, for some integer parameter (the *period interval*) $I > 0$ that we will choose. Whereas we assume that the period signal occurs at the same slow clock tick number, in all correct clock domains, after self-stabilization of the clocking system, we stress that this does not at all require a chip-global reset.

Formally, we require the following properties:

Definition 6.2 (Fast clock parameters):
Let $F_i^k(t)$ be a fast clock in clock domain i, started by the correct slow clock's tick k such that $t_i^k + \epsilon_i^- \leq t_i^{k,0} \leq t_i^k + \epsilon_i^+$ for $0 \leq \epsilon_i^- \leq \epsilon_i^+$, cf. Figure 6.2. It is characterized by

1. the frequency range of the fast clock:
$\forall k \geq 0, \forall b \in \{0, \dots, B-1\} : \frac{1}{f_{i,\max}} \leq t_i^{k,b+1} - t_i^{k,b} \leq \frac{1}{f_{i,\min}}$,

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

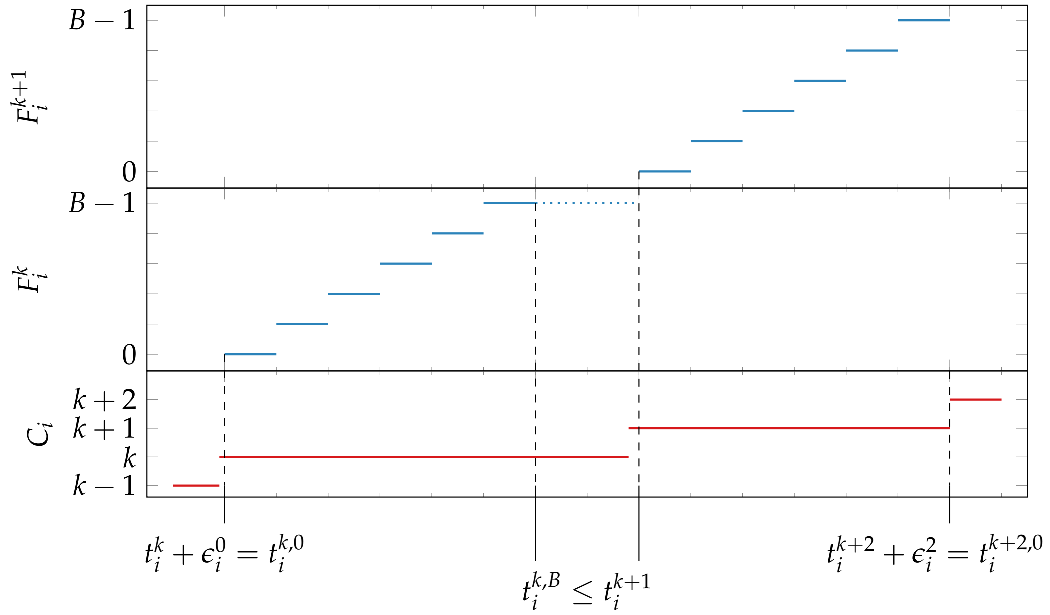


Figure 6.2: Visualization of the slow clock C_i and its corresponding fast clocks F_i^k and F_i^{k+1} with $B = 6$. Note that while ϵ_i^0 and ϵ_i^2 can differ, they both are in $\left[\epsilon_i^-, \epsilon_i^+\right]$.

2. the number of fast clock ticks B, which must satisfy $\frac{B}{f_{i,\max}} \leq t_i^{k,B} - t_i^{k,0} \leq \frac{B}{f_{i,\min}} \leq \Gamma_i$. □

Note that it is impossible to implement a fast clock that satisfies Definition 6.2 on top of a slow clock that does not guarantee some Γ_i' ensuring $\Gamma_i \leq \Gamma_i' - \epsilon_i^+ + \epsilon_i^-$, as it would be impossible to always fit in B fast clock ticks between two slow clock ticks. For non-zero slack $\left[\epsilon_i^-, \epsilon_i^+\right]$, this is even true for the extreme case $B = 1$, where the fast clock is just the slow clock delayed by the slack.

The analysis in the subsequent sections will focus on the times $t_i^{k,0}$, $t_j^{k,0}$ when the fast clocks F_i^k, F_j^k generate their tick 0. The following Lemma 6.3 is an easy consequence of Definition 6.1 and Definition 6.2, which takes into account the (variable) but bounded slack for starting the fast clocks:

Lemma 6.3 (Fast clock skew parameters): *Correct fast clocks F_i^k, F_j^k in correct clock domains i, j satisfy $\sigma_{i,j}^- \leq t_j^{k,0} - t_i^{k,0} \leq \sigma_{i,j}^+$ and $\Gamma_\ell \leq t_\ell^{k+1,0} - t_\ell^{k,0}$ for all $k \geq 0$ and $\ell \in \{i, j\}$, where $\boldsymbol{\sigma}_{i,j} := \left[\sigma_{i,j}'^- - \epsilon_i^+ + \epsilon_j^-, \sigma_{i,j}'^+ - \epsilon_i^- + \epsilon_j^+\right]$ and $\Gamma_i := \Gamma_i' - \epsilon_i^+ + \epsilon_i^-$.* □

Alternatively, synchronized clocks can also be characterized by their *precision bound* π, which is the maximum difference in the clock readings between any two correct clocks at the same real-time. We will consider the *signed* precision $\boldsymbol{\pi}_{i,j} = \left[\pi_{i,j}^-, \pi_{i,j}^+\right]$,

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

which is formally defined via $\pi_{i,j}^- \leq F_i^k(t) - F_j^k(t) \leq \pi_{i,j}^+$ for all $t \geq 0$. Note that $\boldsymbol{\pi}_{i,j}$ and the clock parameters given in Definition 6.2 are of course strongly related. In fact, the analysis in the subsequent sections will make this connection explicit.

In the subsequent sections, node i will always be the reader and j the writer in our point-to-point communication protocol. To keep the notation simple, we will thus drop the subscripts from the parameters in Definition 6.2 and Lemma 6.3: We use $\boldsymbol{\sigma} = [\sigma^-, \sigma^+] := \boldsymbol{\sigma}_{i,j}$, $\boldsymbol{\pi} = [\pi^-, \pi^+] = [-\pi_w, \pi_r] := \left[\pi_{i,j}^-, \pi_{i,j}^+\right]$ (where π_r covers "reader's clock tick occurs before writer's", and $-\pi_w$ covers the situation "writer's clock tick occurs before reader's").

6.2 Description of the Communication Scheme

Our communication scheme, shown in Figure 6.1, relies on a FIFO buffer with an offset between read and write pointer to compensate for the clock skew and frequency deviations. The high-level overview of our implementation is shown in Figure 6.3. The ring buffer (*RB*) is of size M and has dedicated read- and write-ports. The read decoder (*RD*) and write decoder (*WD*) are implemented as counters, clocked by clk^{rd} and clk^{wr}, respectively, which are reset to their initial position when p^{rd} and p^{wr}, respectively, is asserted. Recall that the period signals occur periodically and mark the start of the burst of slow clock tick $k = 0 \bmod I$: When a fast clock tick occurs, and the period signal is not asserted, the corresponding counter is incremented (modulo the size M of the ring buffer *RB*). However, if the period signal is asserted when the fast clock tick occurs, the counter is synchronously reset to a fixed initial position, which is different for reader (0) and writer (τ_w).

As for the data exchange between writer and reader, we assume that the writer wants to send a data item to the reader with every fast clock tick. We stipulate that the reader knows, a priori, if data has been written by the writer in some particular clock tick. With respect to memory access timing, we assume that the writer issues the data along with the clock tick, so that it can be written to *RB* after the counter increment/reset has been completed and the memory address has stabilized. These timing parameters, as well as setup/hold-times and wire delays, are accounted for in the slack $\left[\epsilon_i^-, \epsilon_i^+\right]$, which in turn is included in the clock skew. Similarly, we assume that the reader samples the data read out from *RB* some time after its reader clock tick occurred. We do not restrict the implementation of this in any way, however; for example, the falling edge of the fast clock tick could be used for this purpose.

Given the clock parameters, we will choose τ_w and the size M of *RB* appropriately to ensure the absence of read/write address collisions, and thus metastability-free communication, in the fault-free case. Note that this implies that the very first τ_w data items read, after start-up, are void and must hence be discarded.

Since our solution contains several state-holding components (*RD*, *WD*, and *RB*), which can be corrupted by transient failures, we need to make them self-stabilizing. This is a non-trivial problem, since a corruption of, e.g., *RD* would lead to a persistently

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

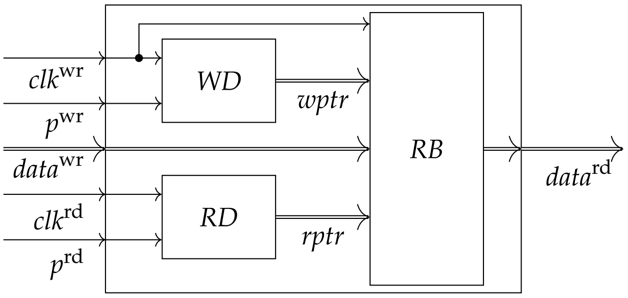


Figure 6.3: Implementation of the communication scheme presented in Figure 6.1. *RB* implements the ring buffer. *RD* and *WD* are the read resp. write decoder.

wrong alignment of the position of the read-out data and the written ones! We are not aware of any existing approach in the literature that solves this problem in a metastability-free way.

Our solution is based on the following simple idea, which does not incur any performance penalty: We choose the period interval I maintained by the underlying clocking system as $I := \mathrm{lcm}(M, B)/B$, where lcm is the least common multiple. This guarantees that, in the absence of transient failures, the read and write decoders point to their initial address every time the period signal is issued along with the first fast clock tick of every I^{th} burst, cf. Figure 6.4. If some counter became corrupted, the period signal corrects it within the next $I \cdot B$ clock ticks at the latest. Albeit the memory may still contain wrongly aligned data by then, at most τ_w clock ticks later the reader will start to read correct data, cf. Figure 6.5. We thus have the following results:

Theorem 6.4 (Self-Stabilization): *If M and τ_w are chosen appropriately to ensure correct communication in the fault-free case, our solution guarantees that correct communication is resumed at most $I \cdot B + \tau_w$ fast clock ticks after the underlying clocking system has stabilized, provided sender and receiver remain correct during the stabilization time. Subsequently, data can be transmitted with every fast clock tick and will be received with a lag of τ_w read events.* $\square$

6.3 Analysis

In the following sections, we will provide a detailed analysis of the minimal memory size required for metastability-free communication for given clock parameters. To ensure metastability-free communication, collision-free memory accesses are required, e.g., the write pointer must be ahead of the read pointer by at least one memory location at any time. Due to the one-to-one correspondence between fast clock ticks and memory addresses, this effectively translates to certain precision requirements, which will be analyzed in this section. They lead to constraints, which must be met in order to guarantee collision-free memory accesses.

We start by defining the address functions corresponding to the progress of the address pointers. In Definition 6.5 below, we will introduce a binary operator $\ominus$ that

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

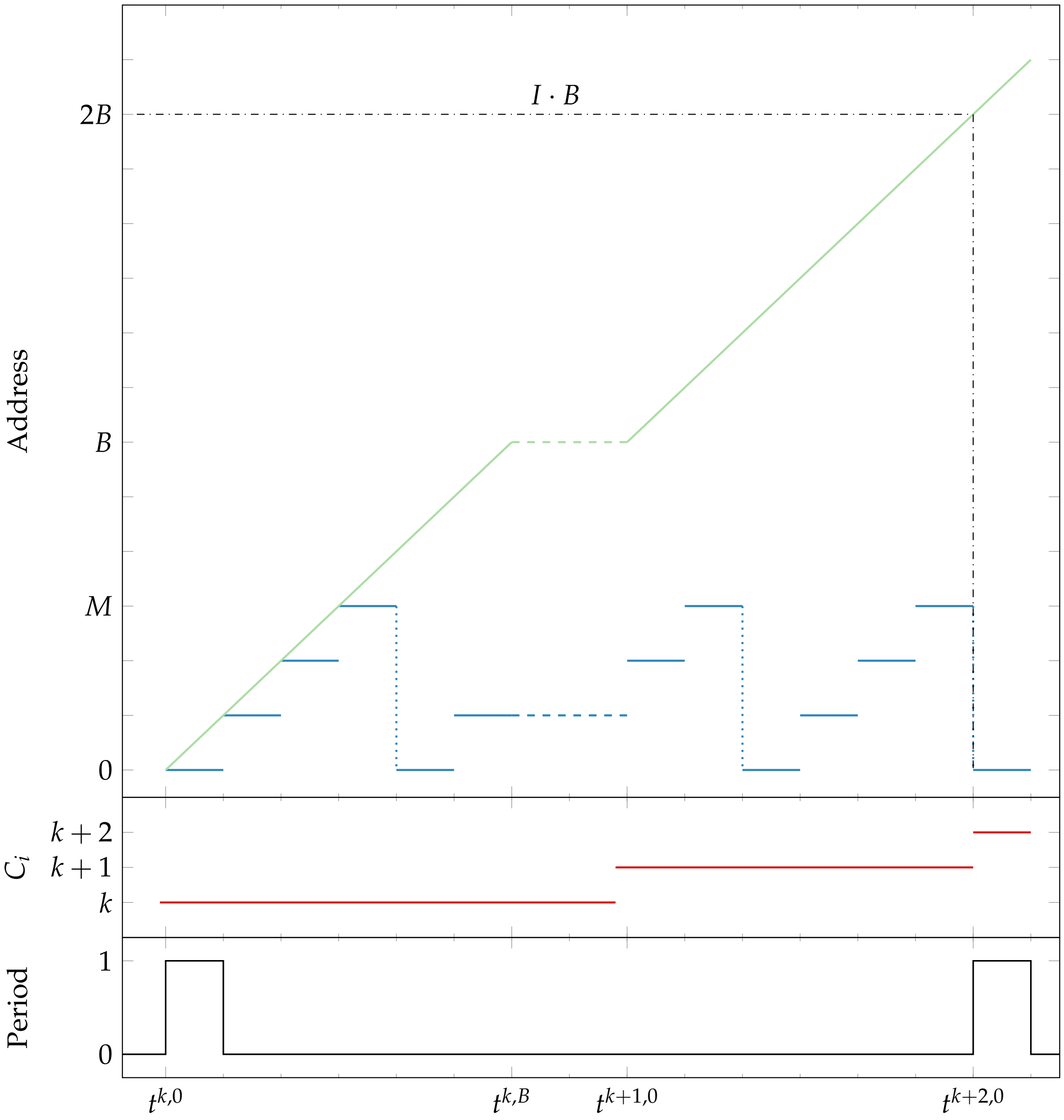


Figure 6.4: A visualization of the effect of a correct dimensioning of I. In this example we have $B = 6$ and $M = 4$, leading to $I = 2$. We see that the period signals occurring at $t^{k,0}$ and $t^{k+2,0}$ correspond with the address pointer (—) pointing to address 0, the initial address. The address function (—) was not floored to improve readability.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

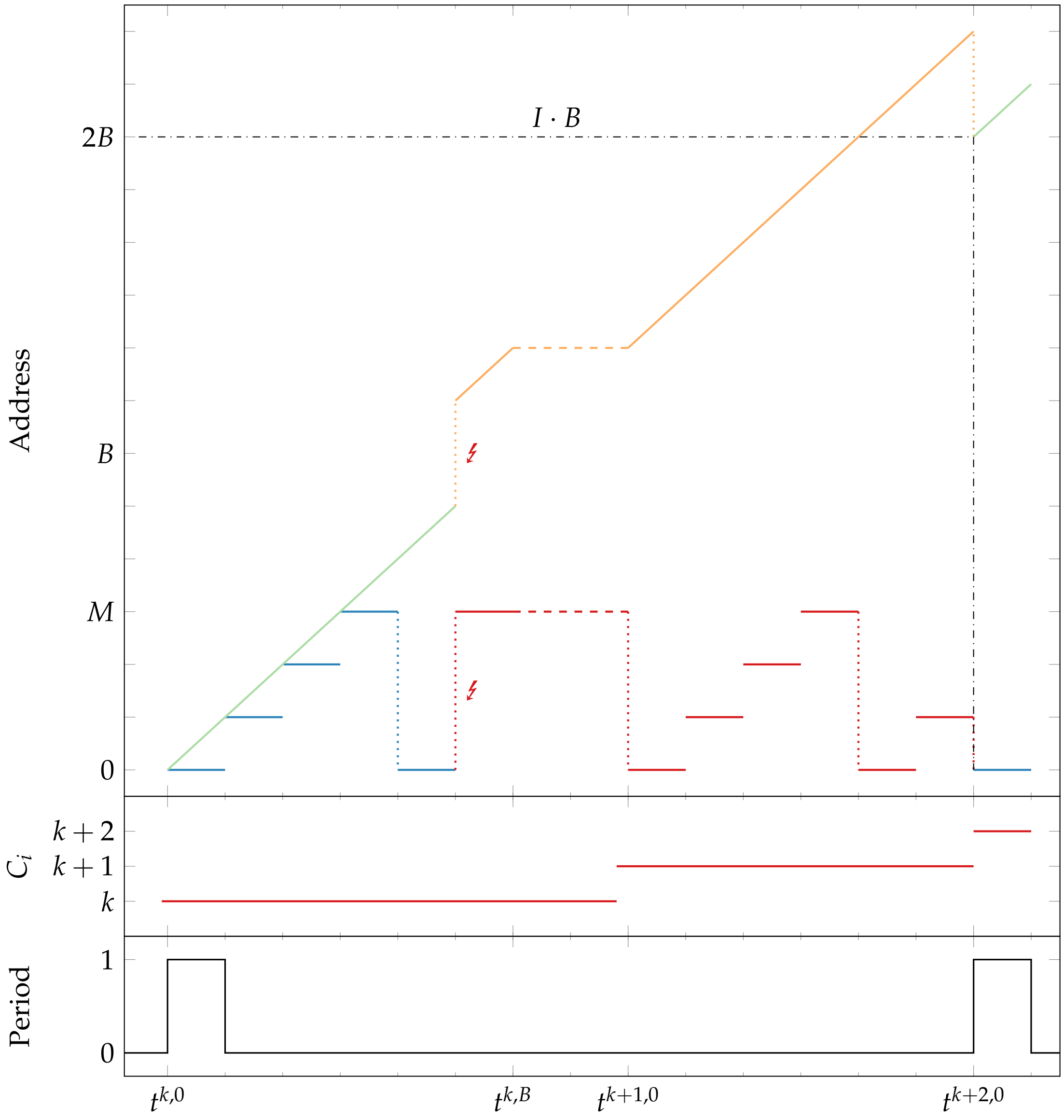


Figure 6.5: The setting of Figure 6.4 in the case where a transient fault occurred at time $t^{k,B-1}$, and corrupted the counter. The 🟥 and 🟧 traces are affected by the fault (marked by ↯). When the period signal is activated at $t^{k+2,0}$, the address pointer is forced to the correct position.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

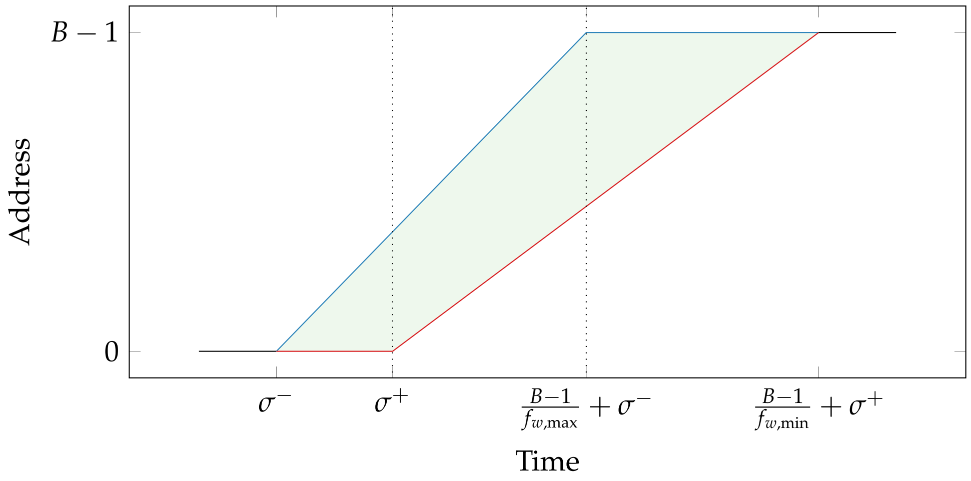


Figure 6.6: The time-cone (🟩) limits the position at which the address functions can be at a given time. Until σ^- both functions stay at 0. At that time, the fastest possible function (—) starts. At time σ^+, the slowest function (—) starts. The fastest possible function stops at value $B-1$ at time $\frac{B-1}{f_{w,\max}}+\sigma^-$. At time $\frac{B-1}{f_{w,\min}}+\sigma^+$, the slowest function also reaches $B-1$ and stays there.

translates the difference of the reader and writer address functions into the difference of the corresponding pointer addresses.

We define an interval $\boldsymbol{\beta}_r(t) := [\beta_r^-(t), \beta_r^+(t)]$ that represents an unbounded time-cone of an address function relating to a single instance $F_r^k(t)$ of the read clock; without loss of generality, we assume $k = 0$ in most parts of this section. As they are B-burst clocks, the interval bounds must be chosen to ensure $F_r^k(t) \in \lfloor \boldsymbol{\beta}_r(t) \rfloor = [\lfloor \beta_r^-(t) \rfloor, \lfloor \beta_r^+(t) \rfloor]$ for all times t. The floor operation accounts for the fact that $F_r^k(t)$ is a right-continuous step function, and $\boldsymbol{\beta}_r(t)$ is defined by:

$$\beta_r^-(t) := \begin{cases} 0 & t < 0 & (6.1a) \\ f_{r,\min} t & t < \frac{B-1}{f_{r,\min}} & (6.1b) \\ B-1 & \text{otherwise} & (6.1c) \end{cases}$$

$$\beta_r^+(t) := \begin{cases} 0 & t < 0 & (6.2a) \\ f_{r,\max} t & t < \frac{B-1}{f_{r,\max}} & (6.2b) \\ B-1 & \text{otherwise} & (6.2c) \end{cases}$$

Note that the clock is not only active in Equations (6.1) and (6.2), but until $\frac{B}{f_{r,\min}}$ resp. $\frac{B}{f_{r,\max}}$, recall Definition 6.2.

For the writer, we analogously define the time-cone $\boldsymbol{\beta}_w(\sigma, t) := [\beta_w^-(\sigma, t), \beta_w^+(\sigma, t)]$ (cf. Figure 6.6) that takes into account the clock skew $\boldsymbol{\sigma}$. It guarantees $F_w^k(t) \in \lfloor \boldsymbol{\beta}_w(\sigma, t) \rfloor$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

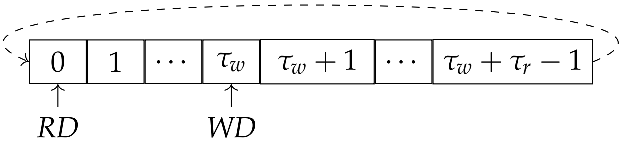


Figure 6.7: The memory layout of the proposed design. Every block represents a memory cell. The initial position of the read- and write-decoder (*RD* resp. *WD*) are marked.

for all times t:

$$\beta_w^-(\sigma, t) := \begin{cases} 0 & t < \sigma^+ \quad (6.3a) \\ f_{w,\min}(t - \sigma^+) & t < \frac{B-1}{f_{w,\min}} + \sigma^+ \quad (6.3b) \\ B - 1 & \text{otherwise} \quad (6.3c) \end{cases}$$

$$\beta_w^+(\sigma, t) := \begin{cases} 0 & t < \sigma^- \quad (6.4a) \\ f_{w,\max}(t - \sigma^-) & t < \frac{B-1}{f_{w,\max}} + \sigma^- \quad (6.4b) \\ B - 1 & \text{otherwise} \quad (6.4c) \end{cases}$$

Since our communication scheme is based on a fixed mapping of clock tick numbers to memory addresses, recall Figure 6.3, there is a trivial relation between the clock precision $\boldsymbol{\pi} = [-\pi_w, \pi_r]$ and the difference of read and write pointers: The read and write pointers are given by $F_r^k(t)$ and $F_w^k(t) + \tau_w$, respectively, where $\tau_w > 0$ denotes the initial value of the write pointer. Given that the initial value of the read pointer is 0, the pointers have an initial offset of τ_w. They are properly contained in the floored time-cones $\delta_r(t) := \lfloor \gamma_r(t) \rfloor$ and $\delta_w(t) := \lfloor \gamma_w(\sigma, t) \rfloor$ defined by:

$$\gamma_r^-(t) := \beta_r^-(t) \tag{6.5}$$

$$\gamma_r^+(t) := \beta_r^+(t) \tag{6.6}$$

$$\gamma_w^-(\sigma, t) := \tau_w + \beta_w^-(\sigma, t) \tag{6.7}$$

$$\gamma_w^+(\sigma, t) := \tau_w + \beta_w^+(\sigma, t) \tag{6.8}$$

$$\delta_r^-(t) := \lfloor \gamma_r^-(t) \rfloor \tag{6.9}$$

$$\delta_r^+(t) := \lfloor \gamma_r^+(t) \rfloor \tag{6.10}$$

$$\delta_w^-(\sigma, t) := \lfloor \gamma_w^-(\sigma, t) \rfloor \tag{6.11}$$

$$\delta_w^+(\sigma, t) := \lfloor \gamma_w^+(\sigma, t) \rfloor \tag{6.12}$$

Now, if the reader and writer clocks maintain $\boldsymbol{\pi} = [-\pi_w, \pi_r]$, then the read and write pointers maintain a difference within $[\tau_w - \pi_w, \tau_w + \pi_r]$. To ensure that the reader is at least one location behind the writer, we hence need $\tau_w > \pi_w$. To ensure that the writer is at least one location behind the reader (modulo the memory size), the size M of the ring buffer must satisfy $M := \tau_r + \tau_w$ for some $\tau_r > \pi_r$ (cf. Figure 6.7).

Note that, implementation-wise, the δ functions should be considered modulo M. For our analysis, however, we keep them unbounded for simplicity. This is without loss

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

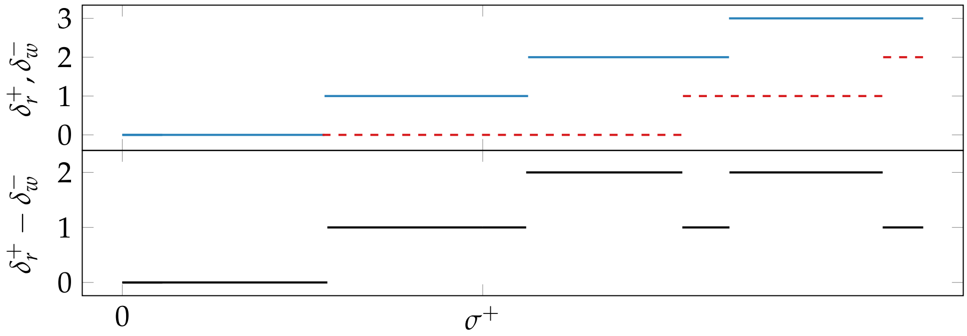


Figure 6.8: In this example we show the issue caused by the non-monotonic behavior of the difference of the δ-functions. We use $B = 200$, $f_{r,\max} = 198$ (δ_r^+, ——), $f_{w,\min} = 200$ (δ_w^-, - - -), and $\sigma^+ = 0.009$ to visualize the issue. This figure focuses on the beginning of the time-frame. As can be seen, the evaluation at a border (σ^+) delivers a difference of $\delta_r^+ - \delta_w^-$ (——) of 1, which increases to 2 shortly thereafter but decreases to 1 later on again.

of generality, as our constraints will ensure that the pointers are never further apart than $M - 1$.

Determining the maximum difference between the reader and writer address function would be trivial if this difference was monotonically increasing. Unfortunately, while δ_r and δ_w have this property, their difference has not, since they are step functions, i.e., involve $\lfloor \cdot \rfloor$. A visualization of this issue is shown in Figure 6.8. To alleviate this issue, we consider the difference of the address functions γ instead of the difference of the pointer functions δ. To determine (a bound on) the difference of the address functions, we introduce the $\ominus$ operator as follows:

Definition 6.5 (Pointer Difference Operator):
Let $\ominus$ be a binary operation, where $a \ominus b$ with variables $a, b \in \mathbb{R}_{\geq 0}$ is defined as $\lceil a - b \rceil$. □

The following lemma establishes its most important property:

Lemma 6.6: *For all times t, it holds for $x, y \in \{w, r\}$ and $m, n \in \{+, -\}$ that:*

$$\delta_x^m(t) - \delta_y^n(t) \leq \gamma_x^m(t) \ominus \gamma_y^n(t) \tag{6.13}$$

□

Proof: We need to prove $\delta_x^m(t) - \delta_y^n(t) = \lfloor a \rfloor - \lfloor b \rfloor \leq a \ominus b = \lceil a - b \rceil = \gamma_x^m(t) \ominus \gamma_y^n(t)$. Using the definition $-\lceil a \rceil = \lfloor -a \rfloor$ and the well-known fact $\lfloor a \rfloor + \lfloor b \rfloor \leq \lfloor a + b \rfloor$ (which is immediately obvious by considering $a = \lfloor a \rfloor + \{a\}$, where $\{a\}$ is the fractional part of a) for $a = x - y \in \mathbb{R}$ and $b = y \in \mathbb{R}$ yields $\lceil y - x \rceil = -\lfloor x - y \rfloor$ and $\lfloor x - y \rfloor \leq \lfloor x \rfloor - \lfloor y \rfloor$. Combining these inequalities proves our lemma. ■

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

As we will evaluate $\ominus$ typically at times when one operand is an integer, we explicitly state the simplified results in such cases, which follow immediately from the definition:

Corollary 6.7: *For specific values of a, b, the following holds:*

$$a \ominus b = \begin{cases} a - b & a, b \in \mathbb{N} & (6.14a) \\ a - \lfloor b \rfloor & a \in \mathbb{N}, b \in \mathbb{R}_{\geq 0} & (6.14b) \\ \lceil a \rceil - b & a \in \mathbb{R}_{\geq 0}, b \in \mathbb{N} & (6.14c) \end{cases}$$

□

After these preparations, we can consider the constraints within a single burst ($k = 0$).

6.4 Single-Burst Memory Parameterization

To ensure metastability-free communication, the address functions must never bump into each other. This entails that A) a fast reader must not catch up with a slow writer, i.e., $\delta_w^-(\sigma, t) > \delta_r^+(t)$, but also that B) a fast writer must not catch up with a slow reader. To abstract away the address wraparound in our analysis, we express the latter condition as $\delta_w^+(\sigma, t) < M + \delta_r^-(t)$.

In the following, we will determine constraints on τ_w and τ_r to ensure that A) and B) are satisfied. Depending on the parameters of the clocks, cf. Lemma 6.3, six different cases $a, \ldots, f$ need to be distinguished: They correspond to the different temporal orders of the four events start and stop of reader and writer, which in turn lead to different interaction of the (floored) time-cones $\delta_r(t)$ and $\delta_w(t)$: Two variants each of partial overlap (a, d), one enveloping the other (b, c), and no overlap (e, f). Each of the six cases is represented by a unique partitioning of the time axis into 5 different regions, see Figure 6.9 for a visualization. We distinguish the outer (= left-most and right-most) and the inner partitions. The sought τ_w (resp. τ_r) must be chosen larger than the maximum value of the difference of reader and writer clock, taken at the same time, within every partition. Note that the difference of the reader and writer clock is inherently zero throughout the outer partitions.

We have shown in Lemma 6.6 that $\ominus$ provides a correct upper bound on the difference between address pointers based on their address functions. The following Lemma 6.8 shows that the latter has a unique maximum at an inner partition boundary.

Lemma 6.8: *For a single burst of reader and writer, $\gamma_x(t) - \gamma_y(t)$ and hence $\gamma_x(t) \ominus \gamma_y(t)$ has a unique maximum and/or minimum, which lies at one of the boundaries between the inner-most partitions.* □

Proof: Since $\gamma_x(t)$ and $\gamma_y(t)$ are continuous, they will assume their minimum and maximum. As $\gamma_x(t) - \gamma_y(t) = 0$ throughout the outer partitions, but not necessarily in the inner partitions, it can reach its minimum and/or maximum only in some inner

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

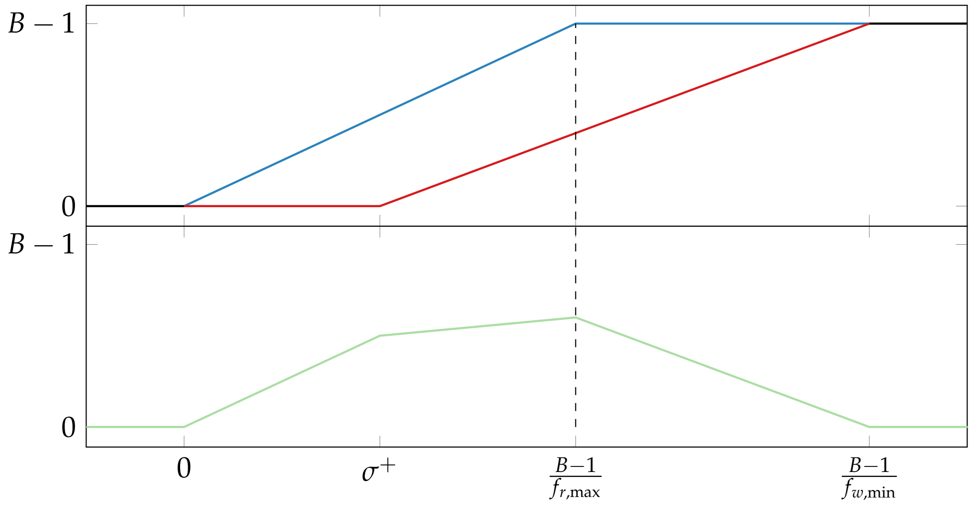


Figure 6.9: Visualization of the case presented in Section 6.4.1.1. To keep the figure clean, we used the address function instead of the pointer functions. The reader (—) starts at time 0 with $f_{r,\max}$. Thus τ_w^a (—) increases until the writer (—) starts at time σ^+. As the writer is slower than the reader, the growth rate of τ_w^a decreased. At time $\frac{B-1}{f_{r,\max}}$ the reader stops, leading to a negative growth of τ_w^a—reaching 0 when the writer stops. The dashed line marks the maximum of τ_w^a.

partition. Since $\gamma_x(t)$ and $\gamma_y(t)$ are piecewise linear, the minimum and maximum of the difference can be assumed at partition boundaries only. As there are only 2 inner boundaries in the single-burst case, the minimum resp. maximum (or both) can be assumed only on one of them.

Since $\lceil x \rceil$ is monotonic, the statement of the lemma also holds for $\gamma_x(t) \ominus \gamma_y(t)$. ∎

We will now derive the conditions that ensure A), i.e., that the reader address function will never catch up with the writer address function, and B), i.e., that the writer address function will never catch up with the reader address function. For every case $x \in \{a, \ldots, f\}$, we will assume that the reader starts at time $t = 0$, and let the writer start (i) with a skew $\sigma \in [\sigma^-, \sigma^+]$ which (ii) maximizes the resulting difference (and hence the safe bound τ_w^x resp. τ_r^x for each case) of the address functions. Obviously, for A), we have to choose $\sigma = \sigma^+$, whereas for B) we need $\sigma = \sigma^-$.

6.4.1 Ensure $\delta_w^-(\sigma, t) > \delta_r^+(t)$, i.e., $\tau_w > \beta_r^+(t) \ominus \beta_w^-(\sigma, t)$

We group the six possible cases according to the sign and magnitude of σ^+ as follows:

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

6.4.1.1 The writer starts during the reader's burst: $0 \leq \sigma^+ \leq \frac{B-1}{f_{r,\max}}$

The reader stops between the writer's start and stop

$$\tau_w^a > \max \begin{cases} 0 \ominus 0 = 0 & t < 0 & \text{(6.15a)} \\ f_{r,\max} t \ominus 0 & t < \sigma^+ & \text{(6.15b)} \\ f_{r,\max} t \ominus f_{w,\min}(t - \sigma^+) & t < \frac{B-1}{f_{r,\max}} & \text{(6.15c)} \\ (B-1) \ominus f_{w,\min}(t - \sigma^+) & t < \frac{B-1}{f_{w,\min}} + \sigma^+ & \text{(6.15d)} \\ (B-1) \ominus (B-1) = 0 & \text{otherwise} & \text{(6.15e)} \end{cases}$$

This case is shown for $f_{r,\max} \geq f_{w,\min}$ in Figure 6.9. By Lemma 6.8, we can reduce the analysis for this case to the center partition (Equation (6.15)). We observe for $t := \sigma^+$ that $\tau_w^{a,1} > \lceil f_{r,\max} \sigma^+ \rceil$ and for $t := \frac{B-1}{f_{r,\max}}$ that $\tau_w^{a,2} > B - 1 - \left\lfloor f_{w,\min} \left(\frac{B-1}{f_{r,\max}} - \sigma^+ \right) \right\rfloor$. We see that the maximum depends on the order of $f_{r,\max}$ to $f_{w,\min}$. Whereas for $f_{r,\max} < f_{w,\min}$

$$\tau_w^{a_1} > \lceil f_{r,\max} \sigma^+ \rceil \tag{6.16}$$

is maximal, for $f_{r,\max} \geq f_{w,\min}$ the maximum is at

$$\tau_w^{a_2} > B - 1 - \left\lfloor f_{w,\min} \left(\frac{B-1}{f_{r,\max}} - \sigma^+ \right) \right\rfloor . \tag{6.17}$$

The reader envelops the writer

$$\tau_w^b > \max \begin{cases} 0 & t < 0 & \text{(6.18a)} \\ f_{r,\max} t \ominus 0 & t < \sigma^+ & \text{(6.18b)} \\ f_{r,\max} t \ominus f_{w,\min}(t - \sigma^+) & t < \frac{B-1}{f_{w,\min}} + \sigma^+ & \text{(6.18c)} \\ f_{r,\max} t \ominus (B-1) & t < \frac{B-1}{f_{r,\max}} & \text{(6.18d)} \\ 0 & \text{otherwise} & \text{(6.18e)} \end{cases}$$

Since the maximum occurs at $t := \sigma^+$, we must choose $\tau_w^b > \lceil f_{r,\max} \sigma^+ \rceil$. Note that this scenario requires $f_{r,\max} < f_{w,\min}$.

6.4.1.2 The writer starts before the reader: $\sigma^+ \leq 0$

The writer envelops the reader

$$\tau_w^c > \max \begin{cases} 0 & t < \sigma^+ & \text{(6.19a)} \\ 0 \ominus f_{w,\min}(t - \sigma^+) & t < 0 & \text{(6.19b)} \\ f_{r,\max} t \ominus f_{w,\min}(t - \sigma^+) & t < \frac{B-1}{f_{r,\max}} & \text{(6.19c)} \\ (B-1) \ominus f_{w,\min}(t - \sigma^+) & t < \frac{B-1}{f_{w,\min}} + \sigma^+ & \text{(6.19d)} \\ 0 & \text{otherwise} & \text{(6.19e)} \end{cases}$$

At $t := \frac{B-1}{f_{r,\max}}$, we observe the maximum (recall that $\sigma^+ \leq 0$). We hence choose $\tau_w^c > B - 1 - \left\lfloor f_{w,\min} \left(\frac{B-1}{f_{r,\max}} - \sigma^+ \right) \right\rfloor$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

The writer stops between the reader's start and stop

$$\tau_w^d > \max \left\{ \begin{array}{lll} 0 & t < \sigma^+ & (6.20a) \\ 0 \ominus f_{w,\min}(t - \sigma^+) & t < 0 & (6.20b) \\ f_{r,\max} t \ominus f_{w,\min}(t - \sigma^+) & t < \frac{B-1}{f_{w,\min}} + \sigma^+ & (6.20c) \\ f_{r,\max} t \ominus (B-1) & t < \frac{B-1}{f_{r,\max}} & (6.20d) \\ 0 & \text{otherwise} & (6.20e) \end{array} \right.$$

Since the writer is always ahead of the reader here, the maximum is 0, so we only have the trivial bound $\tau_w^d > 0$.

The writer stops before the reader starts This trivially requires just $\tau_w^e > 0$.

6.4.1.3 The writer starts after the reader's burst: $0 < \frac{B-1}{f_{r,\max}} < \sigma^+$

$$\tau_w^f > \max \left\{ \begin{array}{lll} 0 & t < 0 & (6.21a) \\ f_{r,\max} t \ominus 0 & t < \frac{B-1}{f_{r,\max}} & (6.21b) \\ (B-1) \ominus 0 & t < \sigma^+ & (6.21c) \\ (B-1) \ominus f_{w,\min}(t - \sigma^+) & t < \frac{B-1}{f_{w,\min}} + \sigma^+ & (6.21d) \\ 0 & \text{otherwise} & (6.21e) \end{array} \right.$$

It is immediately apparent that the maximum is $B-1$ here, so we choose $\tau_w^f > B-1$.

6.4.2 Ensure $\delta_w^+(\sigma, t) < M + \delta_r^-(t)$, i.e., $\tau_r > \beta_w^+(\sigma, t) \ominus \beta_r^-(t)$

We again group the six possible cases as before:

6.4.2.1 The writer starts during the reader's burst: $0 \le \sigma^- \le \frac{B-1}{f_{r,\min}}$

The reader stops between the writer's start and stop

$$\tau_r^a > \max \left\{ \begin{array}{lll} 0 & t < 0 & (6.22a) \\ 0 \ominus f_{r,\min} t & t < \sigma^- & (6.22b) \\ f_{w,\max}(t - \sigma^-) \ominus f_{r,\min} t & t < \frac{B-1}{f_{r,\min}} & (6.22c) \\ f_{w,\max}(t - \sigma^-) \ominus (B-1) & t < \frac{B-1}{f_{w,\max}} + \sigma^- & (6.22d) \\ 0 & \text{otherwise} & (6.22e) \end{array} \right.$$

In this case, the reader is always ahead of the writer, so the maximum is 0. Consequently, $\tau_r^a > 0$ is the only constraint arising here.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

The reader envelops the writer

$$\tau_r^b > \max \begin{cases} 0 & t < 0 & (6.23a) \\ 0 \ominus f_{r,\min} t & t < \sigma^- & (6.23b) \\ f_{w,\max}(t - \sigma^-) \ominus f_{r,\min} t & t < \frac{B-1}{f_{w,\max}} + \sigma^- & (6.23c) \\ (B-1) \ominus f_{r,\min} t & t < \frac{B-1}{f_{r,\min}} & (6.23d) \\ 0 & \text{otherwise} & (6.23e) \end{cases}$$

The maximum occurs in Equation (6.23) at $t := \frac{B-1}{f_{w,\max}} + \sigma^-$. We hence choose $\tau_r^b > B - 1 - \left\lfloor f_{r,\min} \left(\frac{B-1}{f_{w,\max}} - \sigma^- \right) \right\rfloor$.

6.4.2.2 The writer starts before the reader: $\sigma^- \leq 0$

The writer envelops the reader

$$\tau_r^c > \max \begin{cases} 0 & t < \sigma^- & (6.24a) \\ f_{w,\max}(t - \sigma^-) \ominus 0 & t < 0 & (6.24b) \\ f_{w,\max}(t - \sigma^-) \ominus f_{r,\min} t & t < \frac{B-1}{f_{r,\min}} & (6.24c) \\ f_{w,\max}(t - \sigma^-) \ominus (B+1) & t < \frac{B-1}{f_{w,\max}} + \sigma^- & (6.24d) \\ 0 & \text{otherwise} & (6.24e) \end{cases}$$

The maximum is obtained for $t := 0$ here (as $\sigma^- \leq 0$). Note that this scenario requires $f_{w,\max} < f_{r,\min}$. We hence choose $\tau_r^c > \lceil -f_{w,\max}\sigma^- \rceil$.

The writer stops between the reader's start and stop

$$\tau_r^d > \max \begin{cases} 0 & t < \sigma^- & (6.25a) \\ f_{w,\max}(t - \sigma^-) \ominus 0 & t < 0 & (6.25b) \\ f_{w,\max}(t - \sigma^-) \ominus f_{r,\min} t & t < \frac{B-1}{f_{w,\max}} + \sigma^- & (6.25c) \\ (B-1) \ominus f_{r,\min} t & t < \frac{B-1}{f_{r,\min}} & (6.25d) \\ 0 & \text{otherwise} & (6.25e) \end{cases}$$

The maximum occurs at one of the borders of the center partition (Equation (6.25)): For $t = 0$, we find $\tau_r^{d,1} > \lceil -f_{w,\max}\sigma^- \rceil$; for $t := \frac{B-1}{f_{w,\max}} + \sigma^-$, we obtain $\tau_r^{d,2} > B - 1 - \left\lfloor f_{r,\min} \left(\frac{B-1}{f_{w,\max}} + \sigma^- \right) \right\rfloor$. For $f_{w,\max} < f_{r,\min}$, the difference is monotonically decreasing, thus $\tau_r^{d,1} \geq \tau_r^{d,2}$, hence we choose

$$\tau_r^{d_1} > \lceil -f_{w,\max}\sigma^- \rceil. \tag{6.26}$$

Whereas, for $f_{w,\max} \geq f_{r,\min}$ it holds that $\tau_r^{d,2} \geq \tau_r^{d,1}$, hence

$$\tau_r^{d_2} \geq B - 1 - \left\lfloor f_{r,\min} \left(\frac{B-1}{f_{w,\max}} + \sigma^- \right) \right\rfloor. \tag{6.27}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

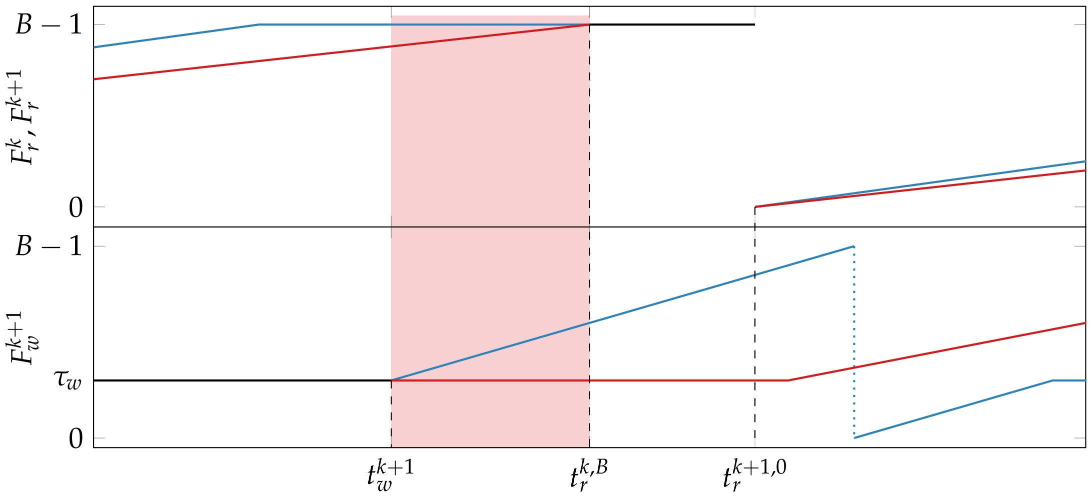


Figure 6.10: Visualization of a scenario which allows a multi-burst collision. The clocks F_r^k and F_r^{k+1} are shown in the same plot (top). For both, reader and writer, we colorize the corresponding $f_{\min}$ with — and $f_{\max}$ with —. Given a small enough σ^-, t_w^{k+1} can happen before $t_r^{k,B}$, causing an overlap during a specific time-interval, shaded red above. Given an incorrectly chosen initial offset τ_w, a collision can occur.

The writer stops before the reader starts

$$\tau_r^e > \max \begin{cases} 0 & t < \sigma^- & \text{(6.28a)} \\ f_{w,\max}(t-\sigma^-) \ominus 0 & t < \frac{B-1}{f_{w,\max}} + \sigma^- & \text{(6.28b)} \\ (B-1) \ominus 0 & t < 0 & \text{(6.28c)} \\ (B-1) \ominus f_{r,\min} t & t < \frac{B-1}{f_{r,\min}} & \text{(6.28d)} \\ 0 & \text{otherwise} & \text{(6.28e)} \end{cases}$$

We immediately obtain a maximum of $B-1$, so we choose $\tau_r^e > B-1$.

6.4.2.3 The reader stops before the writer's burst: $0 < \frac{B-1}{f_{r,\min}} < \sigma^-$

This trivially requires only $\tau_r^f > 0$.

6.5 Multi-Burst Constraints – Writer before Reader

Unfortunately, just considering constraints arising from a single burst are not sufficient for guaranteeing collision-free memory accesses: We also have to consider multi-burst behavior, as, e.g., burst k of the reader may lap into burst $k+1$ of the writer, cf. Figure 6.10. Furthermore, given skews in the order of the pulse separation times,

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

even multiple overlapping burst are possible, e.g., burst k of the reader may lap into burst $k+3$ of the writer. Similar to the single burst constraints, which guarantee a sufficient separation between the pointers of the same burst, we need to find constraints which ensure such a separation between different bursts.

Contrary to the single burst analysis, the following analysis focuses directly on the difference of the address pointers at certain times, rather than the order of the relevant events in time.

In this section, we will consider the case

$$t_w^{k+n,0} < t_r^{k+1,0} \leq t_w^{k+n+1,0}, \text{ for } n \geq 1. \tag{6.29}$$

In the following Section 6.6, the case of $t_r^{k+n,0} < t_w^{k+1,0}$ will be covered, by exploiting some symmetry between reader and writer. Note that our multi-burst formulas do not "degrade" into the single-burst formulas for $n = 0$, as by construction, they require that $n \geq 1$.

The general idea of the analysis provided here is to, first, bind specific time events by *effective skews* towards $t_r^{k,0}$, which is w.l.o.g. set to $t_r^{k,0} = 0$. For this, a case distinction $\Gamma_{w,\min} \leq \Gamma_{r,\min}$ and $\Gamma_{w,\min} \geq \Gamma_{r,\min}$ for the pulse separation times has to be performed. For each case, an upper bound of n can be determined, which in turn allows to determine the maximum difference between writer and reader as follows: We extend the basic requirement for collision free memory access for a single burst ($\tau_r > \beta_w^+(\sigma, t) \ominus \beta_r^-(t)$, cf. Section 6.4.2) by introducing the functions $\beta^+_{w,k+i}\left(\sigma^-_{k+i}, t\right)$ and $\beta^-_{r,k+i}\left(t - \rho^-_{k+i}\right)$ for the $k+i^{\text{th}}$ burst of the writer respectively reader. The constant σ^-_{k+i} represents the effective skew, while ρ^-_{k+i} is a suitably determined constant for the reader. Both are specific settings for each scenario: They are used to appropriately shift the bursts such that the β-functions, which have originally been defined for the k^{th} burst, deliver correct results for the $k+i^{\text{th}}$ burst. This results, for some $t \in \left[t_w^{k+n,0}, t_w^{k+n,B}\right] \cap \left[t_r^{k,0}, t_r^{k,B}\right]$, in

$$\begin{aligned} \tau_r > \left(\sum_{i=k}^{k+n} \beta^+_{w,i}\left(\sigma_i^-, t\right)\right) \ominus \left(\sum_{i=k}^{k+n} \beta^-_{r,i}\left(t - \rho_i^-\right)\right) = \\ = n \cdot B + \beta^+_{w,k+n}(\sigma^-_{k+n}, t) \ominus \beta^-_{r,k}(t - \rho^-_k). \end{aligned} \tag{6.30}$$

The goal is to find the largest bound for τ_r for a given system parameterization. It easy to see that the maximum occurs when the writer bursts are as early (close to $-\infty$) as possible, with maximum absolute skew towards the corresponding reader bursts.

The maximum number of overlapping bursts n is limited by the skew bound, which must hold for every burst-pair. More specifically, $t_r^k + \sigma^- \leq t_w^k$ and $t_w^{k+n} \leq t_r^{k+n} + \sigma^+$ limit the outer shell in which those overlaps can occur. In the case considered here, $t_w^{k+n,0} < t_r^{k+1,0}$ for $n \geq 1$, we only need to care about the σ^- bound, as this causes the maximal distance between writer and reader.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

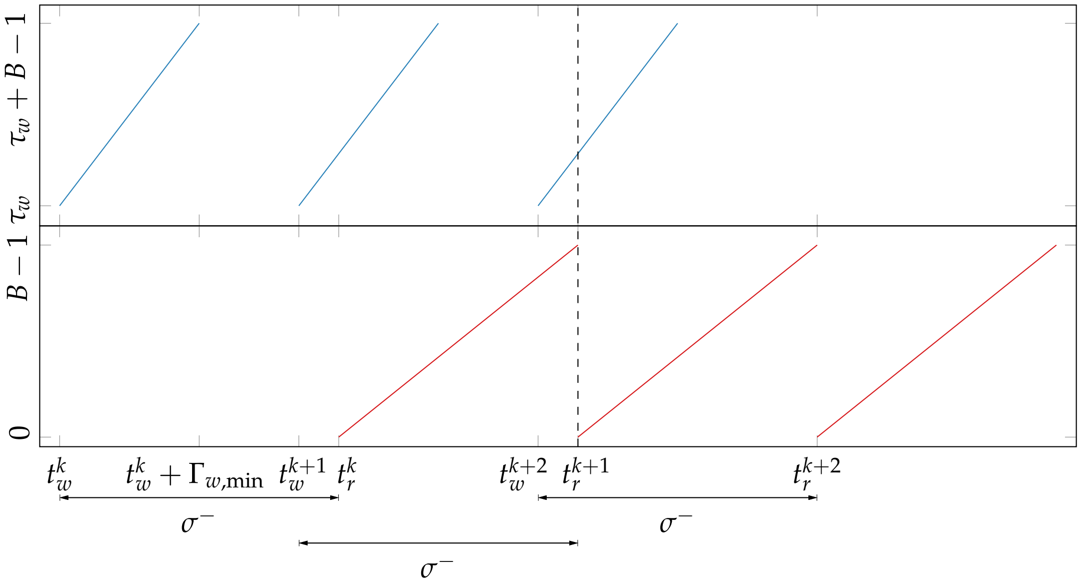


Figure 6.11: A setting where a single burst of the reader (—) overlaps with multiple bursts of the writer (—) with $f_{w,\max} > f_{r,\min}$ and $\Gamma_{r,\min} > \Gamma_{w,\min}$. To avoid clutter, the address wraparound has been omitted. The relevant points in time for the determination of n, t_r^{k+1}, is marked with a dashed line. This type of setting requires that σ^- is negative, while the number of overlaps also depends on the size of $\Gamma_{r,\min}$. Note that the defining parameters here are $\Gamma_{r,\min}$, $\Gamma_{w,\min}$ and σ^-. The frequencies only influence the minimum size of the pulse separation times, but not the construction of the worst-case.

6.5.1 Determination of the effective skews and n in the case of $\Gamma_{w,\min} \leq \Gamma_{r,\min}$

The construction of the worst-case relies on the larger pulse separation time of the reader. As we seek to push the writer as far towards $-\infty$ as possible, reader and writer bursts will be separated by σ^-, see Figure 6.11. From this, it also follows that $t_r^k + n \cdot \Gamma_{r,\min} = t_r^{k+n}$ and from that $t_w^k + n \cdot \Gamma_{r,\min} = t_w^{k+n}$. According to the definition of a multi-burst overlap, a bound for n follows from

$$t_r^{k+1} + (n-1) \cdot \Gamma_{r,\min} + \sigma^- = t_r^{k+n} + \sigma^- \leq t_w^{k+n} < t_r^{k+1} \tag{6.31}$$

$$n < \frac{\Gamma_{r,\min} - \sigma^-}{\Gamma_{r,\min}} \tag{6.32}$$

As n has to be an integer, we can define the maximum of n in this case as

$$n = \left\lceil \frac{\Gamma_{r,\min} - \sigma^-}{\Gamma_{r,\min}} \right\rceil - 1 \tag{6.33}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

To evaluate Equation (6.30) at times relating to events of the writer, we need to bind these times towards the only fixed time: $t_r^{k,0}$. This is done by constructing effective skews between these times and $t_r^{k,0}$. As we will see in Section 6.5.3, the relevant points in time happen during the writers $k+n-1^{\text{st}}$ and $k+n^{\text{th}}$ burst. Hence, we need to find the effective skew bounds $\sigma_{k+n-1}^{r,-}$ and $\sigma_{k+n}^{r,-}$ for the $k+n-1^{\text{st}}$ respectively $k+n^{\text{th}}$ writer burst towards the k^{th} reader. Due to the construction of the worst-case, we immediately see that these effective skews can be constructed via the corresponding reader burst and the skew.

$$t_r^k + \sigma_{k+n}^{r,-} \leq t_w^{k+n}$$

$$\sigma_{k+n}^{r,-} \leq t_w^{k+n} - t_r^k = t_w^{k+n} - t_r^{k+n} + n \cdot \Gamma_{r,\min}$$

$$\sigma_{k+n}^{r,-} \leq \sigma^- + n \cdot \Gamma_{r,\min} \tag{6.34}$$

$$t_r^k + \sigma_{k+n-1}^{r,-} \leq t_w^{k+n-1}$$

$$\sigma_{k+n-1}^{r,-} \leq \sigma^- + (n-1) \cdot \Gamma_{r,\min} \tag{6.35}$$

Furthermore, $\rho_{k+i}^{r,-} = i \cdot \Gamma_{r,\min}$ follows from $t_r^k + n \cdot \Gamma_{r,\min} = t_r^{k+n}$; note that $\rho_k^{r,-} = 0$ here.

6.5.2 Determination of the effective skews and n in the case of $\Gamma_{w,\min} \geq \Gamma_{r,\min}$

The construction of the worst-case is analogous to Section 6.5.1, see Figure 6.12: We maximize the distance between corresponding reader and writer bursts to σ^-, leading to $t_w^k + n \cdot \Gamma_{w,\min} = t_w^{k+n}$ and from that $t_r^k + n \cdot \Gamma_{w,\min} = t_r^{k+n}$. By the definition of a multi-burst overlap, a bound for n can be computed from

$$t_r^{k+1} + (n-1) \cdot \Gamma_{w,\min} + \sigma^- = t_r^{k+n} + \sigma^- \leq t_w^{k+n} < t_r^{k+1} \tag{6.36}$$

$$n = \left\lceil \frac{\Gamma_{w,\min} - \sigma^-}{\Gamma_{w,\min}} \right\rceil - 1 \tag{6.37}$$

Due to the evaluations during bursts of the writer, we again need the effective skews $\sigma_{k+n}^{w,-}$ and $\sigma_{k+n-1}^{w,-}$ of the $k+n-1^{\text{st}}$ respectively $k+n^{\text{th}}$ writer burst towards the k^{th} reader.

$$t_r^k + \sigma_{k+n}^{w,-} \leq t_w^{k+n} = t_w^k + n \cdot \Gamma_{w,\min}$$

$$\sigma_{k+n}^{w,-} \leq \sigma^- + n \cdot \Gamma_{w,\min} \tag{6.38}$$

$$t_r^k + \sigma_{k+n-1}^{w,-} \leq t_w^{k+n-1}$$

$$\sigma_{k+n-1}^{w,-} \leq \sigma^- + (n-1) \cdot \Gamma_{w,\min} \tag{6.39}$$

As $\rho^{w,-}$ is the time between two reader bursts in this case, we can just set $\rho_{k+i}^{w,-} = i \cdot \Gamma_{w,\min}$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

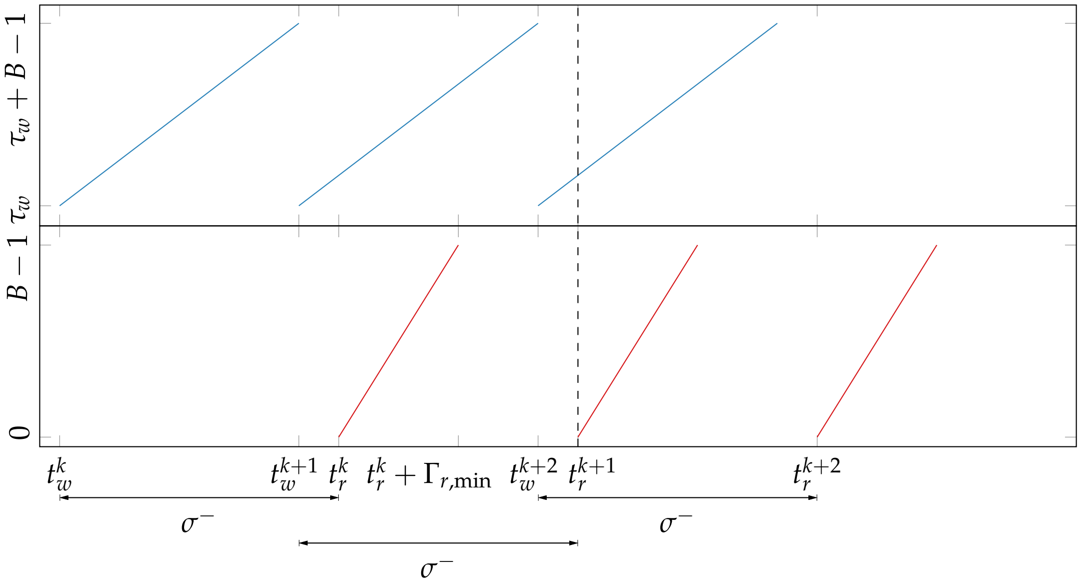


Figure 6.12: A setting where a single burst of the reader (——) overlaps with multiple bursts of the writer (——) with $f_{w,\text{max}} < f_{r,\text{min}}$. To avoid clutter, the address wraparound has been omitted. The relevant points in time for the determination of n, t_r^{k+1}, is marked with a dashed line. This type of setting requires that σ^- is negative, while the number of overlaps also depends on the size of $\Gamma_{r,\text{min}}$. Note that the defining parameters here are $\Gamma_{r,\text{min}}$, $\Gamma_{w,\text{min}}$ and σ^-. The frequencies only influence the minimum size of the pulse separation times, but not the construction of the worst-case.

6.5.3 Bounding τ_r for $\Gamma_{w,\text{min}} \leq \Gamma_{r,\text{min}}$

Using the bound on n given by Equation (6.33), we can focus on finding an upper bound for τ_r. While we already have the general bound with Equation (6.30), this formula only considers the overlap between the $k+n^{\text{th}}$ writer with the k^{th} reader. In some cases, however, we also have to consider the overlap with, e.g., the $k+1^{\text{st}}$ reader.

The relevant points in time, e.g., $t_w^{k+n,0}$, depend on $\Gamma_{r,\text{min}}$. Yet, the decision at which points in time the evaluation must be made depends on where the maximum occurs, and thus, in general, on the relation between $f_{r,\text{min}}$ and $f_{w,\text{max}}$: If the reader is faster than the writer, the difference will decrease during an overlap where both are active, otherwise it will increase.

6.5.3.1 $f_{w,\text{max}} > f_{r,\text{min}}$

In this setting, the writer can increase the difference with every tick, thus we need to evaluate at the latest time: $t_w^{k+n,B}$. Note that, even though this tick may not overlap with the k^{th} reader, it will be the point of the maximum difference towards the k^{th} reader.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

When determining the position of $t_w^{k+n,B}$, we first construct the lower bound of $t_r^{k,0} < t_w^{k+n,B}$. This follows from the case-specific constraints $t_w^{k+n,0} < t_r^{k+1} \leq t_w^{k+n+1,0}$, which implies $t_r^{k,0} \leq t_w^{k+n,0}$. In the following, we consider the three position where $t_w^{k+n,B}$ can be relative to the reader's burst.

Case a: $t_r^{k,0} < t_w^{k+n,B} < t_r^{k,B}$ This setting can be considered the general case, and is straightforward:

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_w^{k+n,B}\right)\right) \ominus \beta^-_{r,k}\left(t_w^{k+n,B} - \rho_k^{r,-}\right) = \\ &= (n+1)\cdot B - \left\lfloor \beta^-_{r,k}\left(t_w^{k+n,B} - \rho_k^{r,-}\right)\right\rfloor. \end{aligned} \tag{6.40}$$

By Equation (6.34) we have for $t_w^{k+n,B} = \sigma^- + n \cdot \Gamma_{r,\min} + \frac{B}{f_{w,\max}}$. As $\rho_k^{r,-} = 0$, we end up with

$$\tau_r > (n+1)\cdot B - \left\lfloor f_{r,\min}\left(\sigma^- + n\cdot\Gamma_{r,\min} + \frac{B}{f_{w,\max}}\right)\right\rfloor. \tag{6.41}$$

Case b: $t_r^{k,B} \leq t_w^{k+n,B} < t_r^{k+1,0}$ This setting is only possible if the minimum pulse separation time of the reader is not tight. The difference calculation is simple, as both reader and writer are at the end of their respective burst:

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_w^{k+n,B}\right)\right) \ominus \beta^-_{r,k}\left(t_w^{k+n,B} - \rho_k^{r,-}\right) = \\ &= n\cdot B. \end{aligned} \tag{6.42}$$

Case c: $t_r^{k+1,0} \leq t_w^{k+n,B} < t_r^{k+1,B}$ Here, we have to consider the $k+1^{\text{st}}$ reader, leading to:

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_w^{k+n,B}\right)\right) \ominus \left(\sum_{i=k}^{k+1} \beta^-_{r,i}\left(t_w^{k+n,B} - \rho_i^{r,-}\right)\right) = \\ &= (n+1)\cdot B - \left(B + \left\lfloor \beta^-_{r,k+1}\left(t_w^{k+n,B} - \rho_{k+1}^{r,-}\right)\right\rfloor\right) = \\ &= n\cdot B - \left\lfloor f_{r,\min}\left(\sigma^- + (n-1)\cdot\Gamma_{r,\min} + \frac{B}{f_{w,\max}}\right)\right\rfloor. \end{aligned} \tag{6.43}$$

Later times do not need to be considered, as with $t_w^{k+n,0} < t_r^{k+1,0}$ and $f_{w,\max} > f_{r,\min}$ it holds that $t_w^{k+n,B} < t_r^{k+1,B}$.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

6.5.3.2 $f_{w,\max} \leq f_{r,\min}$

In this setting, the reader will decrease the difference with every tick. Sadly, it is not clear for every case when the maximum difference occurs. Due to the frequency settings, the $k+n-1^{\text{st}}$ writer burst may lap into the active part of k^{th} burst of reader while the $k+n^{\text{th}}$ writer burst does not. We have to consider three case.

Case a: $t_w^{k+n,0} < t_r^{k,0}$ This case cannot happen, as $t_w^{k+n,0} + \Gamma_{r,\min} = t_w^{k+n+1,0}$ and $t_r^{k,0} + \Gamma_{r,\min} = t_r^{k+1,0}$, we would have $t_w^{k+n+1,0} < t_r^{k+1,0}$, which violates the assumption of $t_w^{k+n,0} < t_r^{k+1,0} \leq t_w^{k+n+1,0}$.

Case b: $t_r^{k,0} \leq t_w^{k+n,0} < t_r^{k,B}$ Here, we have to consider an overlap of the $k+n-1^{\text{st}}$ writer with the k^{th} reader as well, as $f_{w,\max} \leq f_{r,\min}$.

Hence, if $t_r^{k,0} \leq t_w^{k+n-1,0}$ we evaluate at $t_w^{k+n-1,0}$:

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n-1} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_w^{k+n-1,0}\right)\right) \ominus \beta^-_{r,k}\left(t_w^{k+n-1,0} - \rho_k^{r,-}\right) = \\ &= (n-1)\cdot B - \left\lfloor \beta^-_{r,k}\left(t_w^{k+n-1,0} - \rho_k^{r,-}\right)\right\rfloor = \\ &= (n-1)\cdot B - \left\lfloor f_{r,\min}\left(\sigma^- + (n-1)\cdot\Gamma_{r,\min}\right)\right\rfloor .\end{aligned} \tag{6.44}$$

Note that the contribution from the writer's $i = k+n-1^{\text{st}}$ burst is 0.

If $t_w^{k+n-1,0} < t_r^{k,0} \leq t_w^{k+n-1,B}$, we evaluate at $t_r^{k,0}$ and $t_w^{k+n-1,B}$:

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n-1} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_r^{k,0}\right)\right) \ominus \beta^-_{r,k}\left(t_r^{k,0} - \rho_k^{r,-}\right) = \\ &= (n-1)\cdot B + \left\lceil \beta^+_{w,k+n-1}\left(\sigma_{k+n-1}^{r,-}, t_r^{k,0}\right)\right\rceil - 0 = \\ &= (n-1)\cdot B + \left\lceil f_{w,\max}\left(-\sigma^- - (n-1)\cdot\Gamma_{r,\min}\right)\right\rceil ,\end{aligned} \tag{6.45}$$

where we again used Equation (6.34). For $t_w^{k+n-1,B}$, we get

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n-1} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_w^{k+n-1,B}\right)\right) \ominus \beta^-_{r,k}\left(t_w^{k+n-1,B} - \rho_k^{r,-}\right) = \\ &= n\cdot B - \left\lfloor \beta^-_{r,k}\left(t_w^{k+n-1,B} - \rho_k^{r,-}\right)\right\rfloor = \\ &= n\cdot B - \left\lfloor f_{r,\min}\left(\sigma^- + (n-1)\cdot\Gamma_{r,\min} + \frac{B}{f_{w,\max}}\right)\right\rfloor .\end{aligned} \tag{6.46}$$

In case of $t_w^{k+n-1,B} < t_r^{k,0}$, we also evaluate at $t_r^{k,0}$:

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n-1} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_r^{k,0}\right)\right) \ominus \beta^-_{r,k}\left(t_r^{k,0} - \rho_k^{r,-}\right) = \\ &= n\cdot B.\end{aligned} \tag{6.47}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

An evaluation at $t_w^{k+n,0}$ is not required, as by $f_{w,\max} \le f_{r,\min}$ the above setting leads to a higher bound.

However, we must also consider the end of the $k+n^{\text{th}}$ writer burst relatively to the start of the $k+1^{\text{st}}$ reader. The setting $t_w^{k+n,B} < t_r^{k,B}$ cannot occur, as $f_{w,\max} \le f_{r,\min}$ implies $t_w^{k+n,0} < t_r^{k,0}$, which would contradict the scenario handled here $\left(t_r^{k,0} \le t_w^{k+n,0}\right)$. One possible setting is $t_r^{k,B} \le t_w^{k+n,B} < t_r^{k+1,0}$, where we evaluate at $t_w^{k+n,B}$:

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_w^{k+n,B}\right)\right) \ominus \beta^-_{r,k}\left(t_w^{k+n,B} - \rho_k^{r,-}\right) = \\ &= (n+1)\cdot B - B = n \cdot B. \end{aligned} \tag{6.48}$$

The only remaining possibility is $t_r^{k+1,0} \le t_w^{k+n,B}$, where we evaluate at $t_r^{k+1,0}$:

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_r^{k+1,0}\right)\right) \ominus \left(\sum_{i=k}^{k+1} \beta^-_{r,i}\left(t_r^{k+1,0} - \rho_i^{r,-}\right)\right) = \\ &= n \cdot B + \left\lceil \beta^+_{w,k+n}\left(\sigma^{r,-}_{k+n}, t_r^{k+1,0}\right)\right\rceil - B = \\ &= (n-1)\cdot B + \left\lceil f_{w,\max}\left(-\sigma^- - (n-1)\cdot\Gamma_{r,\min}\right)\right\rceil. \end{aligned} \tag{6.49}$$

Fortunately, there is no need to explicitly handle the $k+n+1^{\text{st}}$ writer burst relatively to $k+1^{\text{st}}$ reader burst: The evaluation at $t_w^{k+n+1,0}$ always leads to a smaller bound than in the above settings, due to the frequency setting.

Summarizing the results of this case, we see that Equation (6.44) is always smaller than the other bounds. Further, Equations (6.45) and (6.47) cover the same setting, shifted by $\Gamma_{r,\min}$, as Equations (6.48) and (6.49) leading to the same bounds. Hence, one pair does not need to be considered explicitly.

Case c: $t_r^{k,B} \le t_w^{k+n,0} < t_r^{k+1,0}$ We again have to consider an overlap of the $k+n-1^{\text{st}}$ writer with the reader. For $t_w^{k+n-1,0} < t_r^{k,0} \le t_w^{k+n-1,B}$, we evaluate at $t_r^{k,0}$:

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n-1} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_r^{k,0}\right)\right) \ominus \beta^-_{r,k}\left(t_r^{k,0} - \rho_k^{r,-}\right) = \\ &= (n-1)\cdot B + \left\lceil \beta^+_{w,k+n-1}\left(\sigma^{r,-}_{k+n-1}, t_r^{k,0}\right)\right\rceil = \\ &= (n-1)\cdot B + \left\lceil f_{w,\max}(-\sigma^- - (n-1)\cdot\Gamma_{r,\min})\right\rceil. \end{aligned} \tag{6.50}$$

For $t_r^{k,0} \le t_w^{k+n-1,0} \le t_r^{k,B}$, the evaluation happens at $t_w^{k+n-1,0}$:

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n-1} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_w^{k+n-1,0}\right)\right) \ominus \beta^-_{r,k}\left(t_w^{k+n-1,0} - \rho_k^{r,-}\right) = \\ &= (n-1)\cdot B - \left\lfloor \beta^-_{r,k}\left(t_w^{k+n-1,0} - \rho_k^{r,-}\right)\right\rfloor = \\ &= (n-1)\cdot B - \left\lfloor f_{r,\min}(\sigma^- + (n-1)\cdot\Gamma_{r,\min})\right\rfloor. \end{aligned} \tag{6.51}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

For $t_r^{k,B} \leq t_w^{k+n-1,B}$, we must evaluate at $t_w^{k+n-1,B}$, which leads to

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n-1} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_w^{k+n-1,B}\right)\right) \ominus \beta^-_{r,k}\left(t_w^{k+n-1,B} - \rho_k^{r,-}\right) = \\ &= n \cdot B - B = (n-1) \cdot B. \end{aligned} \tag{6.52}$$

As the k^{th} reader burst has stopped already in this setting, we have to distinguish whether the writer has finished before or after the start of the next reader burst. If $t_w^{k+n,B} < t_r^{k+1,0}$, we have

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_w^{k+n,B}\right)\right) \ominus \beta^-_{r,k}\left(t_w^{k+n,B} - \rho_k^{r,-}\right) = \\ &= (n+1) \cdot B - B = n \cdot B, \end{aligned} \tag{6.53}$$

otherwise

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n} \beta^+_{w,i}\left(\sigma_i^{r,-}, t_r^{k+1,0}\right)\right) \ominus \left(\sum_{i=k}^{k+1} \beta^-_{r,i}\left(t_r^{k+1,0} - \rho_i^{r,-}\right)\right) = \\ &= n \cdot B + \left\lceil \beta^+_{w,k+n}\left(\sigma_{k+n}^{r,-}, t_r^{k+1,0}\right)\right\rceil - B = \\ &= (n-1) \cdot B + \left\lceil f_{w,\max}(-\sigma^- - (n-1) \cdot \Gamma_{r,\min})\right\rceil . \end{aligned} \tag{6.54}$$

Note that these two bounds dominate the evaluations at $t_w^{k+n-1,0}$ and $t_w^{k+n-1,B}$. Moreover, the evaluation at $t_r^{k,0}$ also leads to a bound that is smaller or equal than the two above. Hence, we do not need to consider the bounds of Equations (6.50) to (6.52) explicitly.

Finally, we note that later points in time need not be considered due to $f_{w,\max} \leq f_{r,\min}$.

6.5.4 Bounding τ_r for $\Gamma_{w,\min} \geq \Gamma_{r,\min}$

6.5.4.1 $f_{w,\max} > f_{r,\min}$

In this setting, the writer will increase the difference with every tick, thus we evaluate at the latest time: $t_w^{k+n,B}$. We have to consider 4 cases, which differ w.r.t. the temporal position of the writer burst used for calculating the difference:

Case a: $t_w^{k+n,B} \leq t_r^{k,0}$ Since $t_w^{k+n,B} \leq t_r^{k,0}$ implies $t_w^{k+n+1,B} \leq t_r^{k+1,0}$ here, this case would violate Equation (6.29), i.e., the basic condition $t_w^{k+n,0} < t_r^{k+1,0} \leq t_w^{k+n+1,0}$, and is hence irrelevant.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Case b: $t_r^{k,0} < t_w^{k+n,B} < t_r^{k,B}$ This setting can be considered the general case, and is straightforward:

$$\begin{aligned}\tau_r > &\left(\sum_{i=k}^{k+n} \beta_{w,i}^{+}\left(\sigma_i^{w,-}, t_w^{k+n,B}\right)\right) \ominus \beta_{r,k}^{-}\left(t_w^{k+n,B} - \rho_k^{w,-}\right) = \\ &= (n+1) \cdot B - \left\lfloor \beta_{r,k}^{-}\left(t_w^{k+n,B} - \rho_k^{w,-}\right)\right\rfloor = \\ &= (n+1) \cdot B - \left\lfloor f_{r,\min}\left(\sigma^{-} + n \cdot \Gamma_{w,\min} + \frac{B}{f_{w,\max}}\right)\right\rfloor . \end{aligned} \tag{6.55}$$

Case c: $t_r^{k,B} \leq t_w^{k+n,B} < t_r^{k+1,0}$ Due to $f_{w,\max} > f_{r,\min}$, we also have to consider the difference w.r.t. the $k+n-1$st writer in this case. If $t_w^{k+n-1,B} \leq t_r^{k,0}$, we have

$$\begin{aligned}\tau_r > &\left(\sum_{i=k}^{k+n-1} \beta_{w,i}^{+}\left(\sigma_i^{w,-}, t_w^{k+n-1,B}\right)\right) \ominus \beta_{r,k}^{-}\left(t_w^{k+n-1,B} - \rho_k^{w,-}\right) = \\ &= n \cdot B - 0, \end{aligned} \tag{6.56}$$

if $t_r^{k,0} < t_w^{k+n-1,B} \leq t_r^{k,B}$, we observe

$$\begin{aligned}\tau_r > &\left(\sum_{i=k}^{k+n-1} \beta_{w,i}^{+}\left(\sigma_i^{w,-}, t_w^{k+n-1,B}\right)\right) \ominus \beta_{r,k}^{-}\left(t_w^{k+n-1,B} - \rho_k^{w,-}\right) = \\ &= n \cdot B - \left\lfloor \beta_{r,k}^{-}\left(t_w^{k+n-1,B} - \rho_k^{w,-}\right)\right\rfloor = \\ &= n \cdot B - \left\lfloor f_{r,\min}\left(\sigma^{-} + (n-1) \cdot \Gamma_{w,\min} + \frac{B}{f_{w,\max}}\right)\right\rfloor . \end{aligned} \tag{6.57}$$

Finally, $t_w^{k+n-1,B}$ cannot be after $t_r^{k,B}$, as otherwise $t_r^{k+1,0} \leq t_w^{k+n,0}$ would hold, which contradicts Equation (6.29) as in Case a above.

For the main case, as the kth reader is at the end of its burst, we have

$$\begin{aligned}\tau_r > &\left(\sum_{i=k}^{k+n} \beta_{w,i}^{+}\left(\sigma_i^{w,-}, t_w^{k+n,B}\right)\right) \ominus \beta_{r,k}^{-}\left(t_w^{k+n,B} - \rho_k^{w,-}\right) = \\ &= (n+1) \cdot B - B = n \cdot B. \end{aligned} \tag{6.58}$$

As we can see, the bound given by the $k+n$th writer dominates the bounds for the $k+n-1$st writer, thus only the latter bound need to be considered.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Case d: $t_r^{k+1,0} \leq t_w^{k+n,B} < t_r^{k+1,B}$ Here, the main case involving the $k+1^{\text{st}}$ reader leads to

$$\begin{aligned}
\tau_r &> \left(\sum_{i=k}^{k+n} \beta^+_{w,i}\left(\sigma_i^{w,-}, t_w^{k+n,B}\right)\right) \ominus \left(\sum_{i=k}^{k+1} \beta^-_{r,i}\left(t_w^{k+n,B} - \rho_i^{w,-}\right)\right) = \\
&= (n+1) \cdot B - \left(B + \left\lfloor \beta^-_{r,k+1}\left(t_w^{k+n,B} - \rho_{k+1}^{w,-}\right)\right\rfloor\right) = \\
&= n \cdot B - \left\lfloor f_{r,\min}\left(\sigma^- + (n-1) \cdot \Gamma_{w,\min} + \frac{B}{f_{w,\max}}\right)\right\rfloor . \qquad (6.59)
\end{aligned}$$

Like in Case c, we also need to consider the possible overlap of the $k+n-1^{\text{st}}$ writer with the k^{th} reader. We can ignore the subcase $t_w^{k+n-1,B} \leq t_r^{k,0}$ here, since the separation $\Gamma_{w,\min}$ between the burst of reader and writer would imply $t_w^{k+n,B} < t_r^{k+1,0}$, contradicting the scenario handled here $\left(t_r^{k+1,0} \leq t_w^{k+n,B} \leq t_r^{k+1,B}\right)$. For $t_r^{k,0} < t_w^{k+n-1,B} \leq t_r^{k,B}$, we observe exactly Equation (6.57). Note that this bound is the same as in the main case.

As $t_w^{k+n,0} < t_r^{k+1,0} \leq t_w^{k+n+1,0}$ and $f_{w,\max} > f_{r,\min}$, it follows that $t_w^{k+n,B} < t_r^{k+1,B}$. Hence, we do not need to consider the remaining case $t_w^{k+n,B} \geq t_r^{k+1,B}$ at all.

6.5.4.2 $f_{w,\max} \leq f_{r,\min}$

In this setting, the reader will decrease the difference with every tick. Unfortunately, the case where the maximum difference occurs here depends on the actual setting of the parameters.

We have to consider three different cases here:

Case a: $t_w^{k+n,0} < t_r^{k,0}$ As this would imply $t_w^{k+n+1,0} < t_r^{k+1,0}$, contradicting Equation (6.29), this case is irrelevant.

Case b: $t_r^{k,0} \leq t_w^{k+n,0} \leq t_r^{k,B}$ First, we consider the possible overlap with the $k+n-1^{\text{st}}$ writer. If $t_r^{k,0} \leq t_w^{k+n-1,B}$, we evaluate at $t_r^{k,0}$ and $t_w^{k+n-1,B}$: For the former, recalling that $f_{w,\max} \leq f_{r,\min}$ implies $t_w^{k+n-1,0} < t_r^{k,0}$ providing

$$\begin{aligned}
\tau_r &> \left(\sum_{i=k}^{k+n-1} \beta^+_{w,i}\left(\sigma_i^{w,-}, t_r^{k,0}\right)\right) \ominus \beta^-_{r,k}\left(t_r^{k,0} - \rho_k^{w,-}\right) = \\
&= (n-1) \cdot B + \left\lceil \beta^+_{w,k+n}\left(\sigma_{k+n-1}^{w,-}, t_r^{k,0}\right)\right\rceil - 0 = \\
&= (n-1) \cdot B + \left\lceil f_{w,\max}\left(-\sigma^- - (n-1) \cdot \Gamma_{w,\min}\right)\right\rceil . \qquad (6.60)
\end{aligned}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

For the latter, we get

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n-1} \beta^{+}_{w,i}\left(\sigma_i^{w,-}, t_w^{k+n-1,B}\right)\right) \ominus \beta^{-}_{r,k}\left(t_w^{k+n-1,B} - \rho_k^{w,-}\right) = \\ &= n \cdot B - \left\lfloor \beta^{-}_{r,k}\left(t_w^{k+n-1,B} - \rho_k^{w,-}\right)\right\rfloor = \\ &= n \cdot B - \left\lfloor f_{r,\min}\left(\sigma^{-} + (n-1)\cdot\Gamma_{w,\min} + \frac{B}{f_{w,\max}}\right)\right\rfloor . \qquad (6.61)\end{aligned}$$

If $t_w^{k+n-1,B} < t_r^{k,0}$, we only need to evaluate at $t_r^{k,0}$, leading to

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n-1} \beta^{+}_{w,i}\left(\sigma_i^{w,-}, t_r^{k,0}\right)\right) \ominus \beta^{-}_{r,k}\left(t_r^{k,0} - \rho_k^{w,-}\right) = \\ &= n \cdot B. \qquad (6.62)\end{aligned}$$

For the main case, i.e., the $k+n$th writer and the kth reader, an evaluation of the difference at $t_w^{k+n,0}$ is not required: As $t_w^{k+n-1,B} < t_w^{k+n,0}$ and $f_{w,\max} \leq f_{r,\min}$, this bound would only be smaller than the one computed for $t_w^{k+n-1,B}$ above.

In addition, we also need to consider the relation between the $k+n$th writer and the $k+1$st reader. If there is no overlap, i.e., if $t_r^{k,B} \leq t_w^{k+n,B} < t_r^{k+1,0}$, then also $t_w^{k+n-1,B} < t_r^{k,0}$ holds. Hence, this evaluation will result in the same bound as Equation (6.62). If $t_r^{k+1,0} \leq t_w^{k+n,B}$, the same argument can be made for Equations (6.60) and (6.61).

Case c: $t_r^{k,B} < t_w^{k+n,0} < t_r^{k+1,0}$ Here, we again have to consider a possible overlap with the $k+n-1$st writer. Due to $\Gamma_{w,\min} \geq \Gamma_{r,\min}$, we again only need to consider the case $t_w^{k+n-1,0} < t_r^{k,0}$, which leaves only three remaining cases: If $t_w^{k+n-1,B} < t_r^{k,0}$, we get

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n-1} \beta^{+}_{w,i}\left(\sigma_i^{w,-}, t_r^{k,0}\right)\right) \ominus \beta^{-}_{r,k}\left(t_r^{k,0} - \rho_k^{w,-}\right) = \\ &= n \cdot B - 0. \qquad (6.63)\end{aligned}$$

If $t_r^{k,0} \leq t_w^{k+n-1,B} < t_r^{k,B}$, we evaluate at $t_r^{k,0}$ and $t_w^{k+n-1,B}$, leading to

$$\begin{aligned}\tau_r &> \left(\sum_{i=k}^{k+n-1} \beta^{+}_{w,i}\left(\sigma_i^{w,-}, t_r^{k,0}\right)\right) \ominus \left(\sum_{i=k}^{k+1} \beta^{-}_{r,i}\left(t_r^{k,0} - \rho_i^{w,-}\right)\right) = \\ &= (n-1)\cdot B + \left\lceil \beta^{+}_{w,k+n-1}(\sigma^{w,-}_{k+n-1}, t_r^{k,0})\right\rceil - 0 = \\ &= (n-1)\cdot B + \left\lceil f_{w,\max}\left(-\sigma^{-} - (n-1)\cdot\Gamma_{w,\min}\right)\right\rceil , \qquad (6.64)\end{aligned}$$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Table 6.1: Correspondence table for usage of the results given in Table 6.4 under the setting of Section 6.6.

Section 6.5	Section 6.6
$t_r^{k,0}$	$t_w^{k,0} - \sigma^+$
$\Gamma_{w,\min}$	$\Gamma_{r,\min}$
$\Gamma_{r,\min}$	$\Gamma_{w,\min}$
$f_{r,\min}$	$f_{w,\min}$
$f_{r,\max}$	$f_{w,\max}$
$f_{w,\min}$	$f_{r,\min}$
$f_{w,\max}$	$f_{r,\max}$
σ^-	$-\sigma^+$
σ^+	$-\sigma^-$

$$\begin{aligned}\tau_r > & \left(\sum_{i=k}^{k+n-1} \beta^+_{w,i}\left(\sigma_i^{w,-}, t_w^{k+n-1,B}\right)\right) \ominus \left(\sum_{i=k}^{k+1} \beta^-_{r,i}\left(t_w^{k+n-1,B} - \rho_i^{w,-}\right)\right) = \\ & = n \cdot B - \left\lfloor \beta^-_{r,k}(t_w^{k+n-1,B}) \right\rfloor - 0 = \\ & = n \cdot B - \left\lfloor f_{r,\min}\left(\sigma^- + (n-1)\cdot \Gamma_{w,\min} + \frac{B}{f_{w,\max}}\right) \right\rfloor . \end{aligned} \tag{6.65}$$

The remaining case of $t_r^{k,B} \leq t_w^{k+n-1,B} \leq t_w^{k+n,0}$ is dominated by the main case below, hence not considered explicitly.

For the main case of the $k+n^{\text{th}}$ writer and the k^{th} reader, we consider first the case of $t_w^{k+n,B} < t_r^{k+1,0}$. We immediately see that this implies $t_w^{k+n-1,B} < t_r^{k,0}$, which defined Equation (6.63), hence this case does not need to be considered. For $t_w^{k+n,B} \geq t_r^{k+1,0}$, the same holds with respect to Equations (6.64) and (6.65).

6.6 Multi-Burst Constraints – Reader before Writer

The analysis for this case is symmetric to the one presented in Section 6.5, as it requires only that the writer is considered as the reader and the other way around. Hence, reusing the already gained results requires only a swapping of the parameters to consider the writer here as the reader in the formulas of Section 6.5 listed in Table 6.4. We listed these transformations in Table 6.1. Note that this transformation also holds for the single-burst case, as can be easily seen by comparing Tables 6.2 and 6.3.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 6.2: Single-burst constraints for the condition that $\delta_w^+(\sigma, t) < M + \delta_r^-(t)$, i.e., $\tau_r > \beta_w^+(\sigma, t) \ominus \beta_r^-(t)$

Case	Bound
$0 < \sigma^- < \dfrac{B-1}{f_{r,\min}} < \dfrac{B-1}{f_{w,\max}} + \sigma^-$	$\tau_r^a > 0$
$0 < \sigma^- < \dfrac{B-1}{f_{w,\max}} + \sigma^- < \dfrac{B-1}{f_{r,\min}}$	$\tau_r^b > B - 1 - \left\lfloor f_{r,\min} \left(\dfrac{B-1}{f_{w,\max}} + \sigma^- \right) \right\rfloor$
$\sigma^- < 0 < \dfrac{B-1}{f_{r,\min}} < \dfrac{B-1}{f_{w,\max}} + \sigma^-$	$\tau_r^c > \left\lceil -f_{w,\max} \sigma^- \right\rceil$
$\sigma^- < 0 < \dfrac{B-1}{f_{w,\max}} + \sigma^-$ and $f_{w,\max} < f_{r,\min}$	$\tau_r^{d_1} > \left\lceil -f_{w,\max} \sigma^- \right\rceil$
$\sigma^- < 0 < \dfrac{B-1}{f_{w,\max}} + \sigma^-$ and $f_{w,\max} \geq f_{r,\min}$	$\tau_r^{d_2} > B - 1 - \left\lfloor f_{r,\min} \left(\dfrac{B-1}{f_{w,\max}} + \sigma^- \right) \right\rfloor$
$\dfrac{B-1}{f_{w,\max}} + \sigma^- < 0 < \dfrac{B-1}{f_{r,\min}}$	$\tau_r^e > B - 1$
$0 < \dfrac{B-1}{f_{r,\min}} < \sigma^-$	$\tau_r^f > 0$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 6.3: Single-burst constraints for the condition that $\delta_w^-(\sigma, t) > \delta_r^+(t)$, i.e., $\tau_w > \beta_r^+(t) \ominus \beta_w^-(\sigma, t)$

Case	Bound
$0 < \sigma^+ < \frac{B-1}{f_{r,\max}} < \frac{B-1}{f_{w,\min}} + \sigma^+$ and $f_{r,\max} < f_{w,\min}$	$\tau_w^{a_1} > \lceil f_{r,\max}\sigma^+ \rceil$
$0 < \sigma^+ < \frac{B-1}{f_{r,\max}} < \frac{B-1}{f_{w,\min}} + \sigma^+$ and $f_{r,\max} \geq f_{w,\min}$	$\tau_w^{a_2} > B - 1 - \left\lfloor f_{w,\min}\left(\frac{B-1}{f_{r,\max}} - \sigma^+\right)\right\rfloor$
$0 < \sigma^+ < \frac{B-1}{f_{w,\min}} + \sigma^+ < \frac{B-1}{f_{r,\max}}$	$\tau_w^{b} > \lceil f_{r,\max}\sigma^+ \rceil$
$\sigma^+ < 0 < \frac{B-1}{f_{r,\max}} < \frac{B-1}{f_{w,\min}} + \sigma^+$	$\tau_w^{c} > B - 1 - \left\lfloor f_{w,\min}\left(\frac{B-1}{f_{r,\max}} - \sigma^+\right)\right\rfloor$
$\sigma^+ < 0 < \frac{B-1}{f_{w,\min}} + \sigma^+ < \frac{B-1}{f_{r,\max}}$	$\tau_w^{d} > 0$
$\sigma^+ < \frac{B-1}{f_{w,\min}} + \sigma^+ < 0$	$\tau_w^{e} > 0$
$0 < \frac{B-1}{f_{r,\max}} < \sigma^+$	$\tau_w^{f} > B - 1$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table 6.4: Multi-Burst Constraints for τ_r where $\Gamma_{\text{min}} = \max\{\Gamma_{w,\text{min}}, \Gamma_{r,\text{min}}\}$.

Constraints	Bound
$f_{w,\text{max}} \leq f_{r,\text{min}}$ and $\sigma^- + (n-1) \cdot \Gamma_{\text{min}} + \frac{B}{f_{w,\text{max}}} < 0 \leq \sigma^- + n \cdot \Gamma_{\text{min}}$	$\tau_r > n \cdot B$
$f_{w,\text{max}} \leq f_{r,\text{min}}$ and $\sigma^- + (n-1) \cdot \Gamma_{\text{min}} < 0 \leq \sigma^- + (n-1) \cdot \Gamma_{\text{min}} + \frac{B}{f_{w,\text{max}}}$	$\tau_r > (n-1) \cdot B + \lceil f_{w,\text{max}} (-\sigma^- - (n-1) \cdot \Gamma_{\text{min}}) \rceil$
$f_{w,\text{max}} \leq f_{r,\text{min}}$ and $0 \leq \sigma^- + (n-1) \cdot \Gamma_{\text{min}} + \frac{B}{f_{w,\text{max}}} < \frac{B}{f_{r,\text{min}}}$	$\tau_r > n \cdot B - \left\lfloor f_{r,\text{min}} \left(\sigma^- + (n-1) \cdot \Gamma_{\text{min}} + \frac{B}{f_{w,\text{max}}} \right) \right\rfloor$
$f_{w,\text{max}} \leq f_{r,\text{min}}$ and $\frac{B}{f_{r,\text{min}}} \leq \sigma^- + n \cdot \Gamma_{\text{min}} + \frac{B}{f_{w,\text{max}}} < \Gamma_{\text{min}}$	$\tau_r > n \cdot B$
$f_{w,\text{max}} > f_{r,\text{min}}$ and $0 < \sigma^- + n \cdot \Gamma_{\text{min}} + \frac{B}{f_{w,\text{max}}} < \frac{B}{f_{r,\text{min}}}$	$\tau_r > (n+1) \cdot B - \left\lfloor f_{r,\text{min}} \left(\sigma^- + n \cdot \Gamma_{\text{min}} + \frac{B}{f_{w,\text{max}}} \right) \right\rfloor$
$f_{w,\text{max}} > f_{r,\text{min}}$ and $\frac{B}{f_{r,\text{min}}} \leq \sigma^- + n \cdot \Gamma_{\text{min}} + \frac{B}{f_{w,\text{max}}} < \Gamma_{\text{min}}$	$\tau_r > n \cdot B$
$f_{w,\text{max}} > f_{r,\text{min}}$ and $\Gamma_{\text{min}} \leq \sigma^- + n \cdot \Gamma_{\text{min}} + \frac{B}{f_{w,\text{max}}} < \Gamma_{\text{min}} + \frac{B}{f_{r,\text{min}}}$	$\tau_r > n \cdot B - \left\lfloor f_{r,\text{min}} \left(\sigma^- + (n-1) \cdot \Gamma_{\text{min}} + \frac{B}{f_{w,\text{max}}} \right) \right\rfloor$

6.7 Evaluation

In this section, we apply our theoretical results to some showcase applications. The outcomes experimentally confirm that our approach works as expected, and also demonstrates its practical feasibility.

6.7.1 Simulations

We implemented the schematics shown in Figure 6.3 in VHDL. The synthesis was done with Synopsys® Design Compiler Version M-2016.12 using the UMC65 Low Leakage library. All simulations were conducted with a test bench in Mentor's® ModelSim 10.5c. We experimentally validated correct operation of the design and evaluated the corner cases of the formulas of Section 6.3 for their correctness and tightness. Furthermore, we evaluated the multi-burst formulas of Sections 6.5 and 6.6.

In general, the bounds proved to be as tight as one can expect: Gaps in the tightness can be attributed to the over-approximation due to the rounding of the $\ominus$ operation, recall Lemma 6.6. Hence, the parameters had sometimes to be pushed so that the collision would occur close to the middle of the next memory address to have an actual collision of the pointers. Further, we have to stress that the compensation for the slack (cf. Lemma 6.3), especially in corner cases, is paramount for correctness, as the over-approximation is too small to hide these effects in every case.

Self-stabilization Figure 6.13 shows the self-stabilization of a design with $f_{\min} = 190\,\mathrm{MHz}$, $f_{\max} = 200\,\mathrm{MHz}$, $\sigma = [-1, 1]\,\mathrm{ns}$, and $B = 20$. This results in a parameterization of $\tau_w = \tau_r = 4$, $M = 8$, and $I = 2$. To verify the self-stabilization properties, we set every internal signal in the design to an undefined value (`X`) during the run and continued the simulation. The first vertical (yellow) line in the figure marks the time where `X` was applied for 1 ns. The second line gives the time the slow clock provided the period signal to the reader, followed by the third which marks the time when the read pointer is at τ_w and correct operation is resumed.

Faults To show the effect of incorrect system dimensioning, we use the example shown in Figure 6.14a: it demonstrates the high-level effect of a read pointer colliding with a write pointer. Note that the scenario given can be caused by multiple mismatches: Clock frequencies out of the specified range, an undersized ring buffer, or a skew outside its specified interval. The effect is that the data read is either old data (last G read in Figure 6.14a) or partially invalid, or even metastable.

We also performed a simulation in ModelSim of such a setting. Figure 6.14b shows a system designed for $f_{\min} = 195\,\mathrm{MHz}$, $f_{\max} = 200\,\mathrm{MHz}$, $\sigma = [-1.1, 1]\,\mathrm{ns}$, and $B = 200$. In this example, the writer writes at the falling edge of its clock. This half-cycle shift is accounted for in the skew, resulting in $\sigma = [3.9, 6.128\,206]\,\mathrm{ns}$, without additional slack for setup/hold-timing and signal propagation. This results in a parameterization of $\tau_w = 8$, $\tau_r = 6$, $M = 14$, and $I = 7$. To force a failure, we double B to 400 during the run. In Figure 6.14b, it is apparent that the time-window in which the data read

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

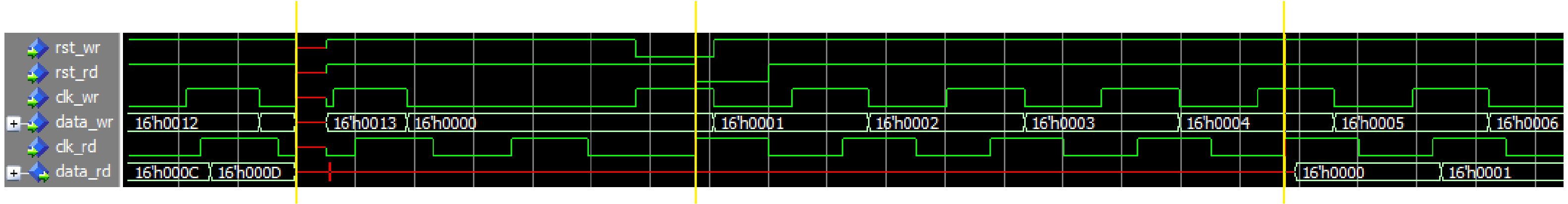
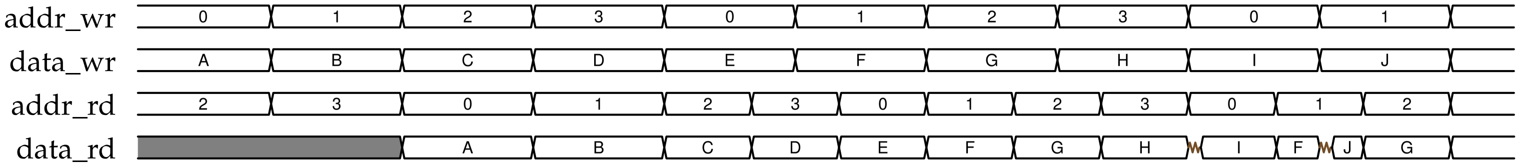
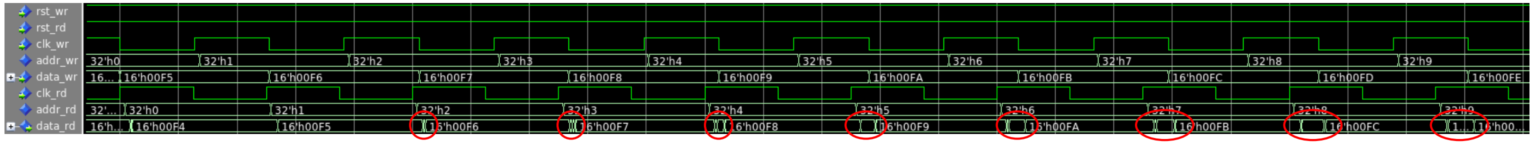


Figure 6.13: A screenshot of the simulation after an invalid state has been forced. The `rst_` signals are the period signals provided by the corresponding slow clocks `clk_`. The signal `data_wr` is the data written into the memory on the falling edge of `clk_wr`, `data_rd` is the data read on `clk_rd`. The white grid lines separate 2 ns intervals. The first vertical (yellow) line in the figure marks the time where the `X` where applied for 1 ns. The second line gives the time the slow clock provided the period signal to the reader, followed by the third which marks the time when the read pointer is at τ_w and correct operation is resumed.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) A high-level view of the effect of a pointer collision. In this scenario we have a memory size $M = 4$. The first two read-outs of the reader are invalid in this scenario, e.g., due to a restart of the system. After the fourth tick, the reader speeds up and runs into the writers current address: When `I` is written, the reader tries to read is too. The effects can range from metastability, over sampling of inconsistent data, to sampling of old data (last read of the reader is `G`).


rst_wr
rst_rd
clk_wr
addr_wr
data_wr
clk_rd
addr_rd
data_rd
32'h0 32'h1 32'h2 32'h3 32'h4 32'h5 32'h6 32'h7 32'h8 32'h9
16'h00F5 16'h00F6 16'h00F7 16'h00F8 16'h00F9 16'h00FA 16'h00FB 16'h00FC 16'h00FD 16'h00FE
16'h00F4 16'h00F5 16'h00F6 16'h00F7 16'h00F8 16'h00F9 16'h00FA 16'h00FB 16'h00FC

(b) A screenshot of the simulation where the burst length has been doubled compared to the designed length to force an error. In this scenario, the faster reader caught up with the slower writer, leading to inconsistent read-outs, marked red. Note that the writer performs its write operations at the falling clock edge here. The white grid lines separate 2 ns intervals.

Figure 6.14: A high-level- (6.14a) and a simulation-view (6.14b) of the effect of a fault on the transmitted data.

(the signal `data_rd`) changes gradually increases. While a short window is normal and caused by delays in gates and wires, the growth is a symptom of the collision: As the relevant read/write clock transitions change their temporal order, the window increases. Note carefully that the data read out during such sufficiently large windows is already stable (but wrong), and at some point, old data will be read in a way that is indistinguishable from fresh data.

6.7.2 Usage in highly-synchronized systems

As a practical example, we consider the Byzantine fault-tolerant self-stabilizing clock synchronization scheme presented in [KHL16], which allows to implement a highly synchronous GALS system. It has a slow clock of 20 kHz, with a measured skew bound of $[-180, 180]$ ps. The fast clock has a frequency in the range of 120 MHz ± 2.6 kHz and a burst length of $B = 6500$. Due to the tight synchronization, the overhead caused by the communication primitive should be as small as possible, which the presented solution achieves perfectly.

In [KHL16], the slow clock is generated from the fast clock by a simple clock divider. As a consequence, slow clock ticks are not numbered and there is no varying gap between subsequent bursts. Nevertheless, this does not impair the applicability of our model, since the clock skew σ is smaller than the fast clock period: Since the multi-pulse constraints collapse to the trivial requirement $\tau_r, \tau_w > 1$, we choose $\tau_w = \tau_r = 2$ and hence $M = 4$ and $I = 1$. The latter confirms that the period signal occurs at every slow tick, so tick numbering is indeed not needed here.

We synthesized the design with 8-bit data-width using UMC65, which results in an pre-routing area of $457.2\,\mu\text{m}^2$. As we do not have access to a library with memory cells, the memory was synthesized with registers, which lead to an area of $415.8\,\mu\text{m}^2$ for the ring buffer alone. We believe that using memory cells would reduce the area significantly, especially as the interconnect area is estimated as $1886.6\,\mu\text{m}^2$. Note that we also synthesized the same design with NanGate 45 nm Cell Library [Nan08], which resulted in a pre-routing area of $254.3\,\mu\text{m}^2$, also without memory cells.

6.8 Conclusions

We presented a simple self-stabilizing communication protocol for multi-synchronous GALS systems. Compared to related work, e.g., [CN04; Gre95], the presented approach is metastability-free in the fault-free case. This is due to the separation of the control circuitry between the clock domains of reader and writer. The analysis of the required memory sizing and initial offsets between reader and writer ensure sufficient separation of the pointers in the memory. Hence, both components will never access the same memory location at the same time, ensuring a separation of the data level as well. The correctness of the analysis has been validated with simulations of a VHDL implementation.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

TU WIEN Bibliothek, Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

> Beyond mountains, there are mountains.
>
> — HAITIAN PROVERB

PROVIDING A CIRCUIT with well-synchronized clock signals is hard. As fault-tolerance is non-optional in some application domains, we considered ways to generate a clock signal in a fault-tolerant and self-stabilizing way in a small area of the IC and distribute it over the whole chip without sacrificing fault-tolerance and self-stabilization.

We first considered techniques for distributing a clock signal in a way that preserves its fault-tolerant properties and is able to achieve self-stabilization after the occurrence of an overwhelming number of transient faults. Our first approach, termed HEX, is a hexagonal grid with special nodes located at the grid points. Apart from a theoretical analysis of the worst-case skew between neighbor nodes in the grid, we conducted extensive simulations in order to assess the average-case skew as well. Our results revealed that there is a large gap between the worst-case and the average-case skew. Furthermore, the fault-tolerance and self-stabilization properties have been studied in these simulations. As expected, HEX in particular, achieves self-stabilization faster than the theoretically proven upper bound. However, faults have a significant impact on the skews in the system, albeit these effects are restricted to the neighborhood of the faulty nodes.

As the skew in HEX tends to grow in adverse settings, we searched for improvements of the overall architecture. Extensive exploratory simulations have been conducted to survey alternative topologies and firing behaviors. These simulations have shown that, depending on the quality of the underlying clock generation and the fault model, different combinations of topology and firing behavior achieve the best result.

Building upon the assumption that a well-behaved fault-tolerant clock synchronization is available, we selected a particular combination, termed TRIX, as the most favorable result of our survey. Whereas HEX still lacks an optimized design of its quite complex nodes, i.e., a transistor-level implementation, we provided such an implementation for TRIX. It also showcased that, as expected, transistor-level implementations lead to a significant reduction of the delay variance of the nodes. Furthermore, the lower abstraction level of the implementation allows a significant reduction of the required area, as certain fault-tolerance aspects did not require explicit handling anymore. In particular, TRIX is able to mask a metastable signal from one of its neighbors, which HEX does not.

The availability of a fault-tolerant and self-stabilizing clock signal also enabled a solution for the next layer of a holistic fault-tolerant and self-stabilizing system design: communication between components. We presented a FIFO-based communication link implementation, which is able to transmit data from the sender to the receiver with every cycle of a clock signal as distributed by HEX/TRIX augmented by a fast clock. Due to the resilience of the distributed clock, our buffer management is able to achieve self-stabilization by recovering corrupted FIFO pointers without any overhead in the fault-free case and without the need to detect corruptions.

Therefore, we affirm the solvability of the two research question posed in Chapter 1:

- We can build a clock distribution that is fault-tolerant, self-stabilizing and implementable in practice.
- We can utilize the properties of the distributed clock signal to build a communication primitive that is comparable in performance to what is achievable with a global clock.

7.1 Future Work

The research questions of this thesis focused only on a small part of our vision: A holistic fault-tolerant and self-stabilizing system design. There is still a lot of work to be done to actually achieve this goal, however. The obvious next step is to facilitate chip-wide communication: A fault-tolerant and self-stabilizing routing methodology has to be designed for this purpose, which builds on our point-to-point communication solution. If this was available, distributed algorithms could be executed on top of it.

Besides this big picture, there are many open directions, unfinished work, and space for improvement left in the wake of this thesis, which we will briefly discuss below.

Adaption of the HEX analysis from $\mathcal{A}^2_{\text{adj}}$ to $\mathcal{A}^2_{\text{any}}$: While we have seen in the simulations that $\mathcal{A}^2_{\text{any}}$ performs better than $\mathcal{A}^2_{\text{adj}}$ in all cases, the question remains if this also holds in the worst-case. Due to the additional firing rules, the path construction using left zig-zag paths is not possible anymore. Nonetheless, it is simple to show that a node can only trigger earlier under $\mathcal{A}^2_{\text{any}}$ by enumeration of all cases. However, showing that

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

this does not violate the skew potential bound within a layer is not trivially achievable and quite likely requires an argument including causal path construction over multiple layers.

Analysis of TRIX: As TRIX has been selected to replace HEX, a worst-case analysis of its skew bounds is appropriate before further use. Contrary to the analysis of HEX under $\mathcal{A}^2_{\text{any}}$, even extended zig-zag path constructions cannot be used here, due to the absence of intra-layer links. This leaves an interesting research challenge open.

Furthermore, we considered an analytic study of the average-case behavior. Even tough TRIX can be modeled in the Min-Max-Plus algebra [HOW06], this does not simplify the analysis: General solutions for this kind of algebra are rare and depend on specific distributions of the involved random variables like the delays of the links. Distributions for which solutions exist, however, are suitable for the delay distributions used for modeling a CDN.

Approaching the average-case analysis from a probabilistic point of view, via order statistics, gives the intuition that TRIX should be able to deliver the result suggested by the simulations. It can be shown that the median-based determination of the firing time ensures that neighboring nodes are likely to fire close to each other.

Improving TRIX's Implementation: The achieved delay variations of the presented transistor-level implementation of TRIX are quite good. However, chip-wide clock trees are able to achieve better results, yet, they are unable to provide fault-tolerance as TRIX does. We hence consider that an alternative transistor-cell construction, along the lines of [DC93], should be able to reduce the delay variations further.

As CDNs consume a major fraction of the power required by an IC, the power required by a TRIX-based CDN is of interest. However, the specific loads of the circuits attached to a CDN have a huge impact on its power requirement. Therefore, a comparison has to be conducted, by comparing the same design clocked by a traditional and a TRIX-based CDN. A first target could be a GALS architecture, where TRIX performs the distribution of the clock signal to the synchronous islands. Note that HEX and TRIX never intended to completely remove the need for clock trees at the lowest level of the CDN. While GALS seem to be an ideal basis for the comparison, not every communication system will be able to properly handle the skew guarantees provided by the two CDNs without adaptions, however.

Overall, we conjecture that our fault-tolerant design will require more power; however, the question is how big the overhead is.

Implementation of the point-to-point link: The simulations of the point-to-point link in Chapter 6 were based on a pre-layout synthesis of the design, so that wire delays were not included. However, considering that we were not able to use SRAM memory for the FIFO buffer, a post-layout synthesis of the design would have significantly suffered from the induced routing delays. This is due to the massive area overhead caused by the construction of the memory via registers. Therefore, an evaluation of the presented

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

solution via a fully placed & routed implementation with SRAM memory should be considered. We expect that this implementation would be significantly smaller in area and hence exhibit lower delays due to the shorter paths between the counters and the memory. Finally, the feasibility of our link as a communication primitive in a GALS system should be investigated.

Fault-Tolerant and Self-Stabilizing Clock Generation: During our search for an approach to analyze TRIX, we found that a clock generation similar to the underlying concept of, e.g., [LE09] may be able to deliver better results than the clock generation and distribution approach considered in this thesis. In particular, we conjecture that the clock generation approach proposed by Fairbanks and Moore [Fai06; FM05], already presented in Section 2.4, is a very promising candidate. However, the authors did not provide any consideration with regard to the fault-tolerance of the circuit.

We hence built the circuit, following the design recommendations given in [FM05], with a UMC 65 nm process. While we were able to build a circuit behaving as intended, there are a few caveats: The scaling required for the transistors lead to sizings well outside of the range for which parameterized transistors are provided by the library. We, hence, had to manually parameterize some transistors. In combination with the omitted parasitic capacitances and wire models inside the cells of the circuit,[1] we are uncertain if the behavior of the circuit, especially for boundary conditions, is representative and scalable to other technologies.

Furthermore, we conducted simple fault-injection experiments, which led to an increase of the skew, due to frequency variations of $\approx 15\%$ between the cells. We have to stress that, due to the strong coupling between the cells of the circuit, our simple fault-injections may not have modeled faults faithfully. We conjecture that, with a suitable parameterization of the circuit, the frequency variations can be reduced to an acceptable level. Nonetheless, for a preliminary evaluation, we consider this a promising result.

As an analysis of the transistor-level circuit may not provide sufficient insights into the actual principals behind this alternative clock generation method, we consider a simplified digital model a promising approach. Preliminary simulations of such a digital model provided impressive results with respect to fault-tolerance and self-stabilization. However, the results also suggest that the function that determines the switching delay of a cell, based on the arrival times of the signals from its neighbors, is crucial for the stability of the system. Overall, we believe that this approach may eventually lead to an impressive fault-tolerant and self-stabilizing clock generation scheme.

---

[1] We used a RCR-element to model the wires between the cells.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

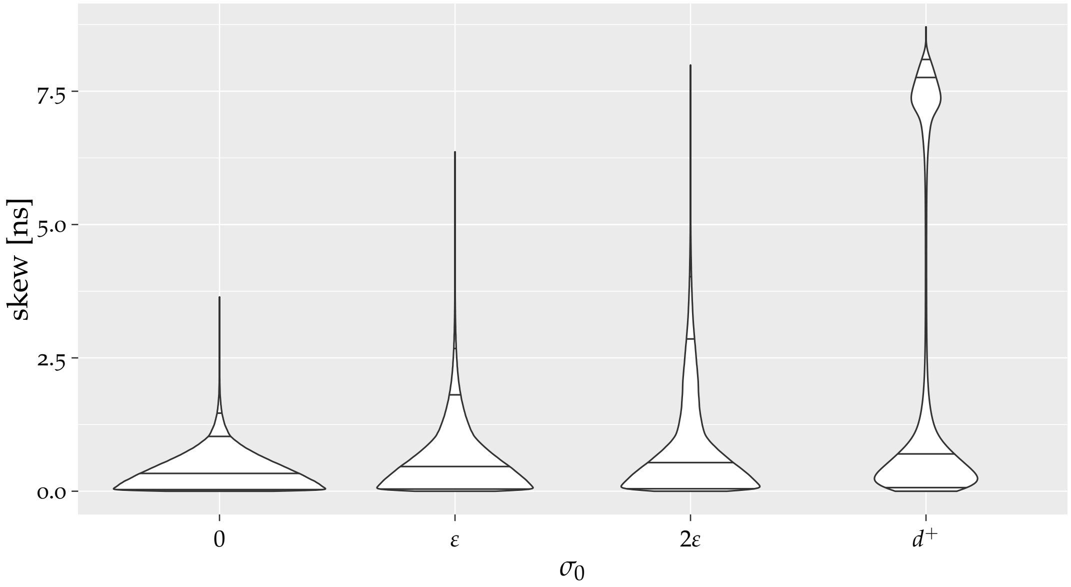
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

APPENDIX A

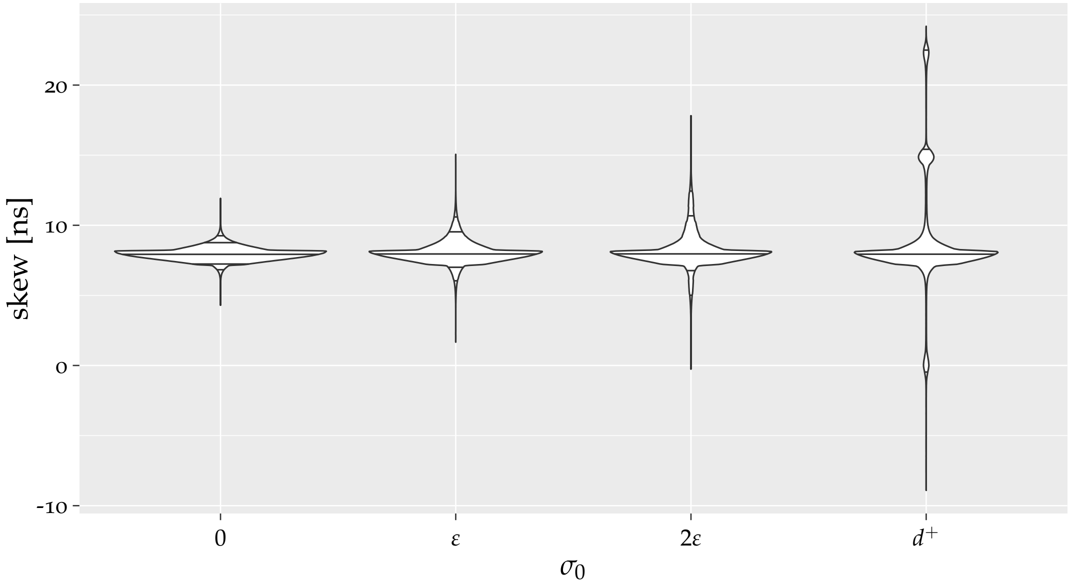
ADDITIONAL VISUALIZATIONS OF THE SIMULATIONS OF CHAPTER 4

A.1 Plots of the Pulse-based Algorithms in $\mathcal{G}_H$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



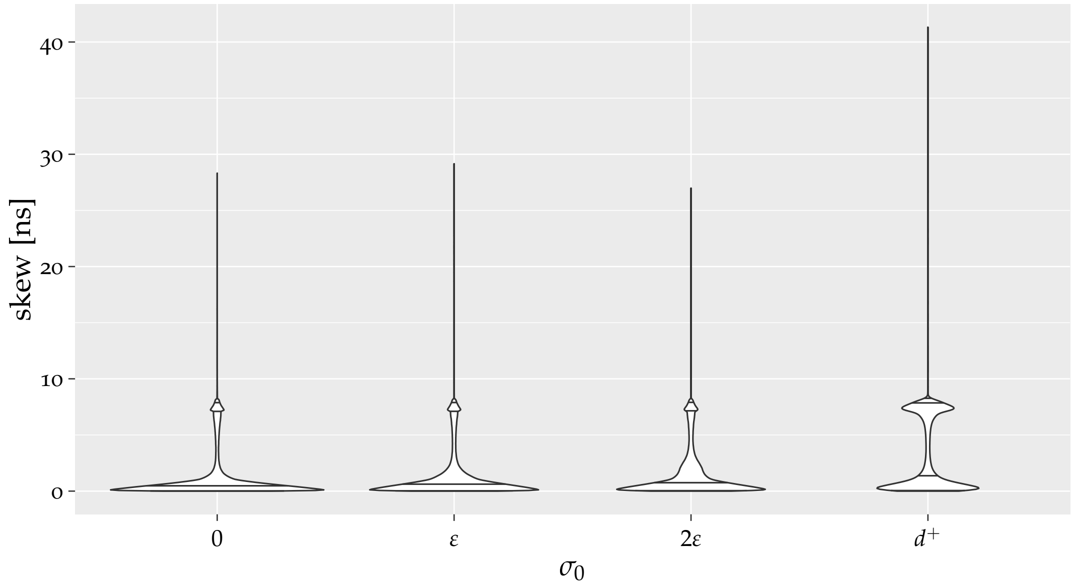
(a) intra-layer skews



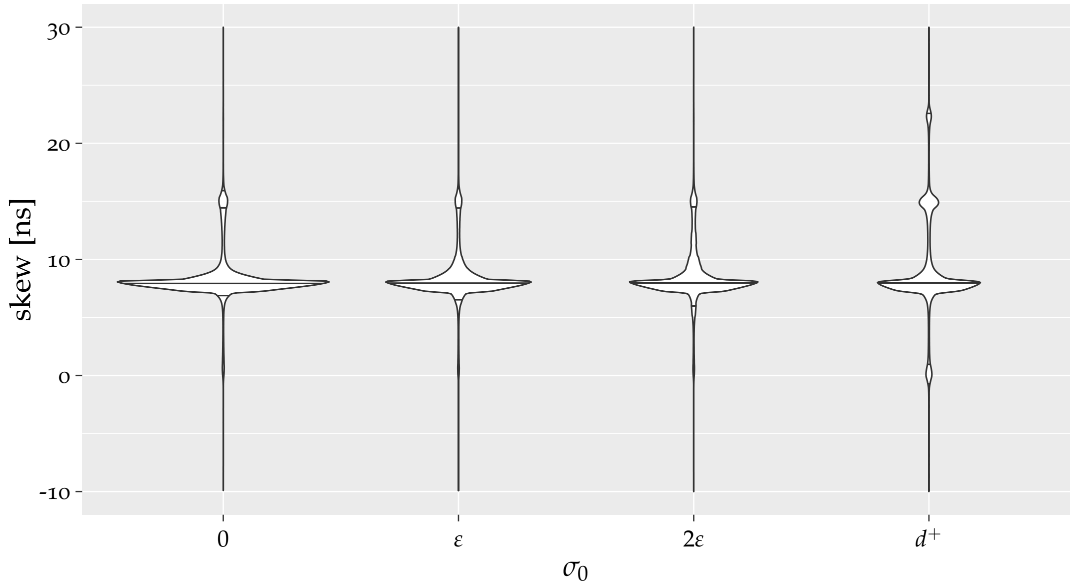
(b) inter-layer skews

Figure A.1: Skew in $\mathcal{G}_H$ with $\mathcal{A}^2_{\mathrm{adj}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



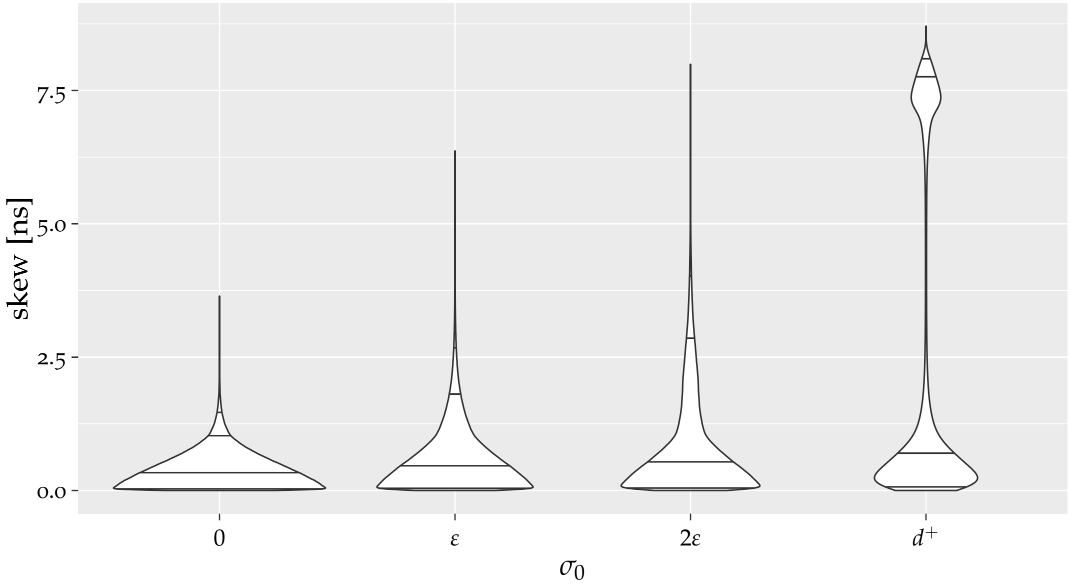
(a) intra-layer skew



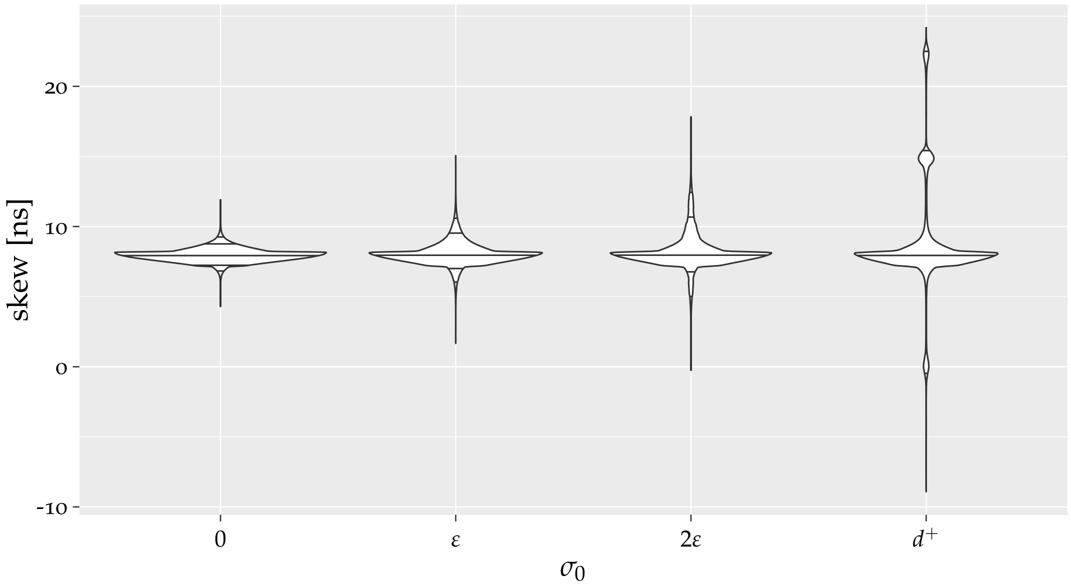
(b) inter-layer skew

Figure A.2: Skew in $\mathcal{G}_H$ with $\mathcal{A}^2_{\text{adj}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the inter-layer skews have been limited to be within $[-10, 30]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



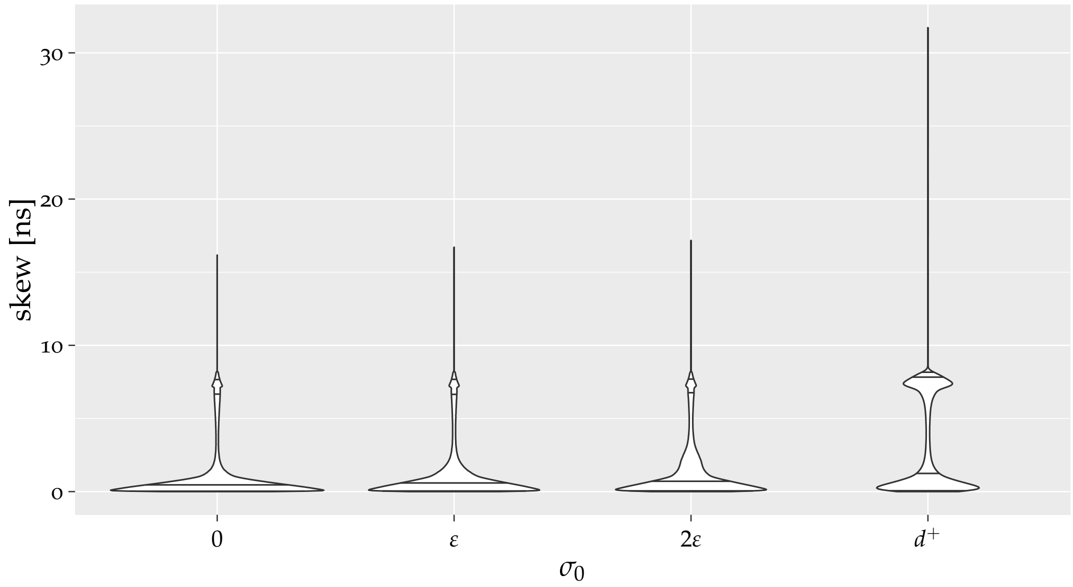
(a) intra-layer skew



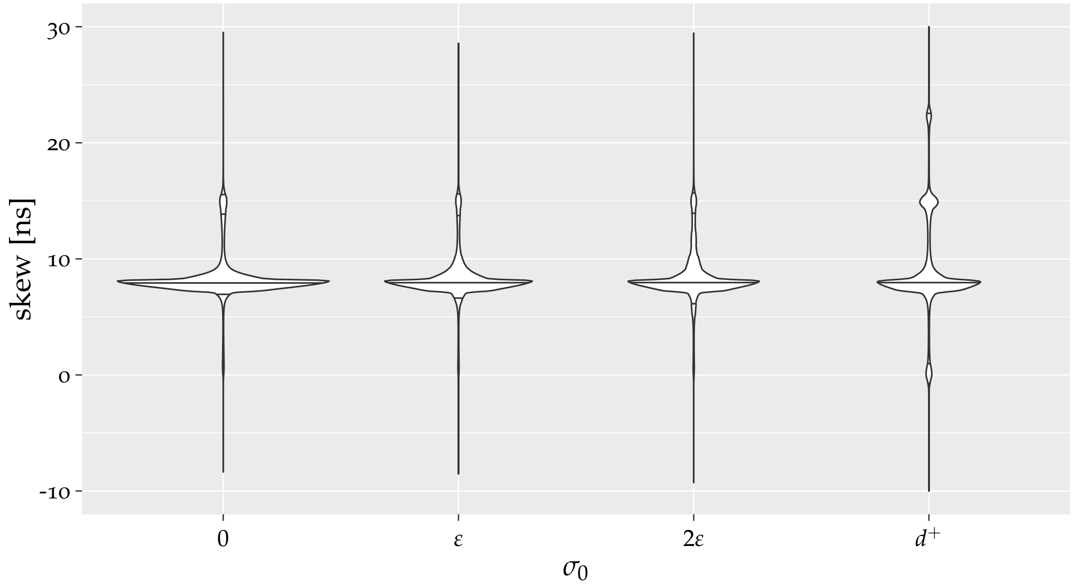
(b) inter-layer skew

Figure A.3: Skew in $\mathcal{G}_H$ with $\mathcal{A}^2_{\text{any}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skew



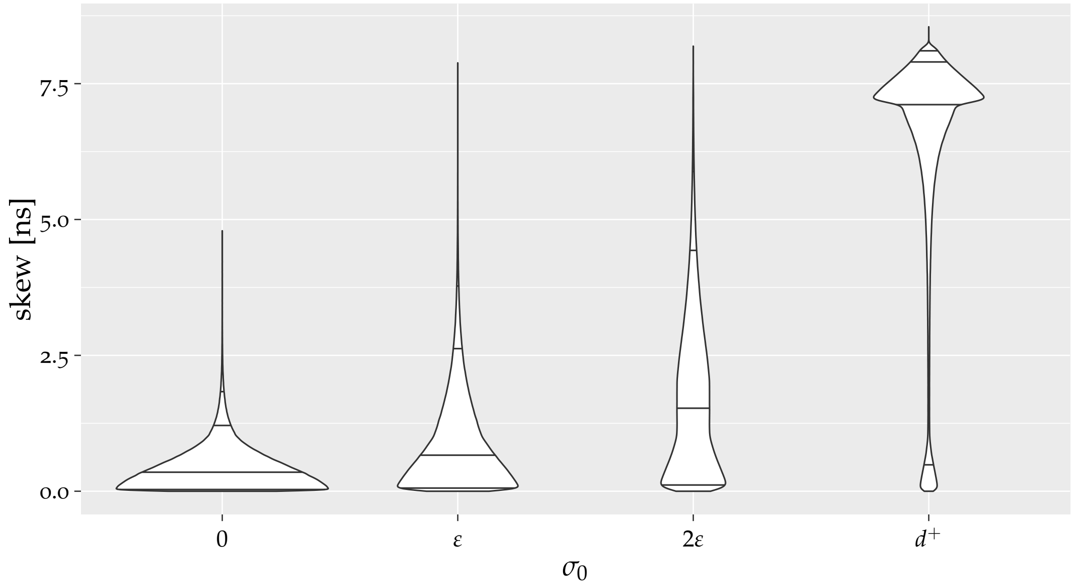
(b) inter-layer skew

Figure A.4: Skew in $\mathcal{G}_H$ with $\mathcal{A}^2_{\text{any}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the inter-layer skews have been limited to be within $[-10, 30]$ ns.

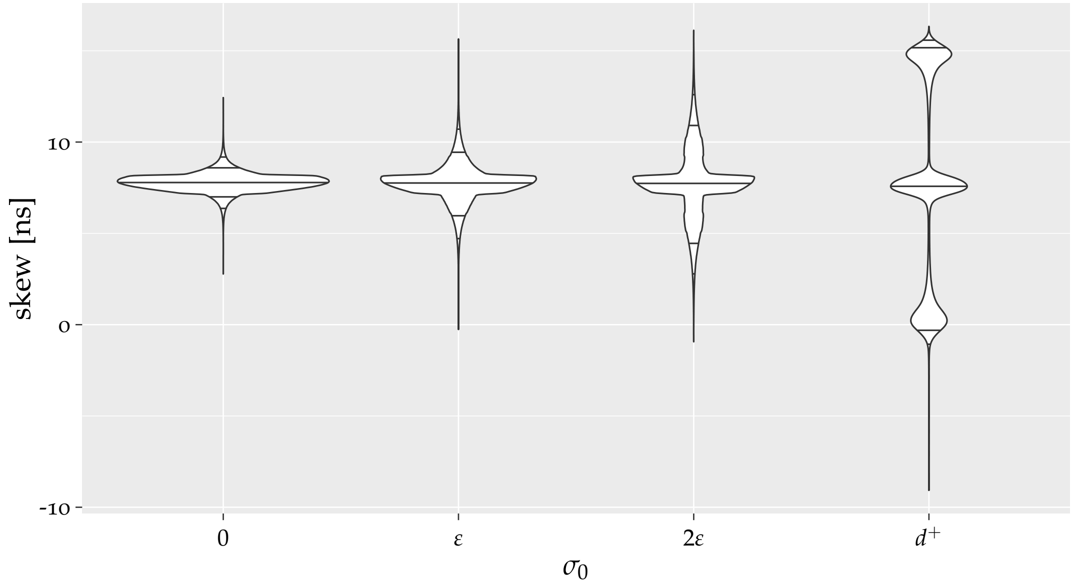
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

A.2 Plots of the Pulse-based Algorithms in $\mathcal{G}_A$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



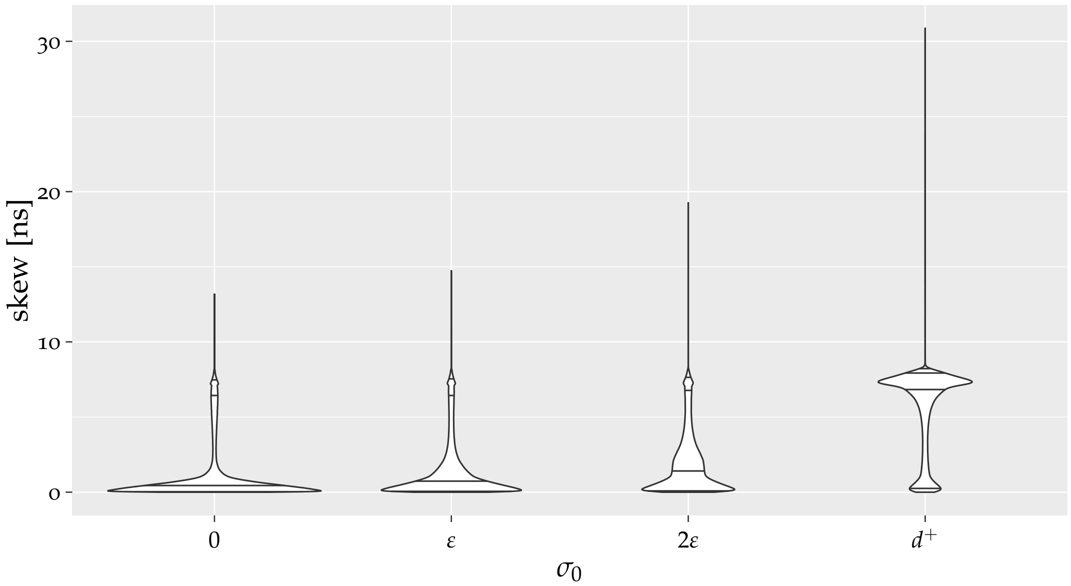
(a) intra-layer skew



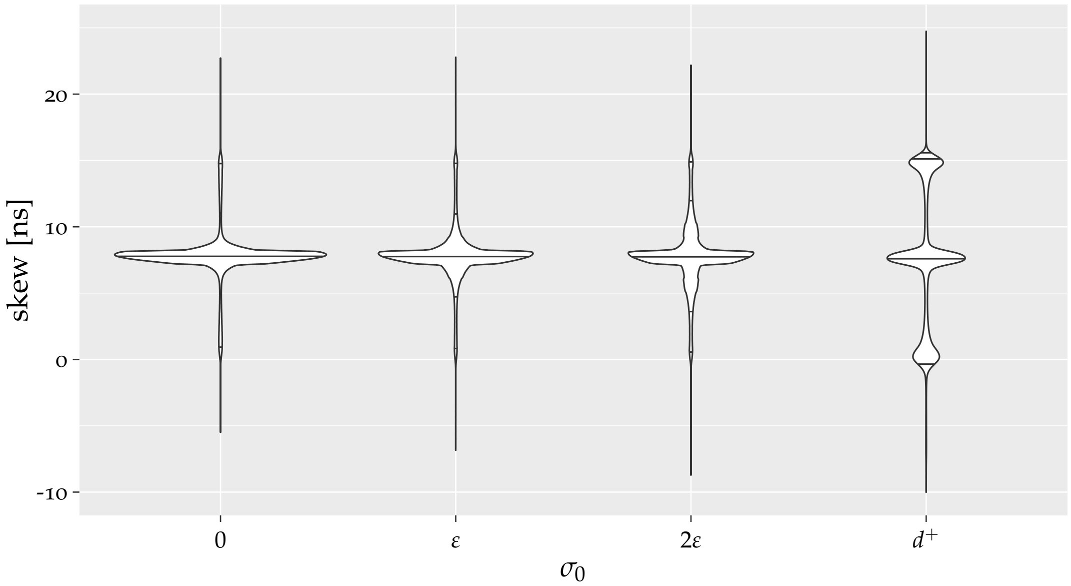
(b) inter-layer skew

Figure A.5: Skew in $\mathcal{G}_A$ with $\mathcal{A}^2_{\text{adj}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



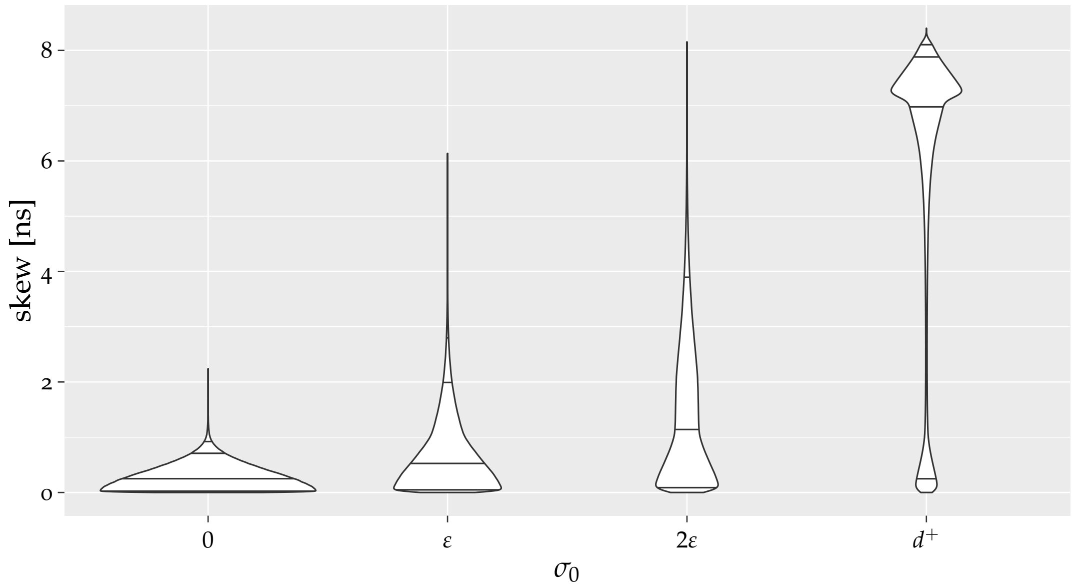
(a) intra-layer skew



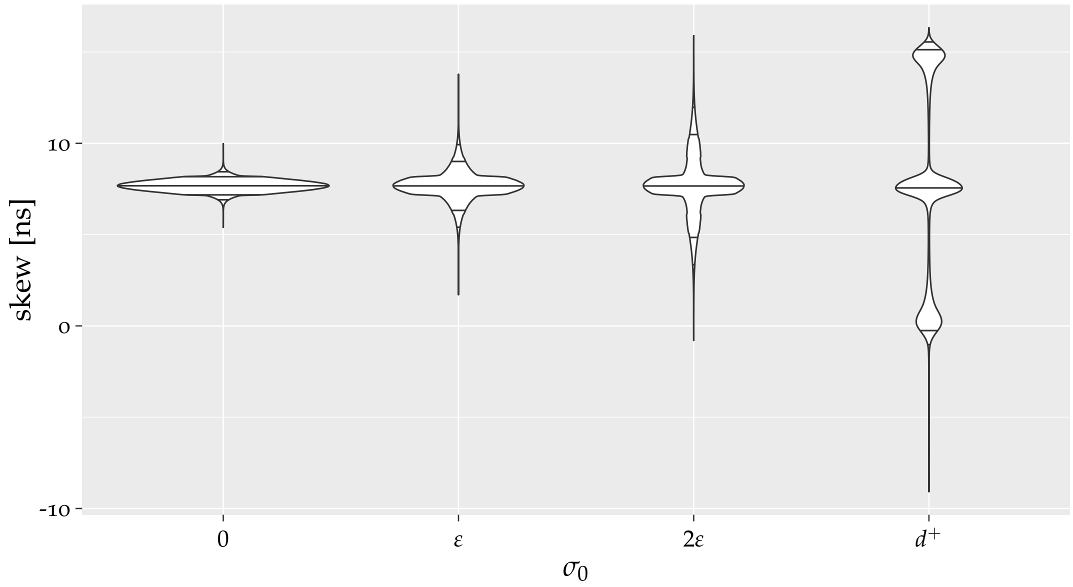
(b) inter-layer skew

Figure A.6: Skew in $\mathcal{G}_A$ with $\mathcal{A}^2_{\text{adj}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the inter-layer skews have been limited to be within $[-10, 25]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



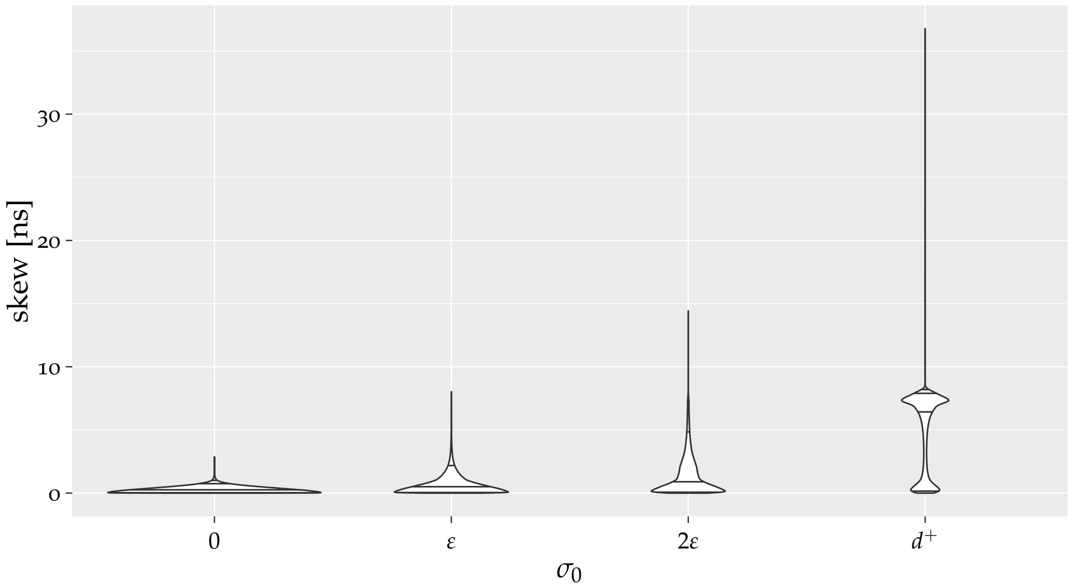
(a) intra-layer skew



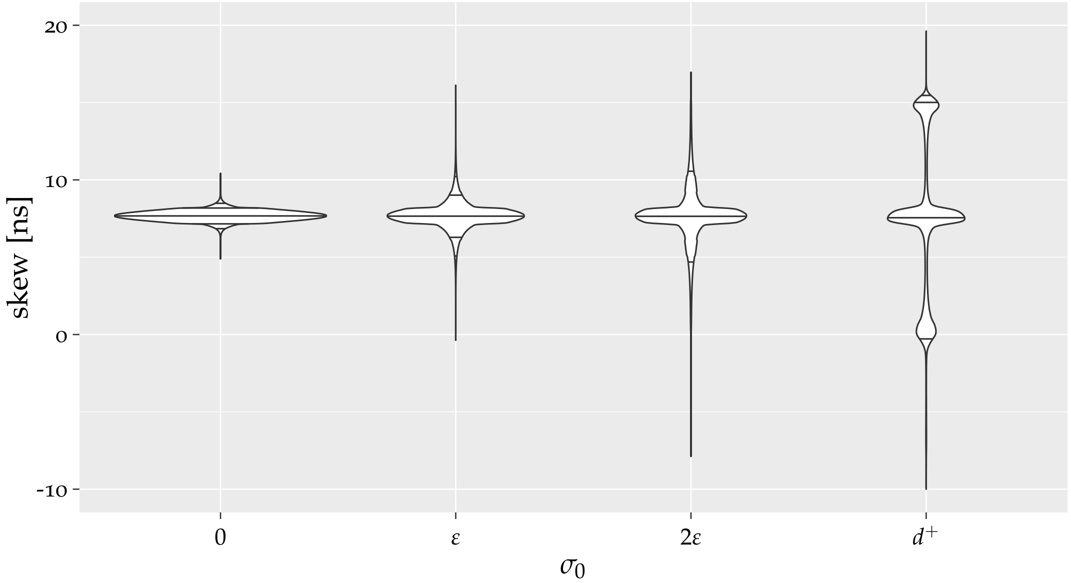
(b) inter-layer skew

Figure A.7: Skew in $\mathcal{G}_A$ with $\mathcal{A}^2_{\text{any}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



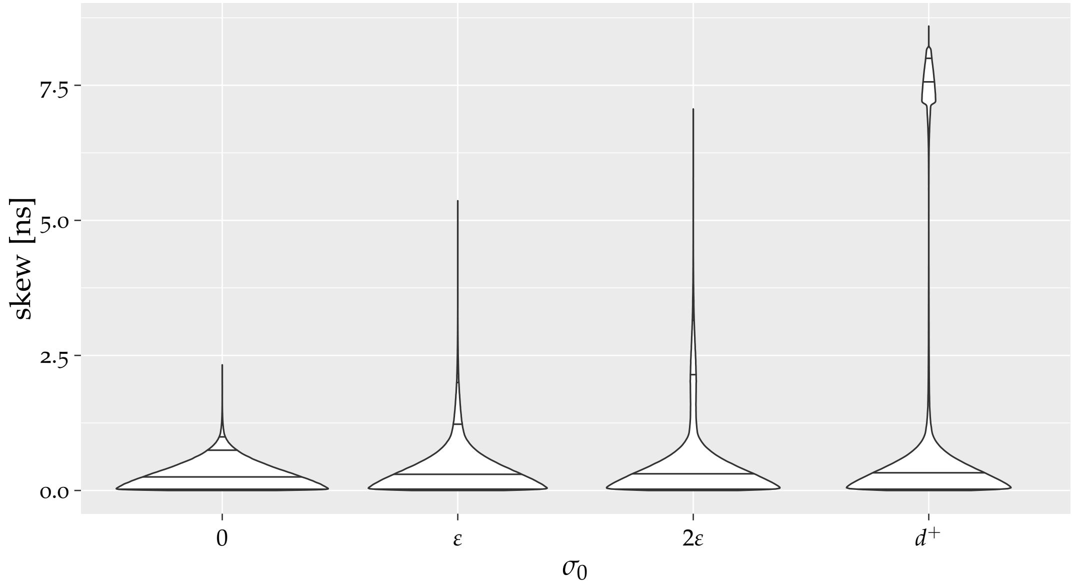
(a) intra-layer skew



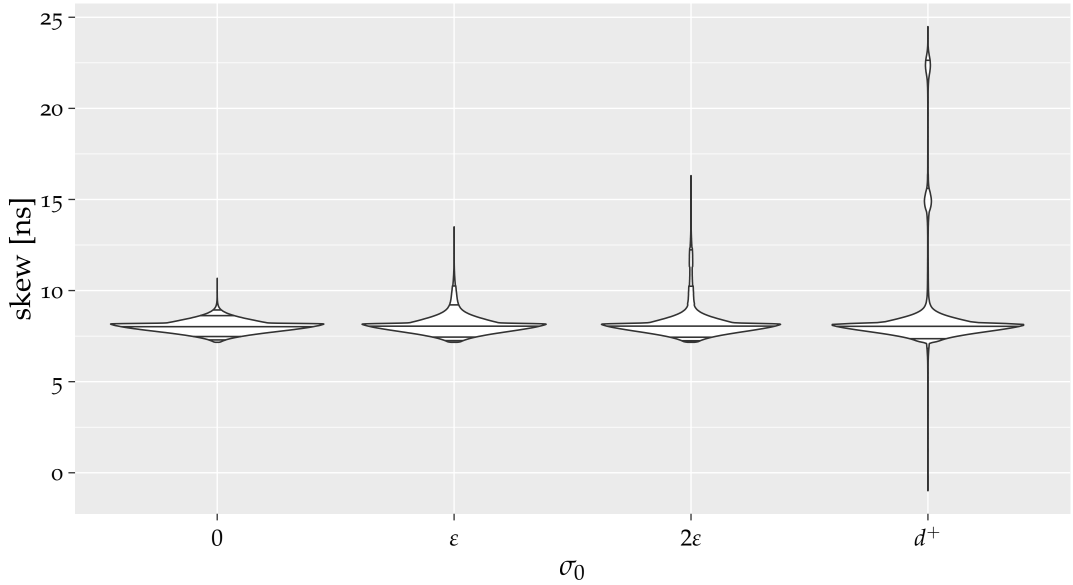
(b) inter-layer skew

Figure A.8: Skew in $\mathcal{G}_A$ with $\mathcal{A}^2_{\text{any}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skew

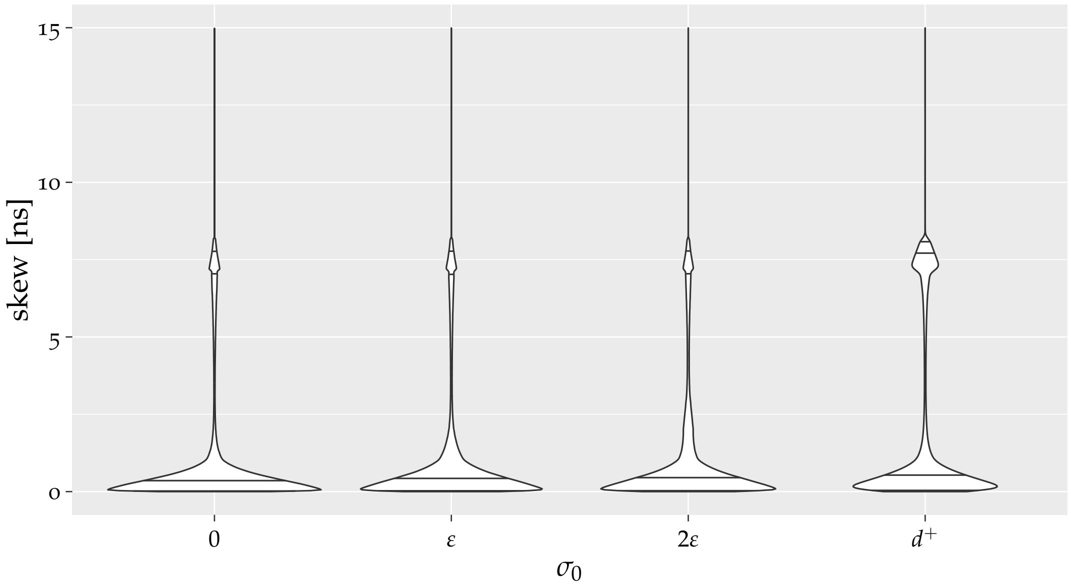


(b) inter-layer skew

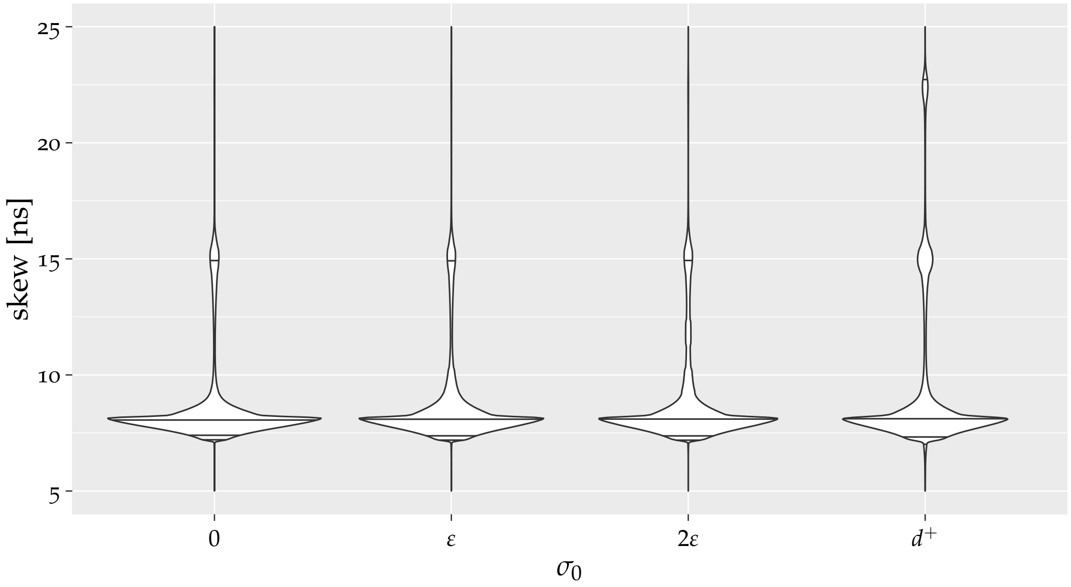
Figure A.9: Skew in $\mathcal{G}_A$ with $\mathcal{A}^3_{\text{any}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



(a) intra-layer skew

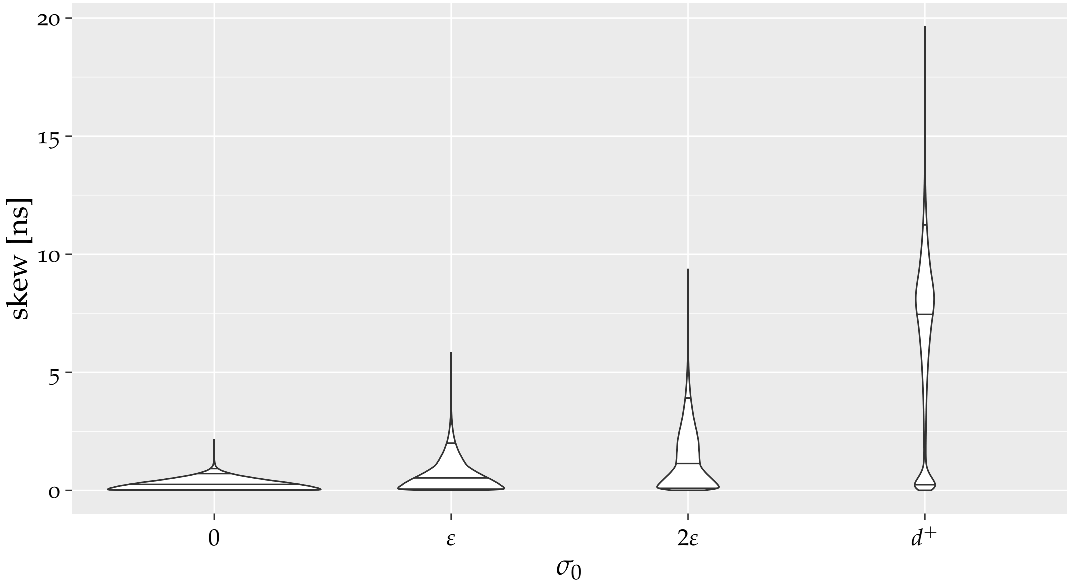


(b) inter-layer skew

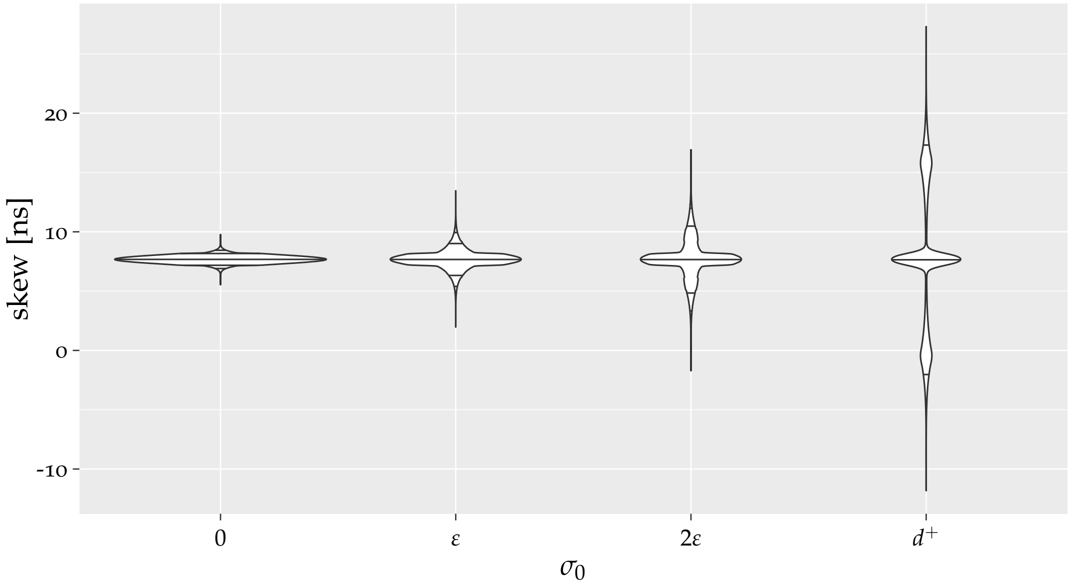
Figure A.10: Skew in $\mathcal{G}_A$ with $\mathcal{A}^3_{\text{any}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[5, 25]$ ns.

A.3 Plots of the Pulse-based Algorithms in $\mathcal{G}_B$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub



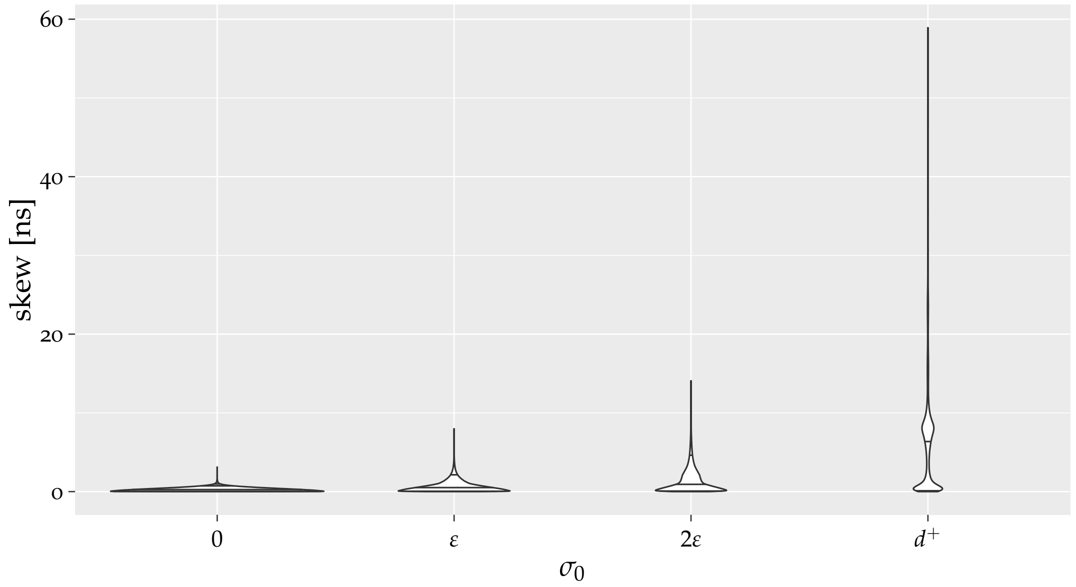
(a) intra-layer skew



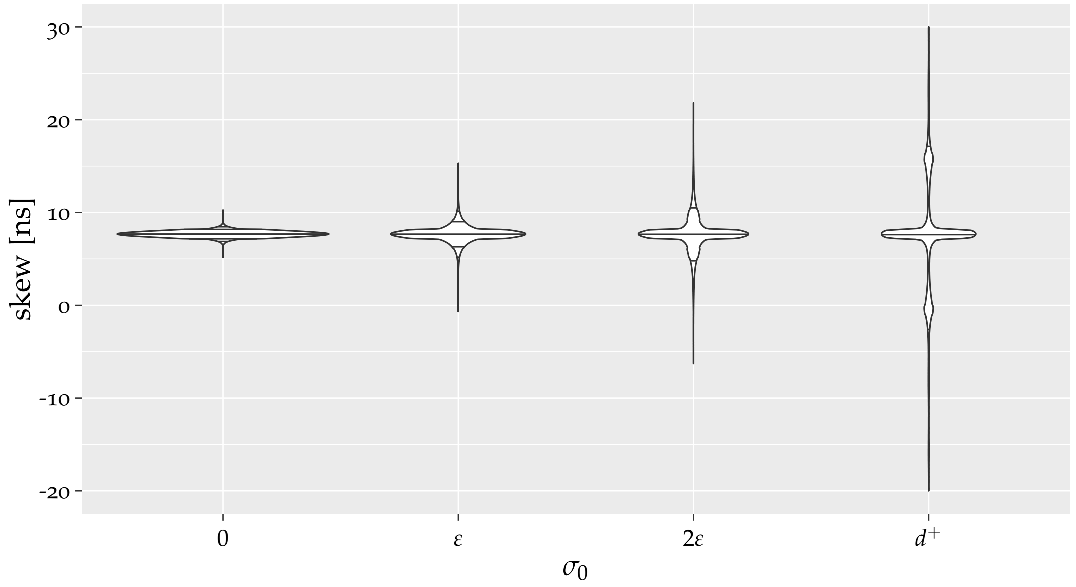
(b) inter-layer skew

Figure A.11: Skew in $\mathcal{G}_B$ with $\mathcal{A}^2_{\text{any}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub



(a) intra-layer skew



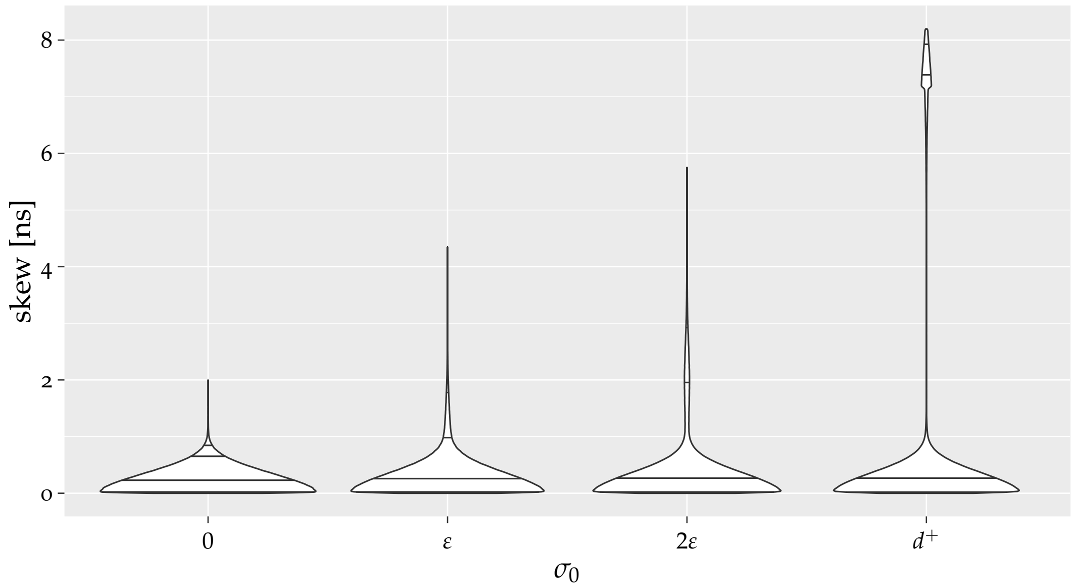
(b) inter-layer skew

Figure A.12: Skew in $\mathcal{G}_B$ with $\mathcal{A}^2_{\text{any}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the inter-layer skews have been limited to be within $[-20, 30]$ ns.

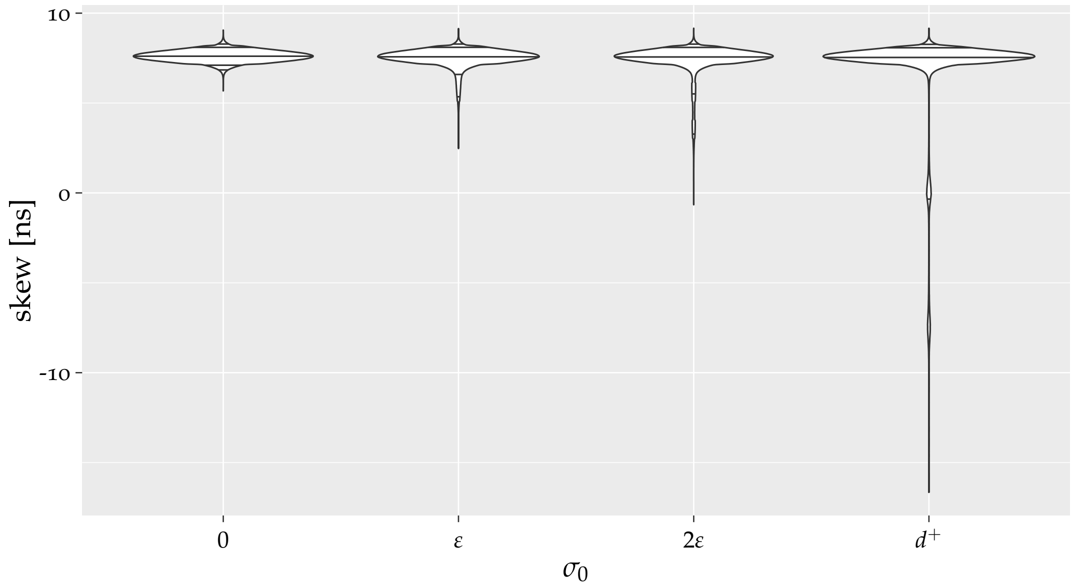
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

A.4 Plots of the Pulse-Based Algorithms in $\mathcal{G}_C$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



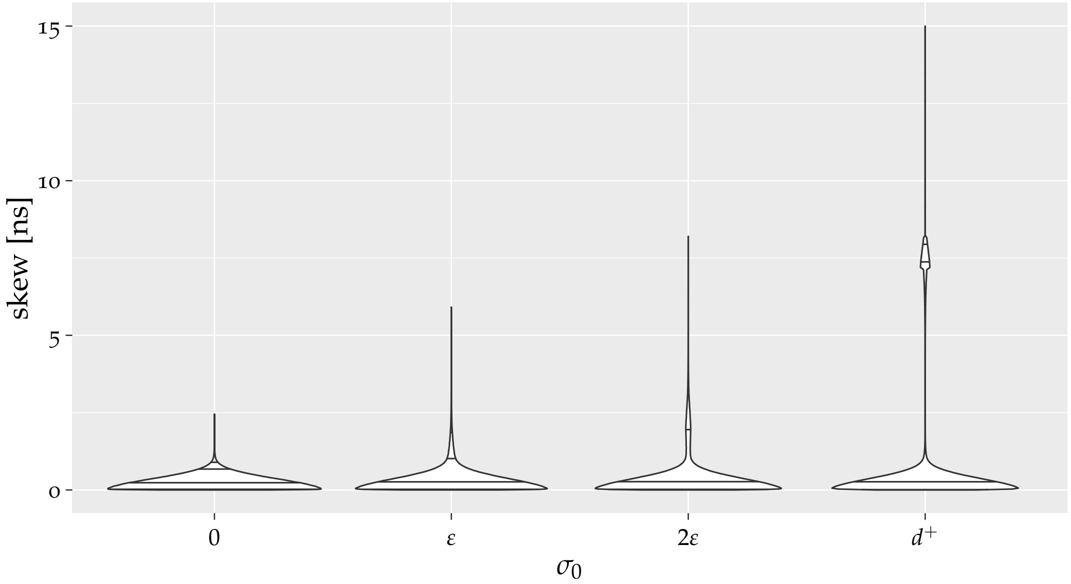
(a) intra-layer skew



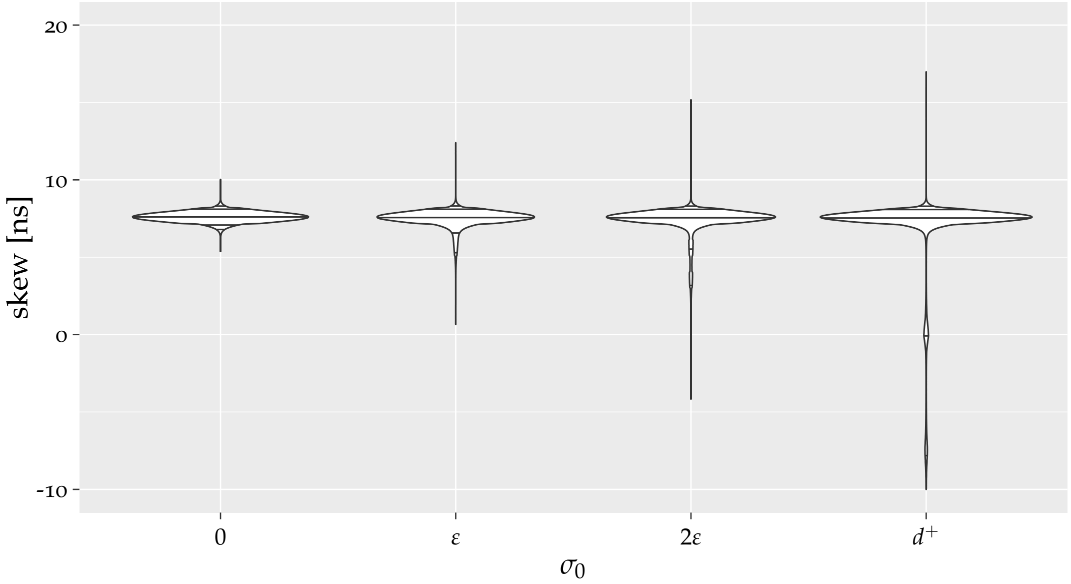
(b) inter-layer skew

Figure A.13: Skew in $\mathcal{G}_C$ with $\mathcal{A}^2_{\text{adj}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



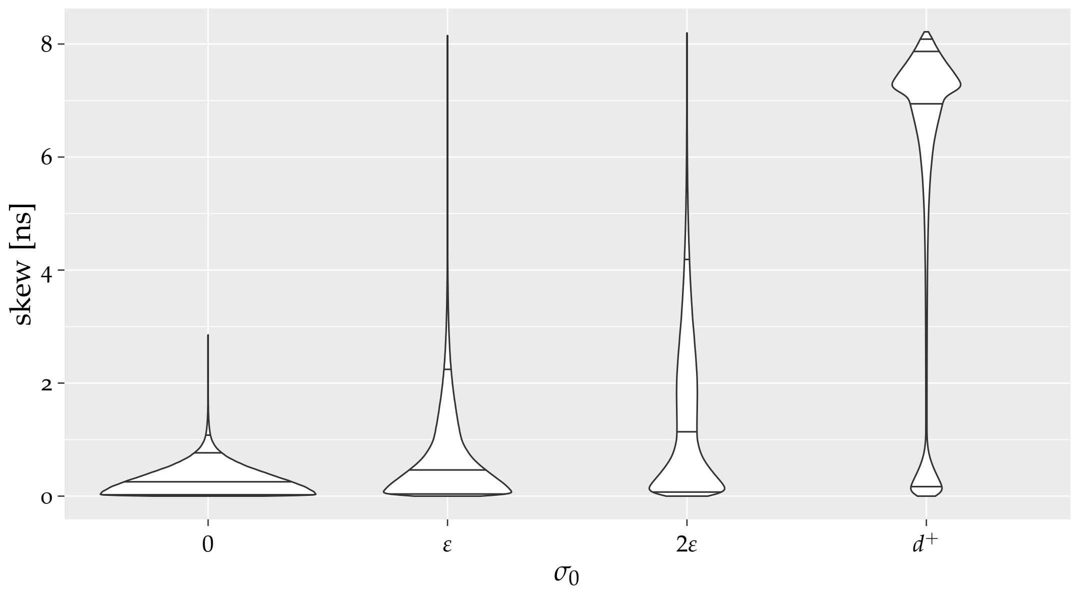
(a) intra-layer skew



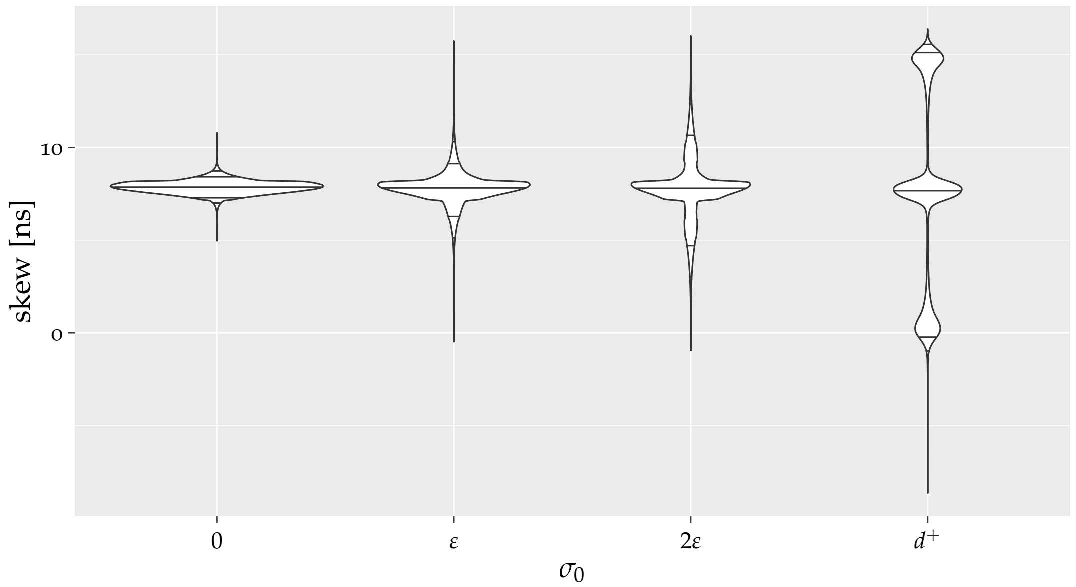
(b) inter-layer skew

Figure A.14: Skew in $\mathcal{G}_C$ with $\mathcal{A}^2_{\text{adj}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



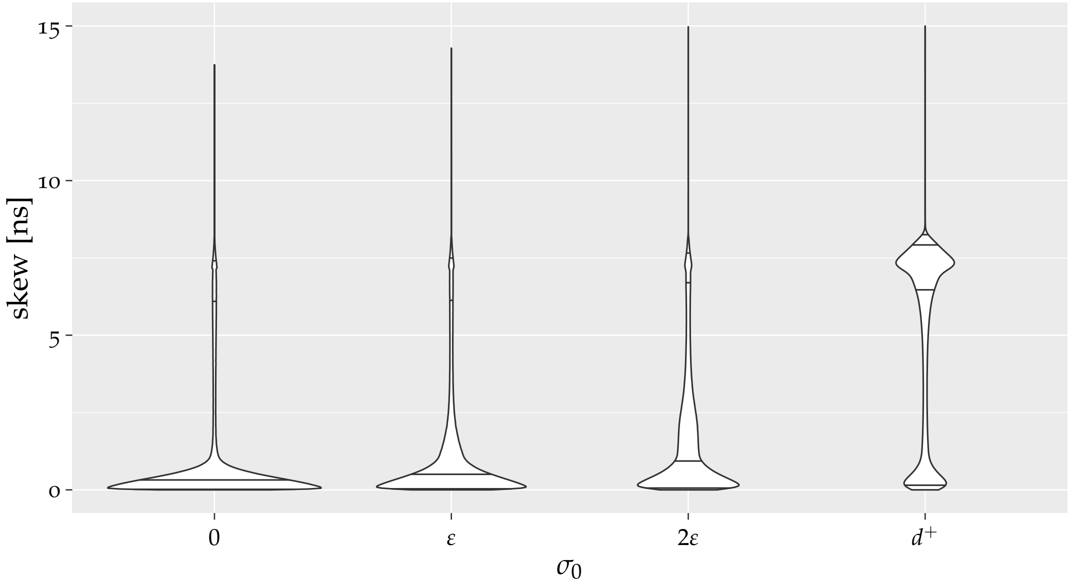
(a) intra-layer skew



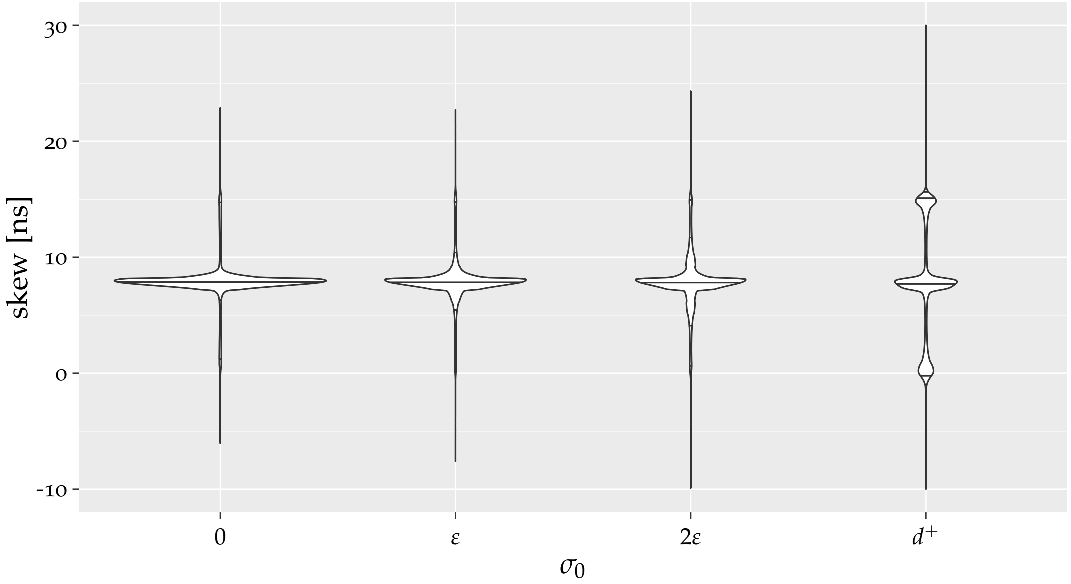
(b) inter-layer skew

Figure A.15: Skew in $\mathcal{G}_C$ with $\mathcal{A}^3_{\text{adj}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



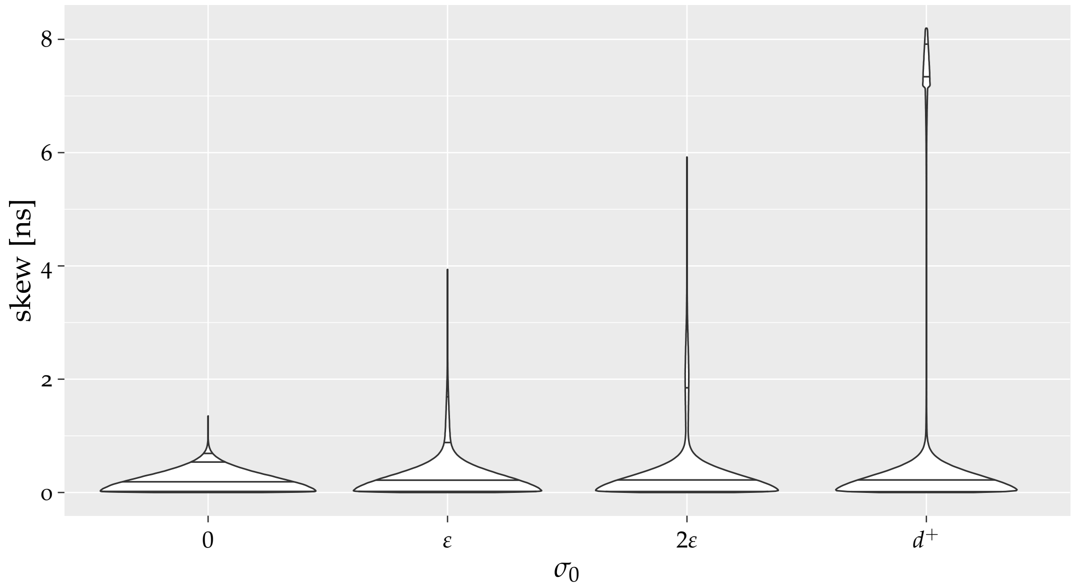
(a) intra-layer skew



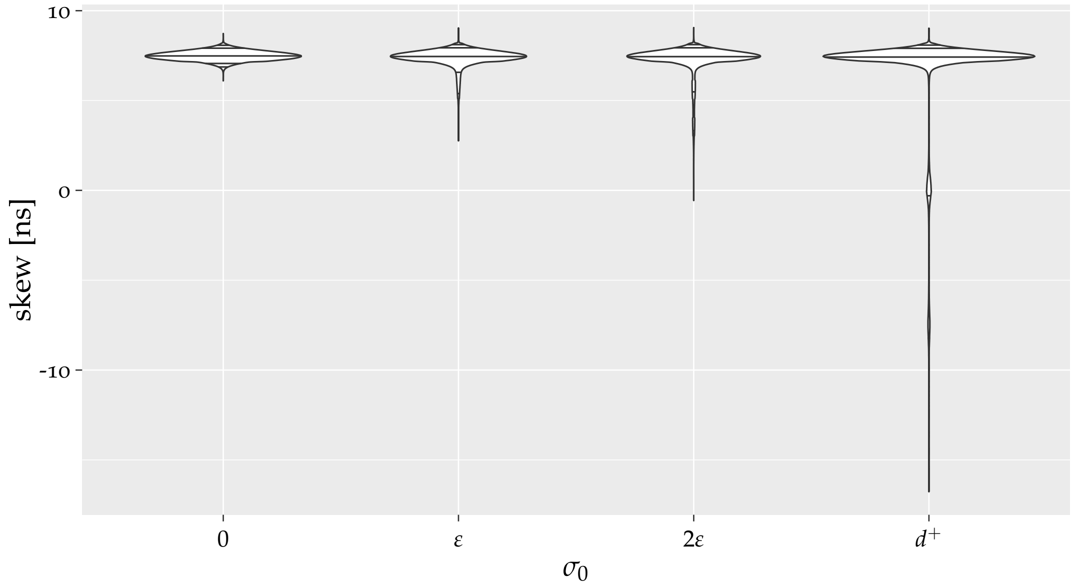
(b) inter-layer skew

Figure A.16: Skew in $\mathcal{G}_C$ with $\mathcal{A}^3_{\text{adj}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-10, 30]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skew

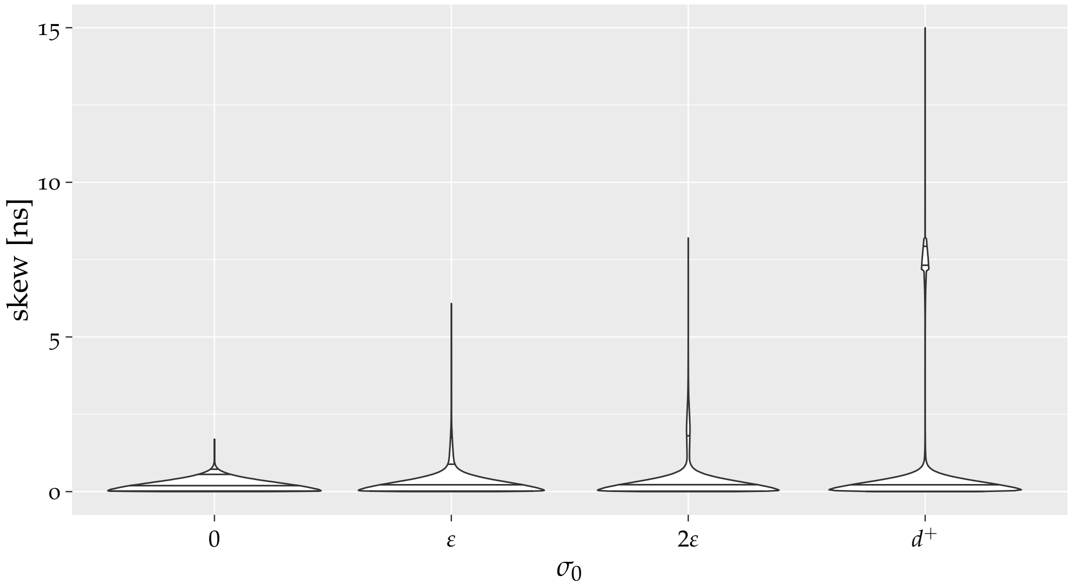


(b) inter-layer skew

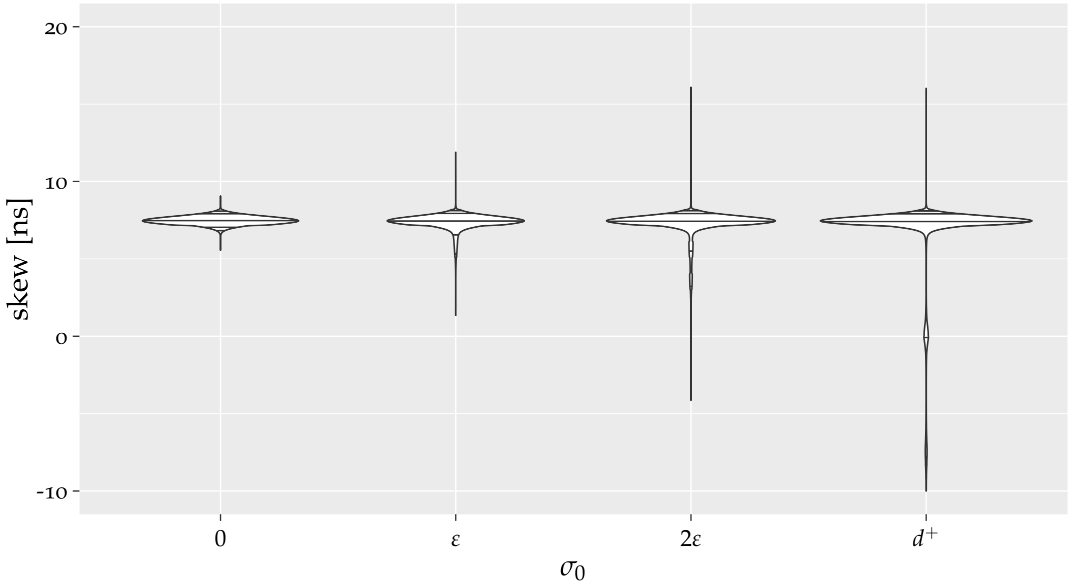
Figure A.17: Skew in $\mathcal{G}_C$ with $\mathcal{A}^2_{\text{any}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

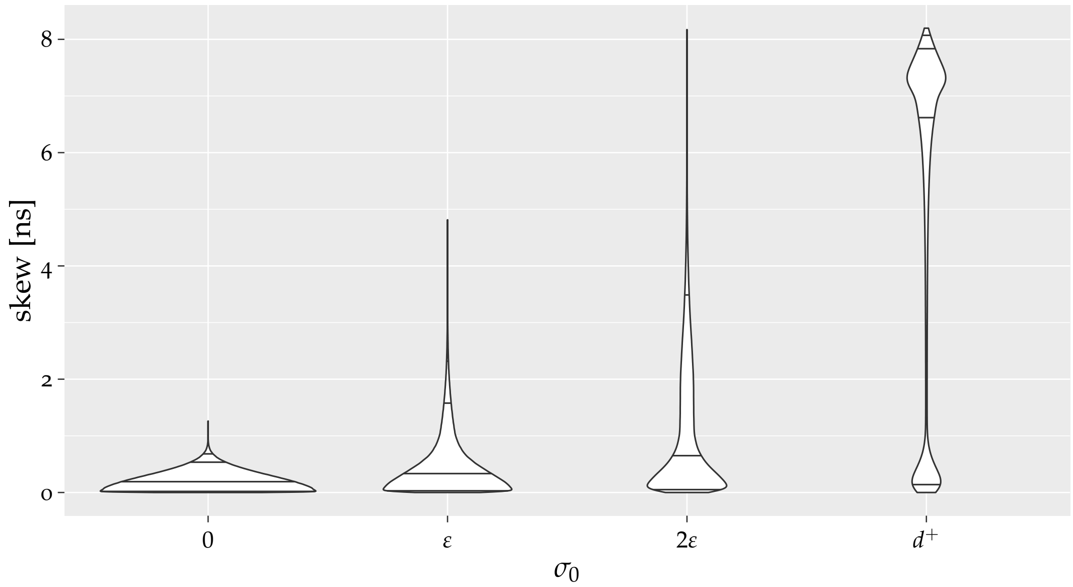


(a) intra-layer skew

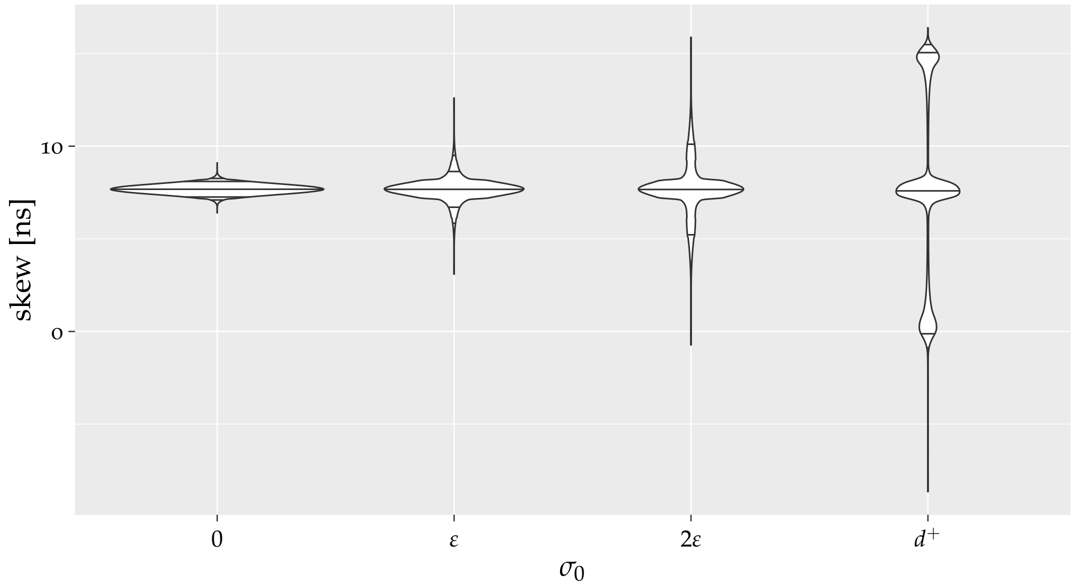


(b) inter-layer skew

Figure A.18: Skew in $\mathcal{G}_C$ with $\mathcal{A}^2_{\text{any}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.



(a) intra-layer skew

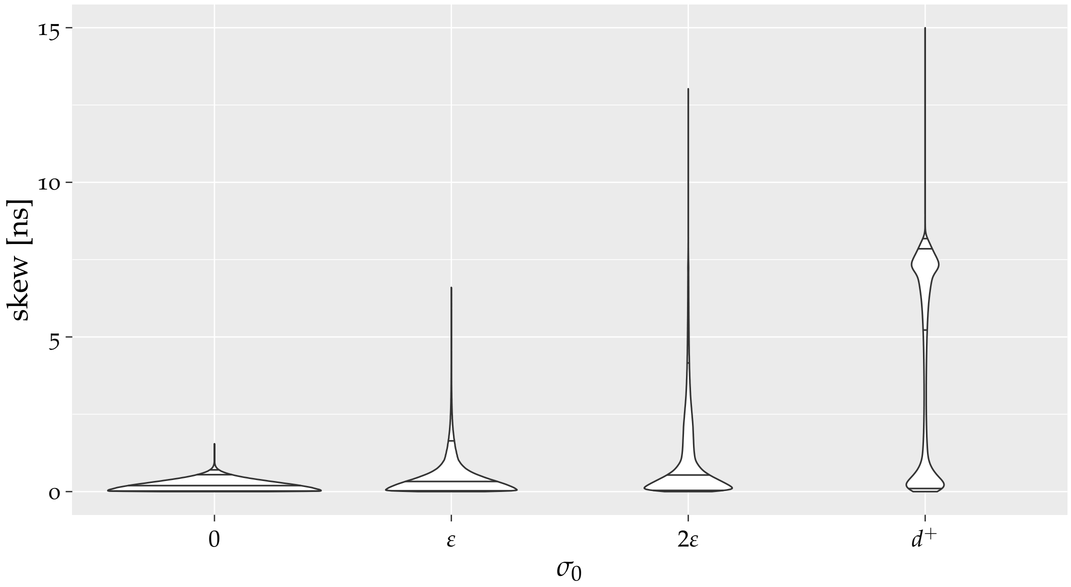


(b) inter-layer skew

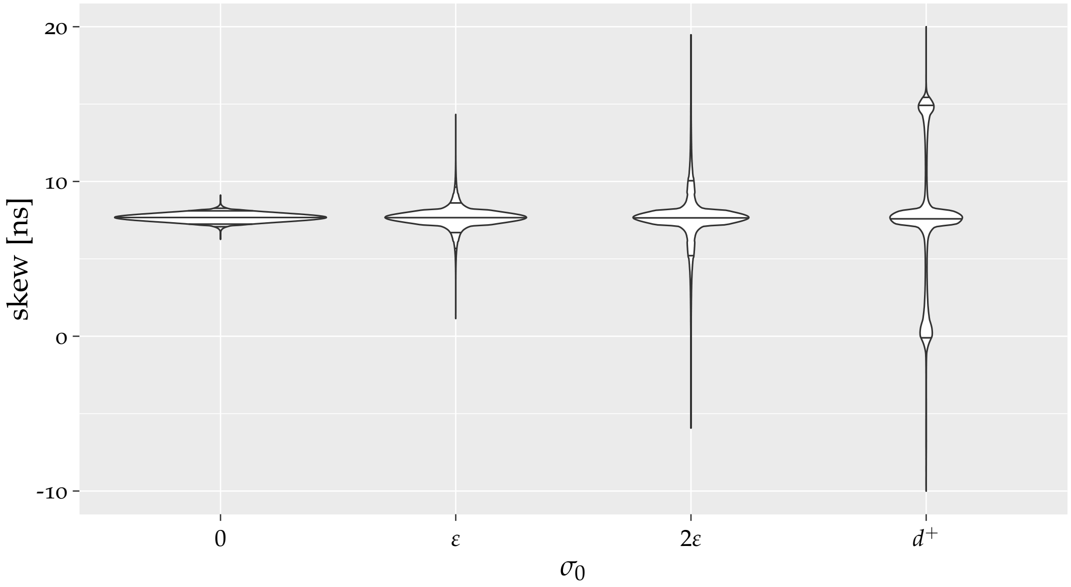
Figure A.19: Skew in $\mathcal{G}_C$ with $\mathcal{A}^3_{\text{any}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skew



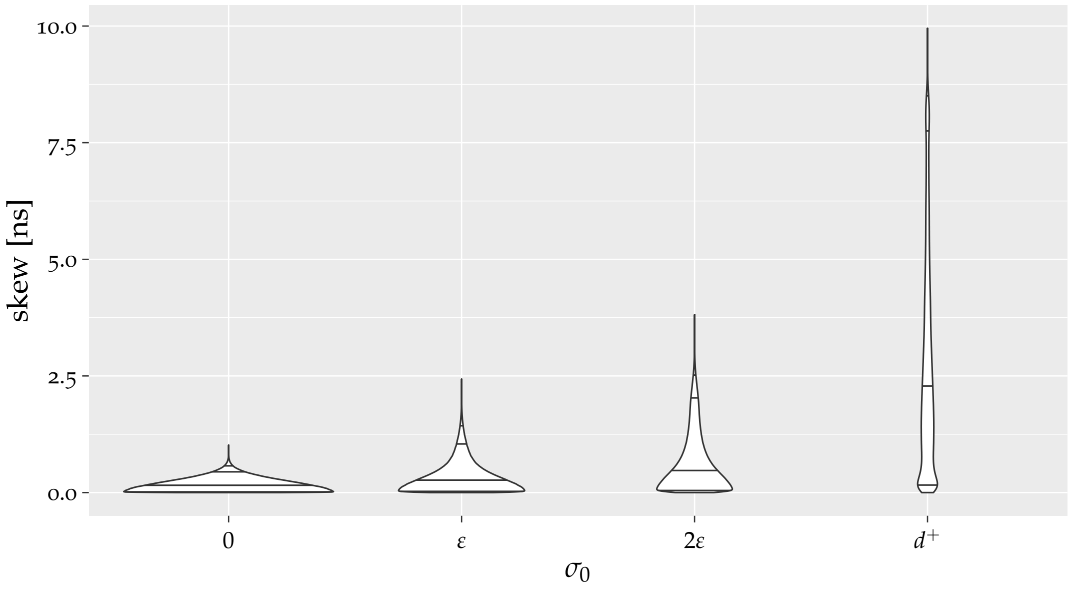
(b) inter-layer skew

Figure A.20: Skew in $\mathcal{G}_C$ with $\mathcal{A}^3_{\text{any}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

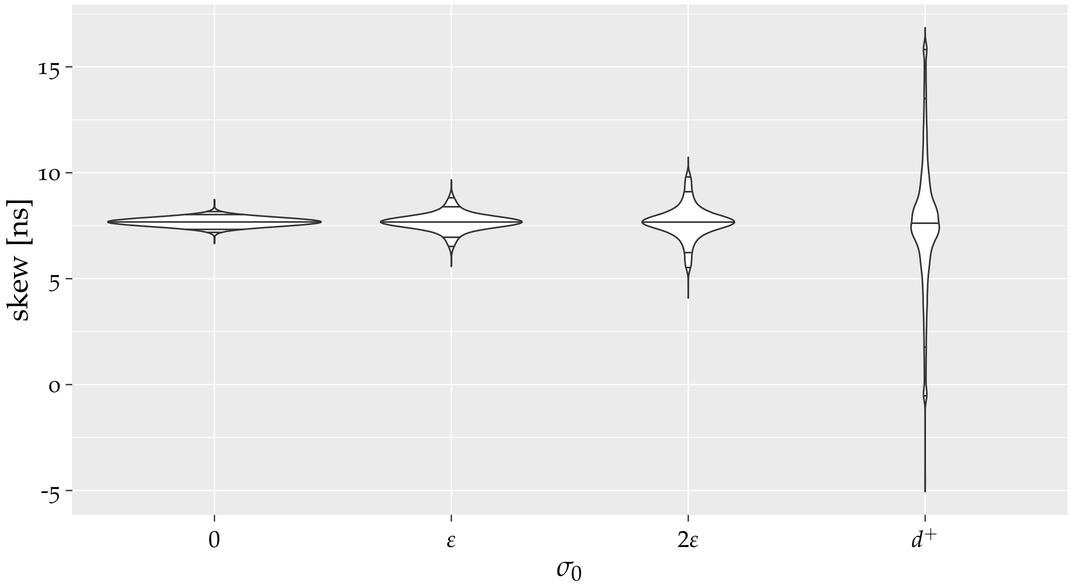
A.5 Plots of the Time-based Algorithms with $d = 2$ in $\mathcal{G}_C$

In this section we present the plots of the algorithms $\mathcal{A}^2_{\text{avg},\,25\,\text{ns}}$ and $\mathcal{A}^2_{\text{mid},\,25\,\text{ns}}$. We omit the common offset of 25 ns in the algorithm description. Further, recall that the provided inter-layer skews have been compensated by this offset 25 ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub



(a) intra-layer skew

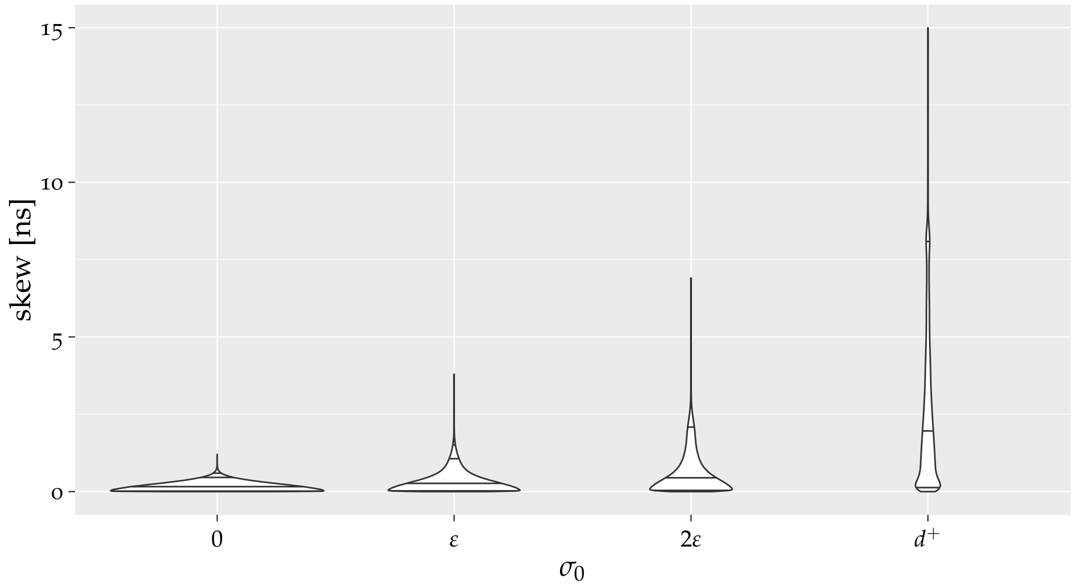


(b) inter-layer skew

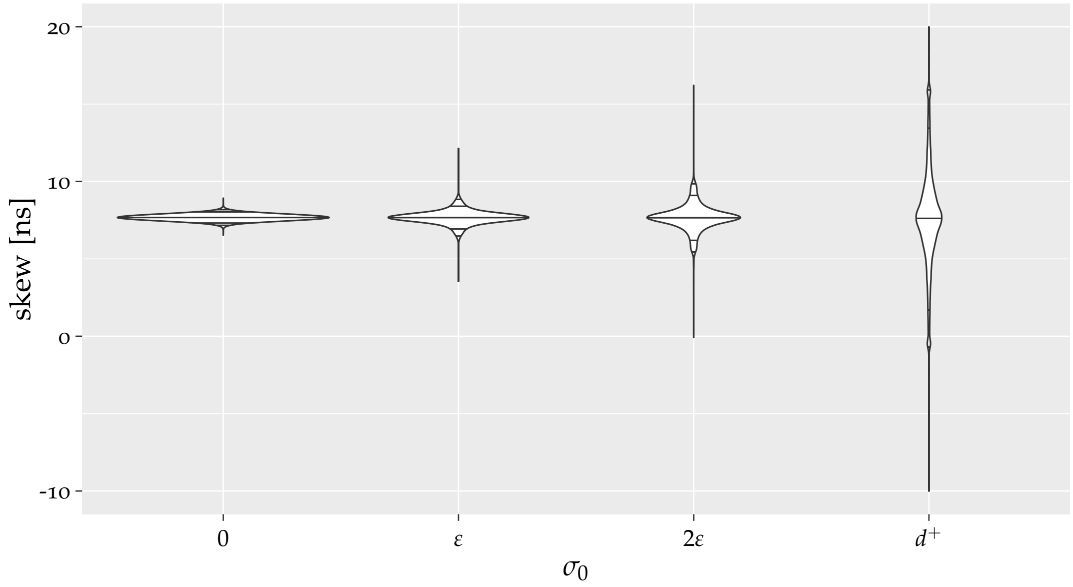
Figure A.21: Skew in $\mathcal{G}_C$ with $\mathcal{A}^2_{\text{mid},\,25\,\text{ns}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

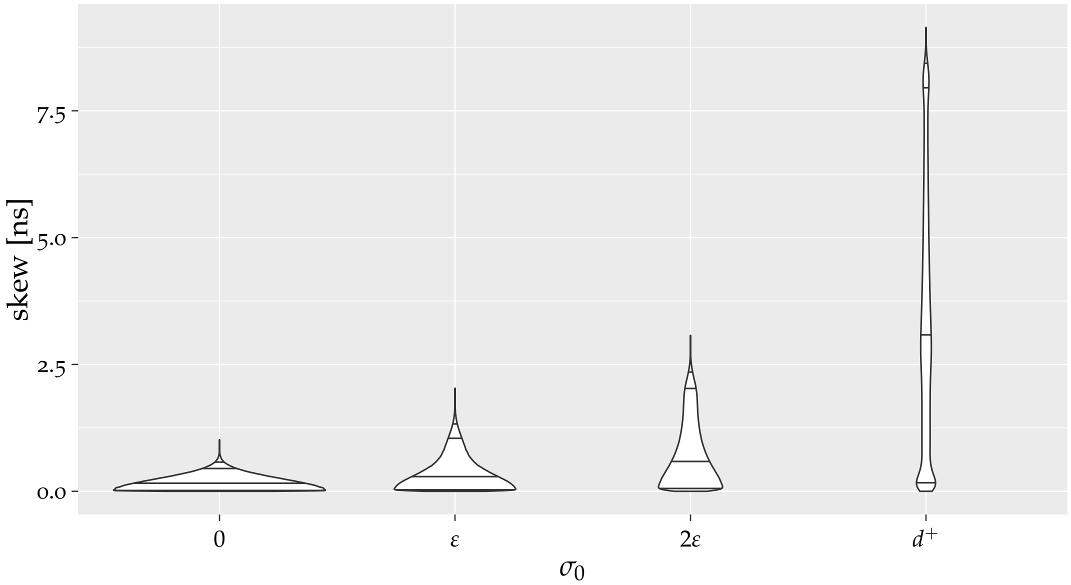


(a) intra-layer skew

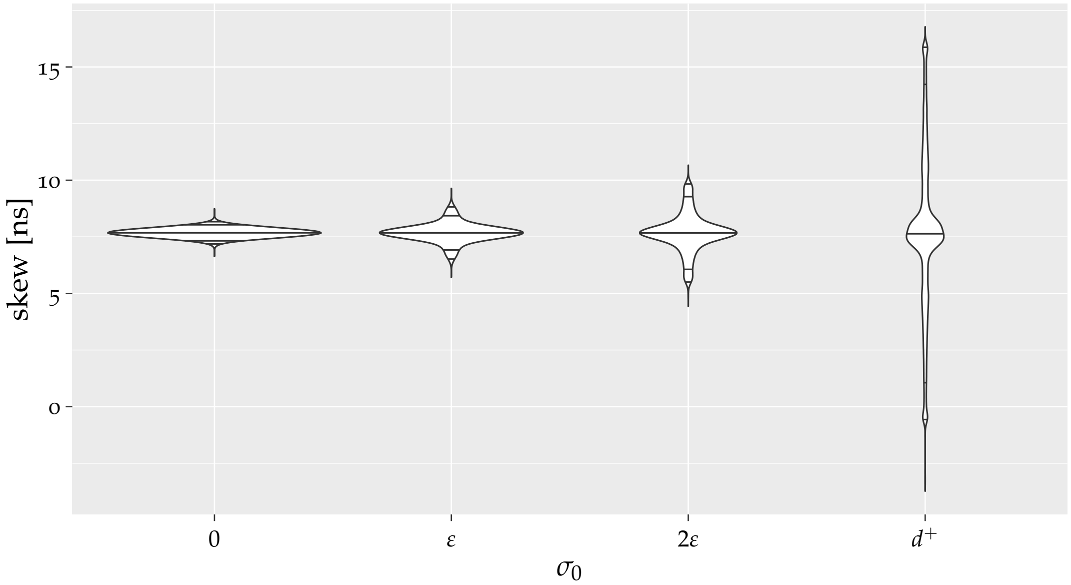


(b) inter-layer skew

Figure A.22: Skew in $\mathcal{G}_C$ with $\mathcal{A}^2_{\text{mid}, 25\,\text{ns}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.



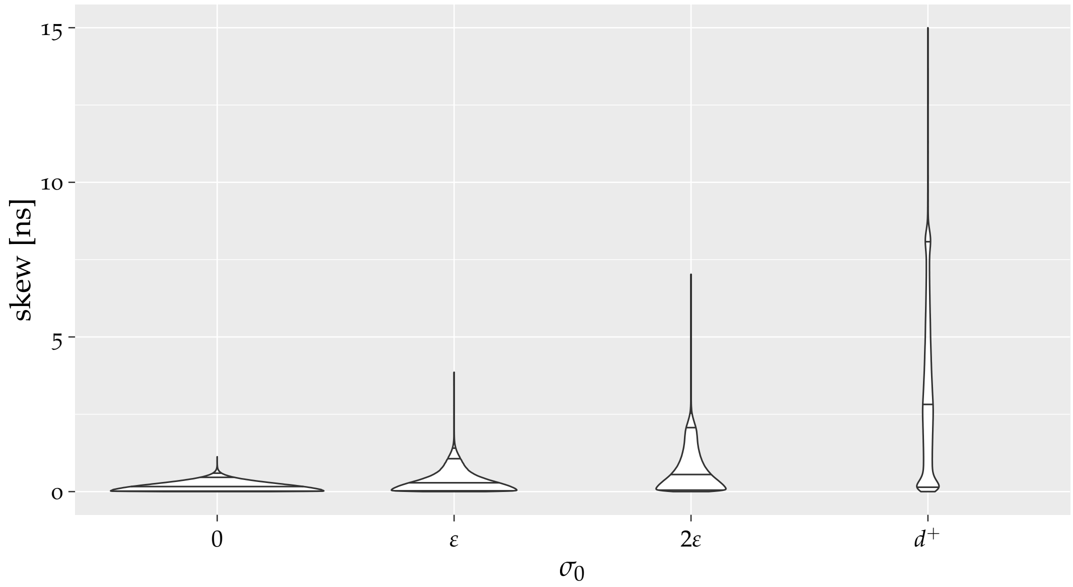
(a) intra-layer skew



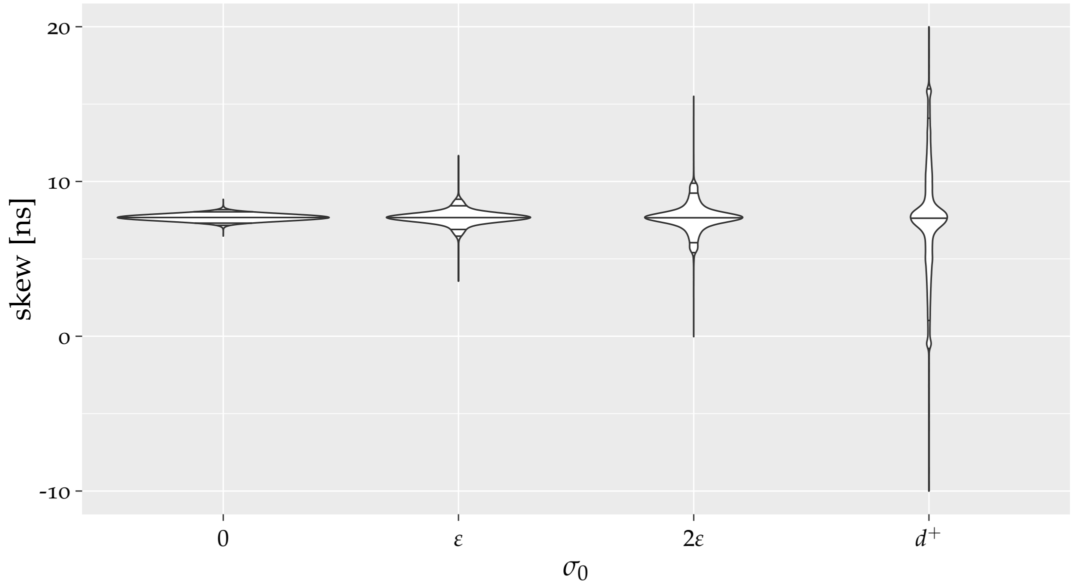
(b) inter-layer skew

Figure A.23: Skew in $\mathcal{G}_C$ with $\mathcal{A}^2_{\text{avg},\,25\,\text{ns}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub



(a) intra-layer skew



(b) inter-layer skew

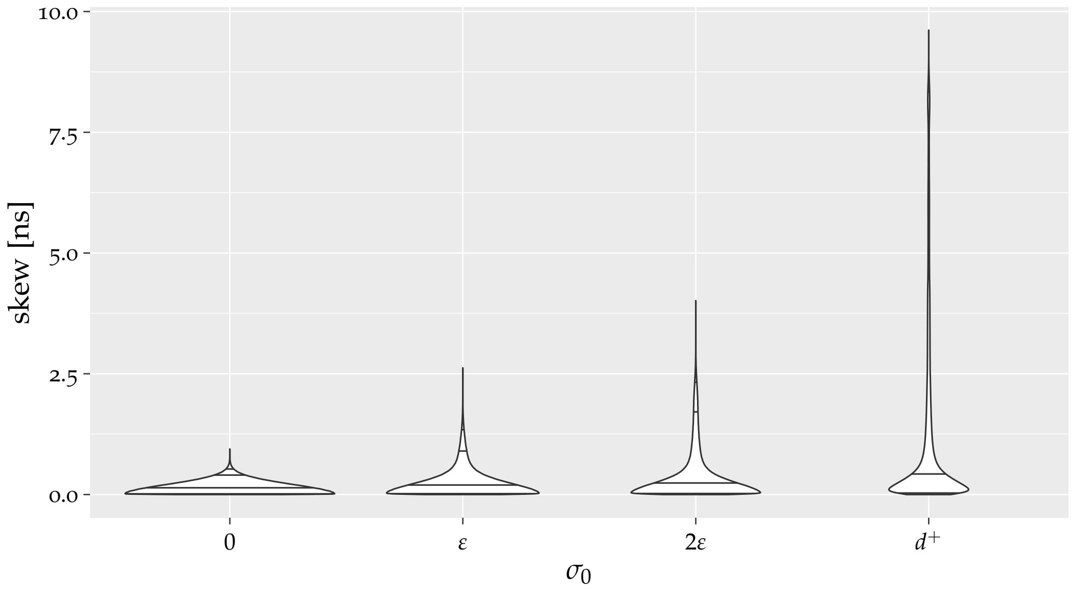
Figure A.24: Skew in $\mathcal{G}_C$ with $\mathcal{A}^2_{\text{avg},\,25\,\text{ns}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Wien Bibliothek Your knowledge hub

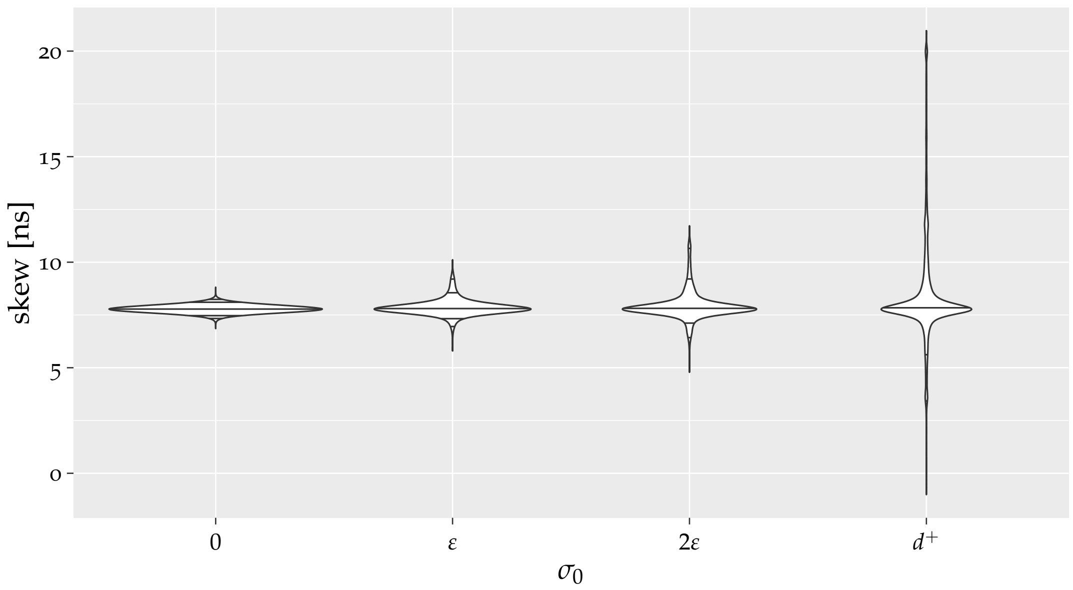
A.6 Plots of the Time-based Algorithms with $d = 1$ in $\mathcal{G}_C$

In this section we present the plots of the algorithms $\mathcal{A}^1_{\text{avg}, 25\,\text{ns}}$ and $\mathcal{A}^1_{\text{mid}, 25\,\text{ns}}$. We omit the common offset of 25 ns in the algorithm description. Further, recall that the provided inter-layer skews have been compensated by 25 ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



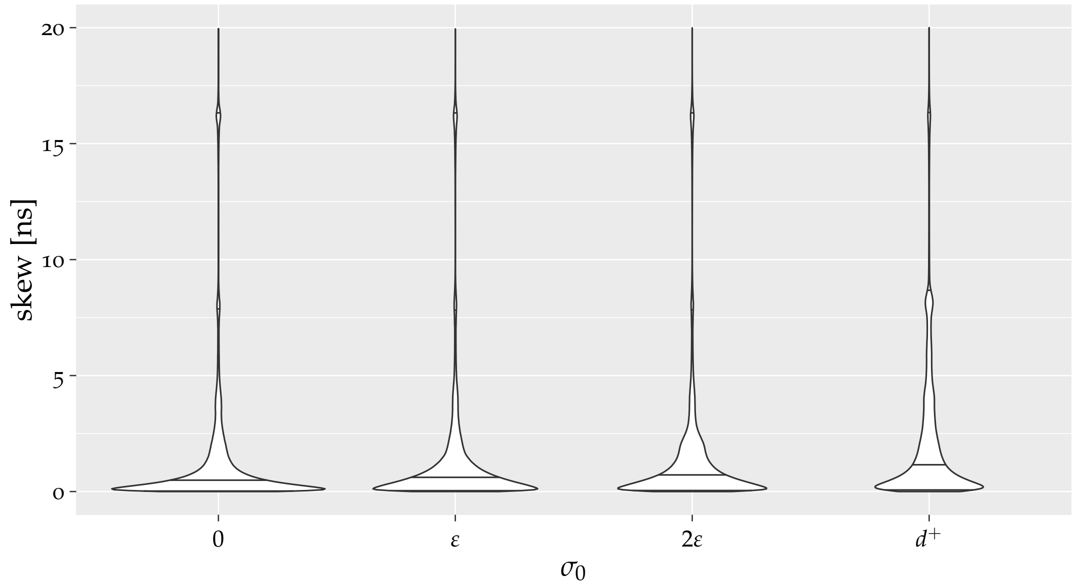
(a) intra-layer skews



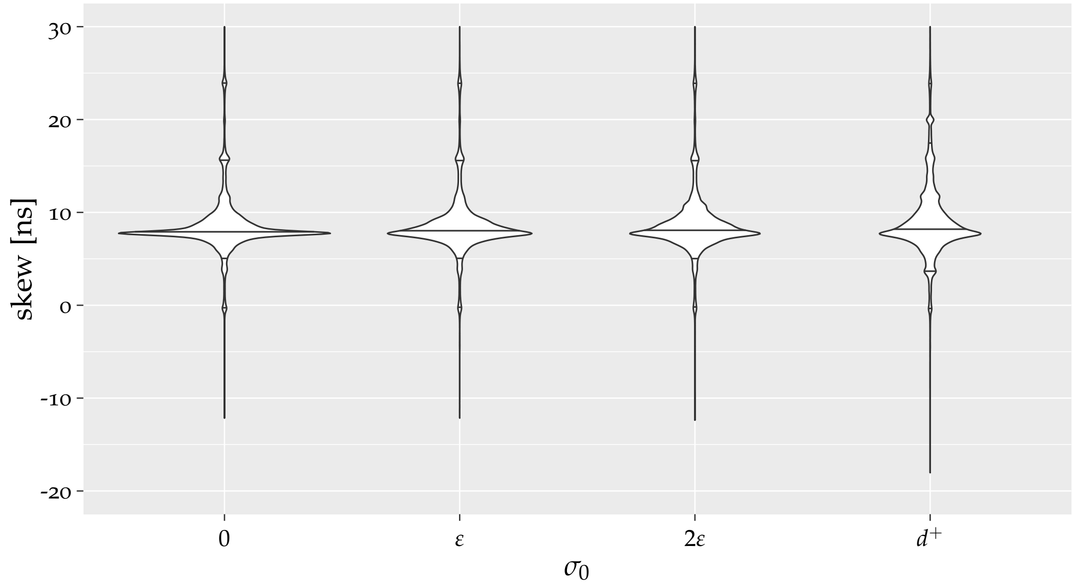
(b) inter-layer skews

Figure A.25: Skew in $\mathcal{G}_C$ with $\mathcal{A}^1_{\mathrm{mid},\,25\,\mathrm{ns}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



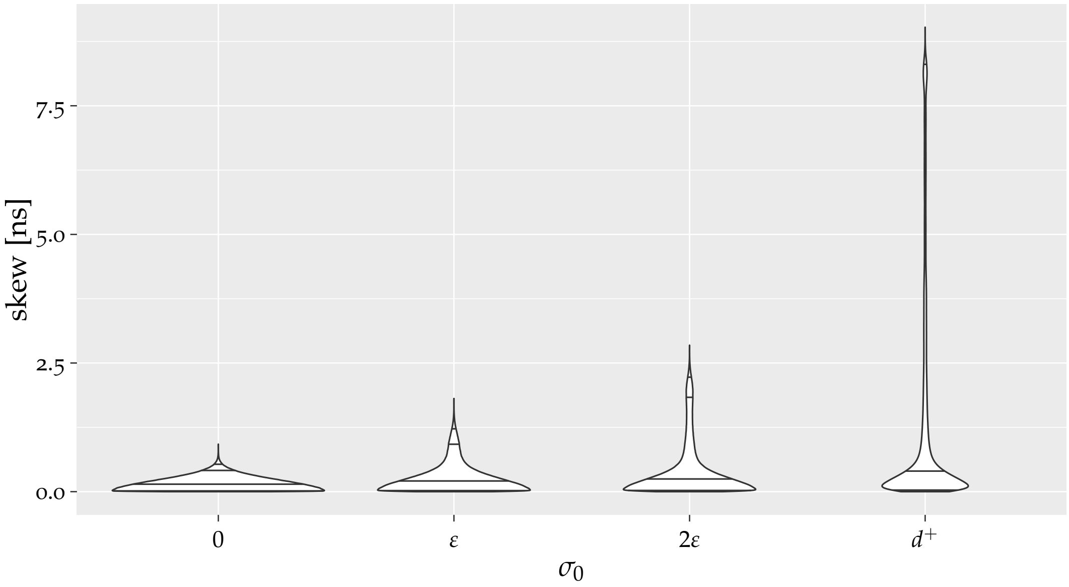
(a) intra-layer skews



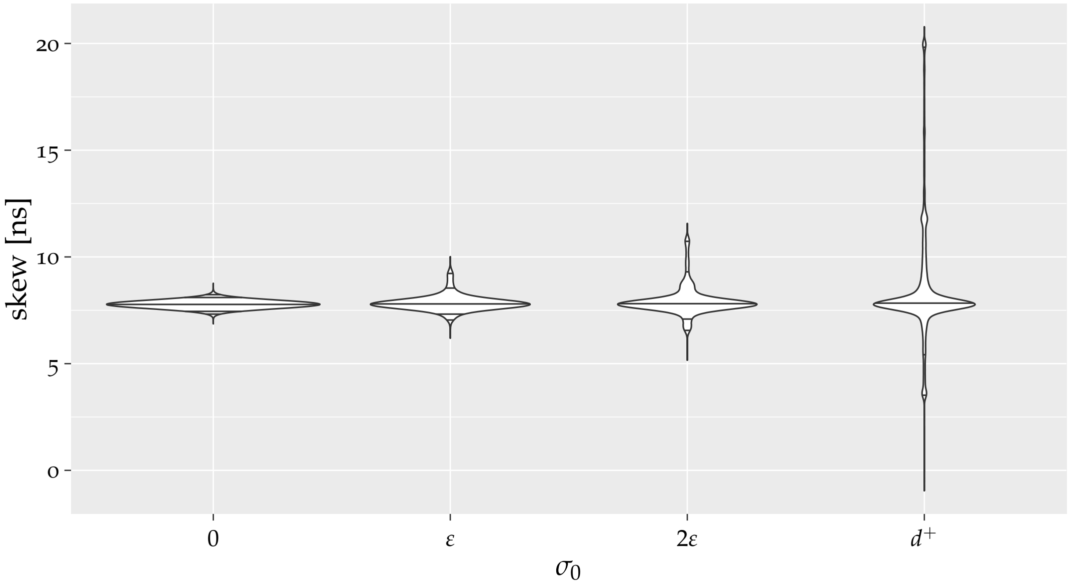
(b) inter-layer skews

Figure A.26: Skew in $\mathcal{G}_C$ with $\mathcal{A}^1_{\mathrm{mid},\,25\,\mathrm{ns}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-20, 30]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



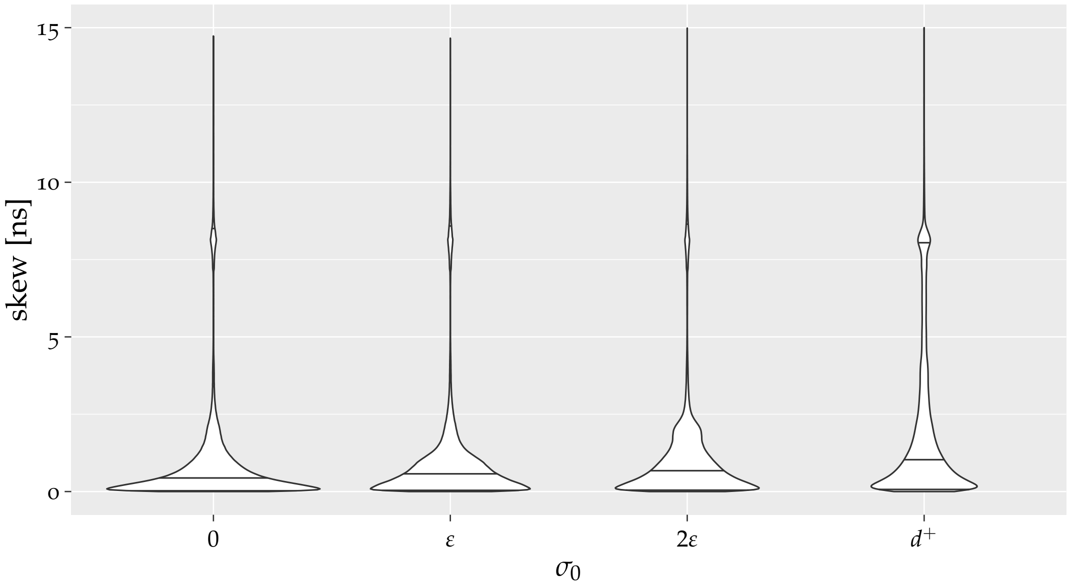
(a) intra-layer skews



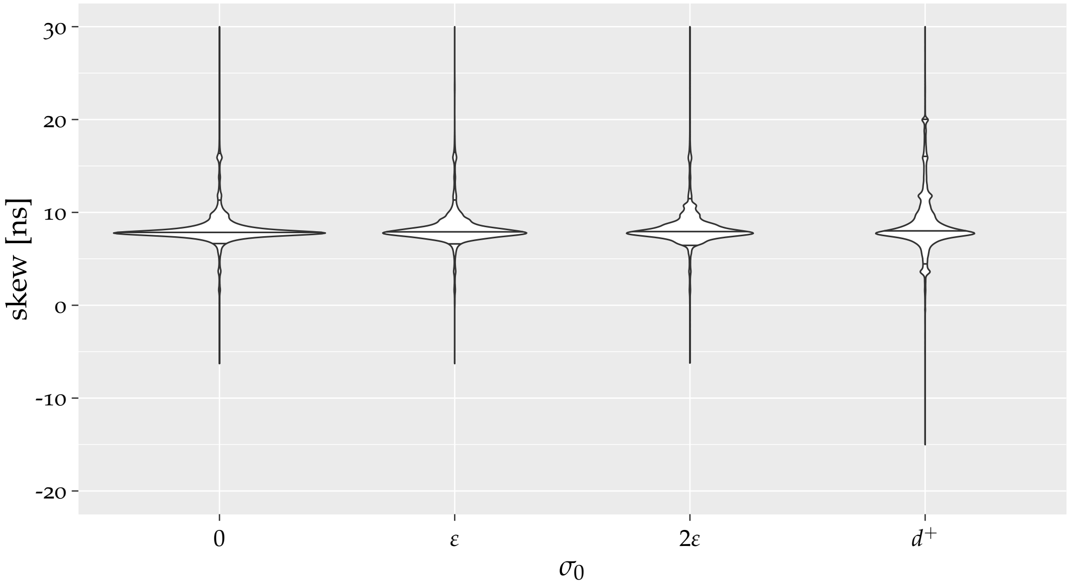
(b) inter-layer skews

Figure A.27: Skew in $\mathcal{G}_C$ with $\mathcal{A}^1_{\text{avg}, 25\,\text{ns}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skews



(b) inter-layer skews

Figure A.28: Skew in $\mathcal{G}_C$ with $\mathcal{A}^1_{\text{avg},\,25\,\text{ns}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-20, 30]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

A.7 Tables of the Time-based Algorithms with $d = 1$ in $\mathcal{G}_C$

In this section we present the tables of the algorithms $\mathcal{A}^1_{\text{avg},\,25\,\text{ns}}$ and $\mathcal{A}^1_{\text{mid},\,25\,\text{ns}}$. We omit the common offset of 25 ns in the algorithm description. Further, recall that the provided inter-layer skews have been compensated by 25 ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Wien Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.1: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology with time-based algorithms parameterized with $d = 1$ and input skews 0 and ε. The inter-layer skews have been compensated by 25 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^1_{\text{mid},\,25\,\text{ns}}$	0	0	0	0.00	0.01	0.16	0.14	0.40	0.53	0.94	6.86	7.34	7.47	7.78	7.78	8.10	8.24	8.81
		15	0	0.01	0.04	1.62	0.49	7.84	16.34	22.08	−12.14	−0.23	5.10	8.75	7.95	15.76	24.24	55.39
		15	1	0.01	0.03	1.22	0.46	4.65	14.84	19.00	−6.63	2.32	5.70	8.51	7.94	13.51	20.13	49.67
$\mathcal{A}^1_{\text{avg},\,25\,\text{ns}}$	0	0	0	0.00	0.01	0.17	0.14	0.41	0.53	0.92	6.87	7.32	7.45	7.78	7.78	8.10	8.23	8.76
		15	0	0.01	0.03	0.99	0.44	3.71	8.51	16.68	−6.28	3.27	6.68	8.26	7.88	11.37	16.41	47.13
		15	1	0.01	0.03	0.76	0.42	2.32	7.82	10.34	−3.17	5.34	6.88	8.13	7.88	10.17	15.13	45.42
$\mathcal{A}^1_{\text{mid},\,25\,\text{ns}}$	ε	0	0	0.00	0.02	0.28	0.20	0.90	1.34	2.62	5.80	6.96	7.33	7.85	7.81	8.56	9.21	10.10
		15	0	0.01	0.05	1.68	0.62	7.79	16.34	22.02	−12.15	−0.17	5.11	8.78	8.07	15.75	24.28	54.56
		15	1	0.01	0.05	1.29	0.59	4.70	14.77	18.97	−6.56	2.35	5.68	8.55	8.06	13.48	20.15	48.13
$\mathcal{A}^1_{\text{avg},\,25\,\text{ns}}$	ε	0	0	0.00	0.02	0.29	0.21	0.92	1.22	1.81	6.20	7.05	7.33	7.84	7.80	8.55	9.23	10.01
		15	0	0.01	0.05	1.06	0.58	3.71	8.59	16.41	−6.27	3.23	6.65	8.30	7.94	11.38	16.53	46.27
		15	1	0.01	0.05	0.84	0.55	2.37	7.77	10.41	−3.16	5.31	6.84	8.18	7.94	10.24	15.10	43.96

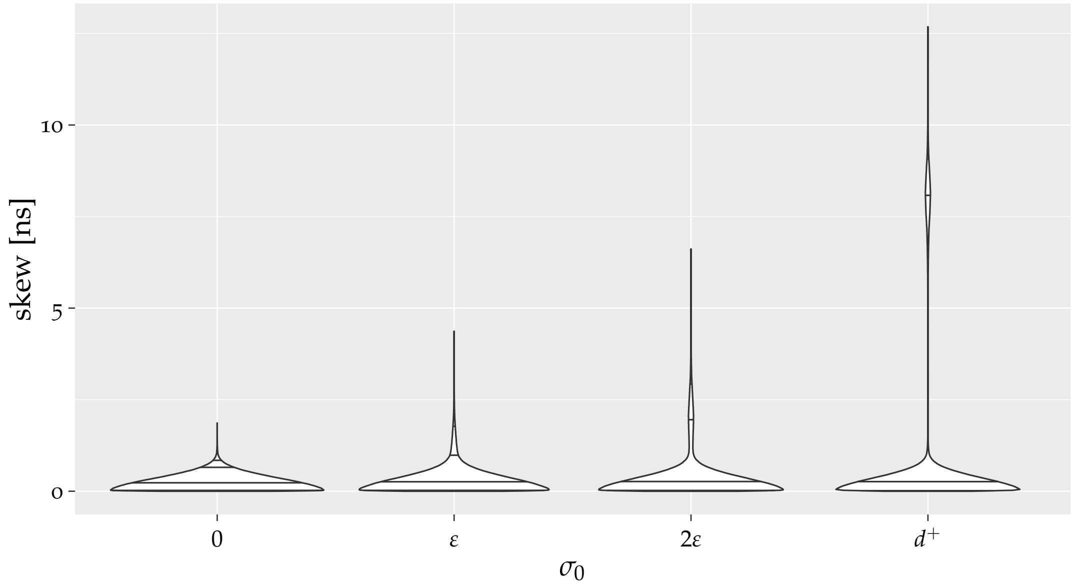
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.2: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology with time-based algorithms parameterized with $d = 1$ and input skews 2ε and d^+. The inter-layer skews have been compensated by 25 ns.

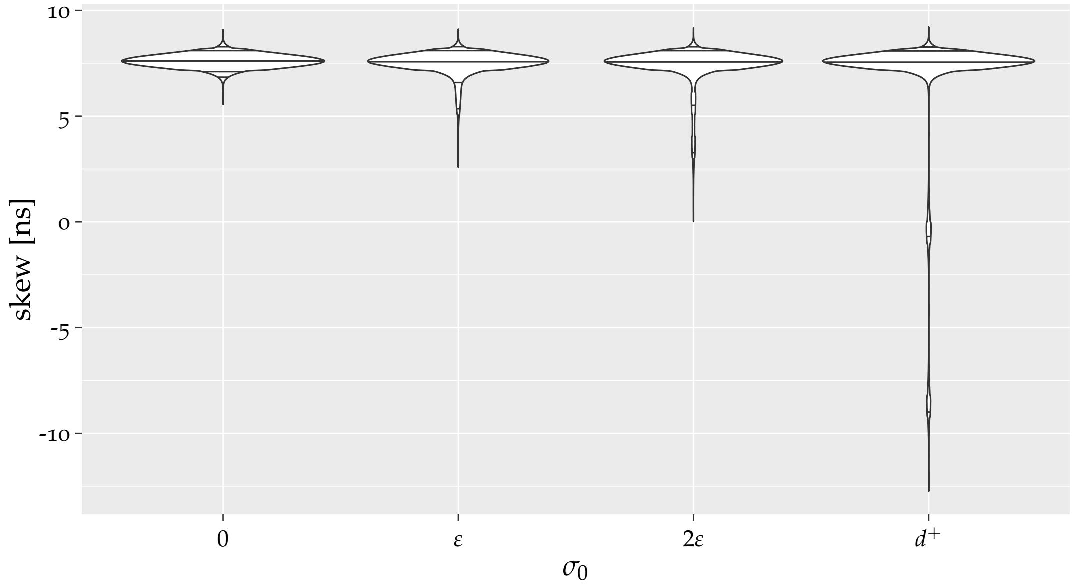
				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^1_{\text{mid}, 25\,\text{ns}}$	2ε	0	0	0.00	0.02	0.43	0.24	1.71	2.33	4.01	4.79	6.44	7.12	7.93	7.82	9.21	10.66	11.71
		15	0	0.01	0.05	1.77	0.71	7.79	16.34	22.03	−12.37	−0.14	5.08	8.83	8.12	15.74	24.30	53.05
		15	1	0.01	0.05	1.38	0.67	4.75	14.64	18.85	−6.54	2.33	5.64	8.60	8.11	13.47	20.16	49.40
$\mathcal{A}^1_{\text{avg}, 25\,\text{ns}}$	2ε	0	0	0.00	0.02	0.46	0.25	1.83	2.23	2.84	5.16	6.57	7.10	7.93	7.82	9.31	10.74	11.56
		15	0	0.01	0.05	1.19	0.68	3.86	8.65	16.53	−6.22	3.23	6.49	8.37	7.97	11.53	16.65	45.30
		15	1	0.01	0.05	0.96	0.64	2.50	7.72	10.31	−3.31	5.27	6.65	8.25	7.96	10.73	15.09	42.88
$\mathcal{A}^1_{\text{mid}, 25\,\text{ns}}$	d^+	0	0	0.01	0.03	1.37	0.41	6.51	8.33	9.61	−1.01	3.45	5.64	8.41	7.86	13.10	19.73	20.97
		15	0	0.01	0.07	2.52	1.10	8.63	16.36	24.93	−18.03	−0.30	3.72	9.20	8.24	17.85	24.37	52.93
		15	1	0.01	0.06	2.16	1.02	7.99	14.12	20.91	−10.03	1.32	4.16	8.99	8.22	16.14	20.52	49.41
$\mathcal{A}^1_{\text{avg}, 25\,\text{ns}}$	d^+	0	0	0.01	0.03	1.49	0.38	7.38	8.30	9.02	−0.95	3.54	5.44	8.50	7.86	14.11	19.85	20.78
		15	0	0.01	0.06	2.13	1.00	8.07	9.39	24.17	−15.01	2.42	4.50	8.87	8.05	16.09	20.07	51.05
		15	1	0.01	0.06	1.92	0.94	7.77	8.45	13.74	−7.93	3.36	4.99	8.77	8.04	15.49	19.91	40.51

A.8 Plots of the Pulse-based Algorithms in $\mathcal{G}_D$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skew

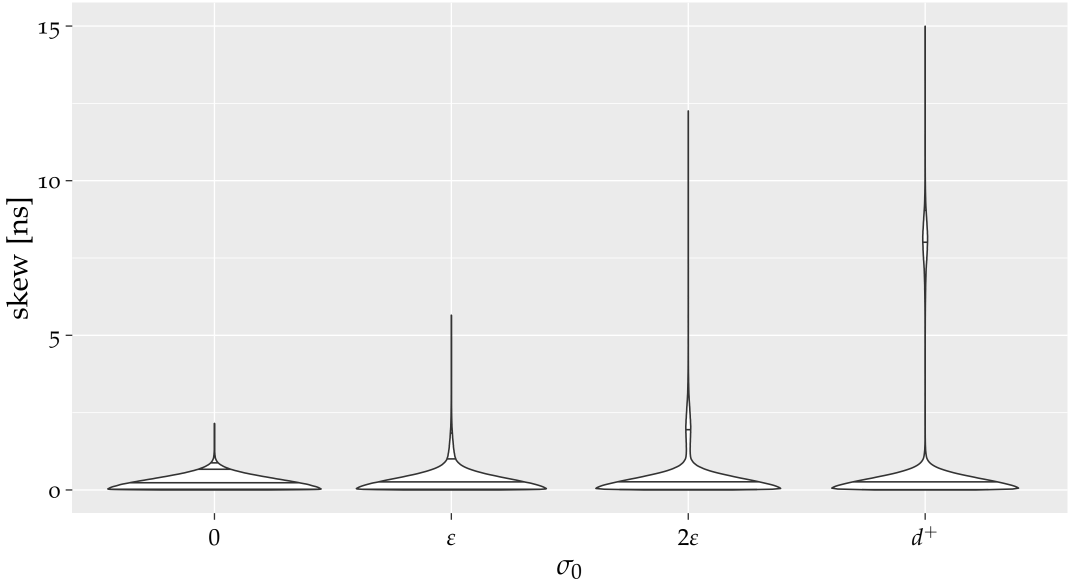


(b) inter-layer skew

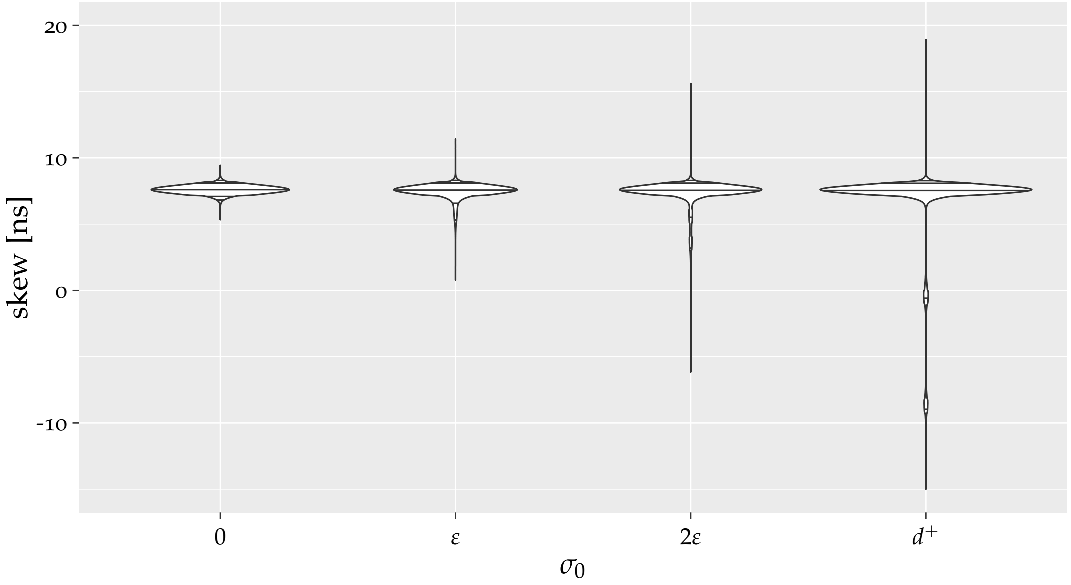
Figure A.29: Skew in $\mathcal{G}_D$ with $\mathcal{A}^2_{\text{adj}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

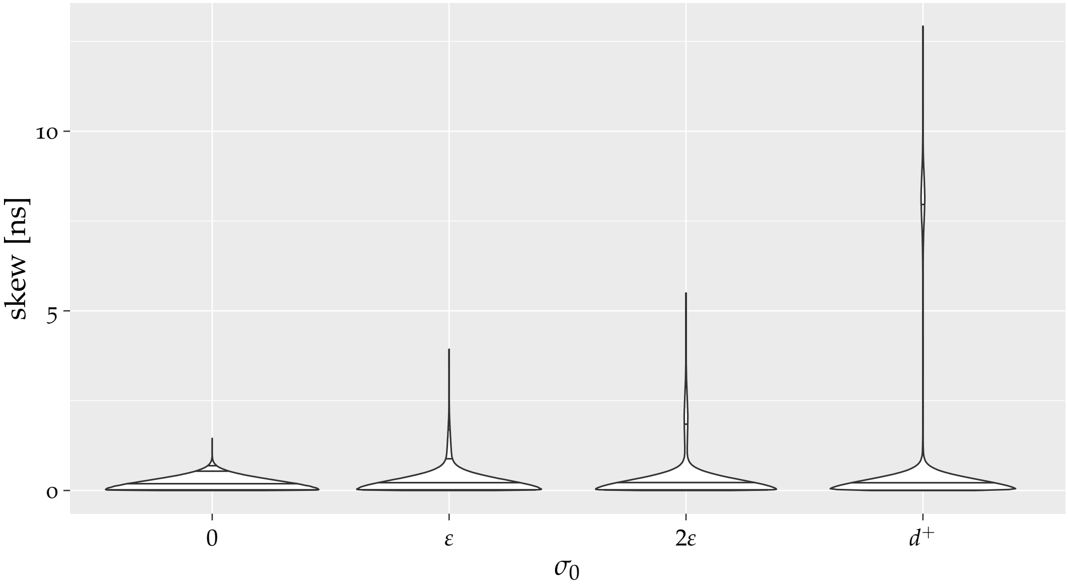


(a) intra-layer skew

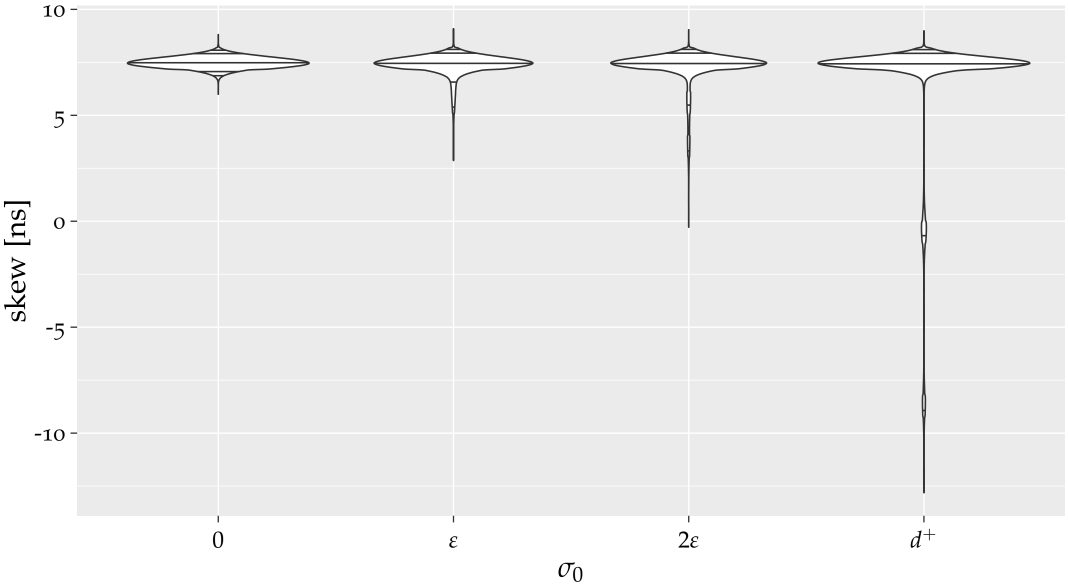


(b) inter-layer skew

Figure A.30: Skew in $\mathcal{G}_D$ with $\mathcal{A}^2_{\text{adj}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-15, 20]$ ns.



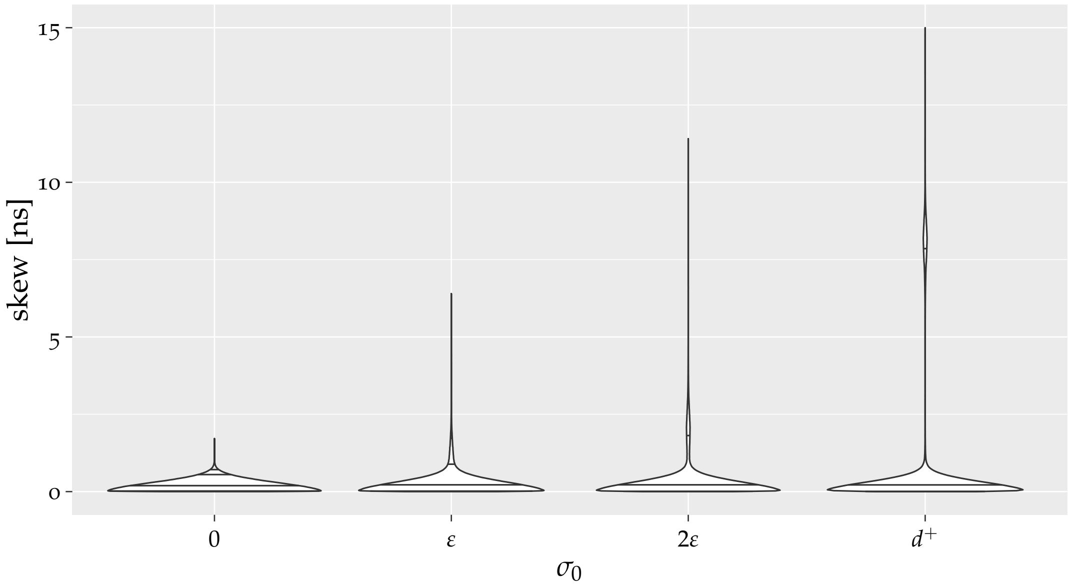
(a) intra-layer skew



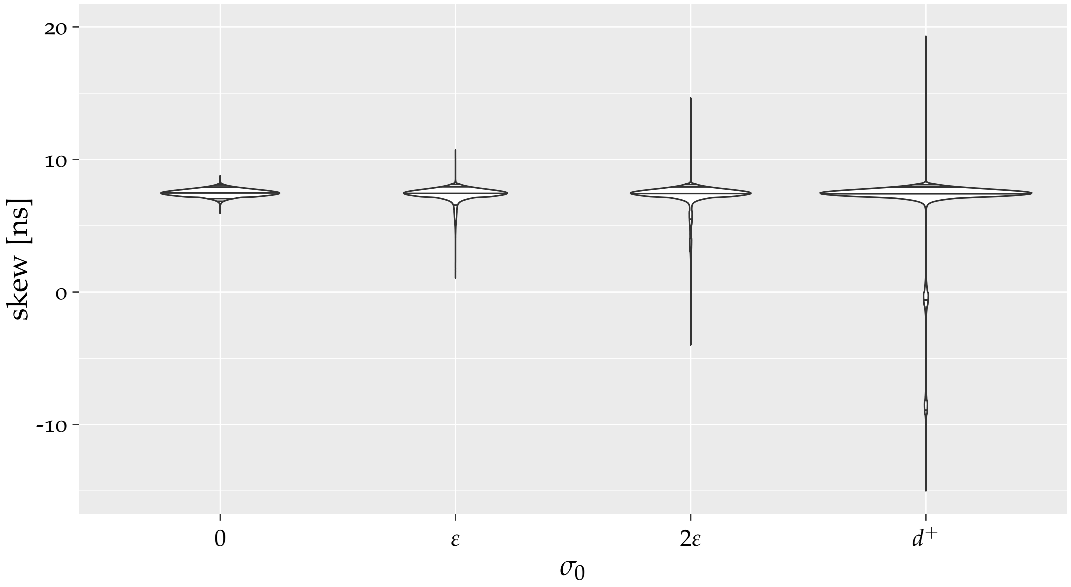
(b) inter-layer skew

Figure A.31: Skew in $\mathcal{G}_D$ with $\mathcal{A}^2_{\text{any}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



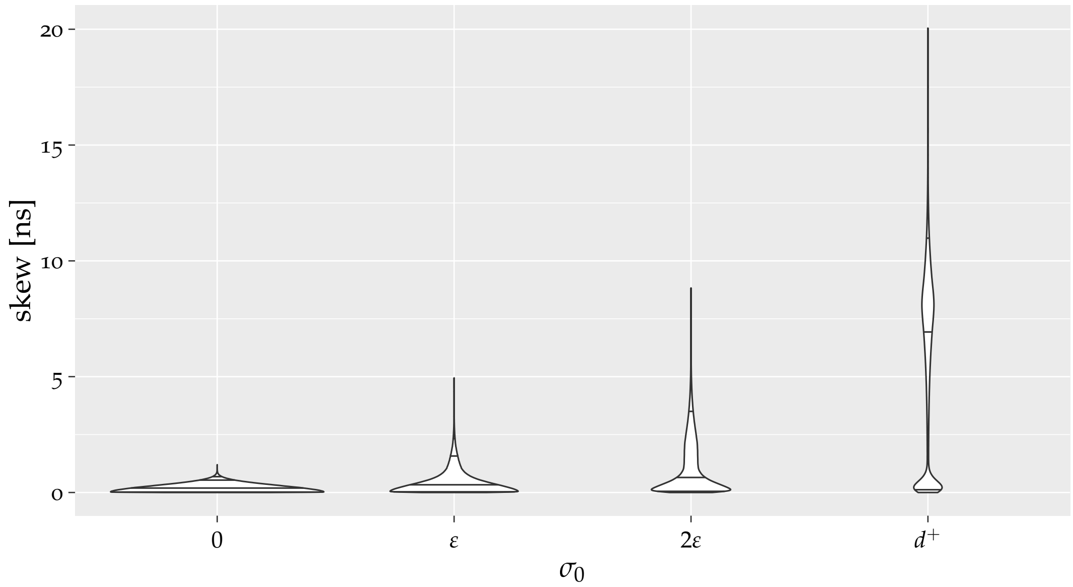
(a) intra-layer skew



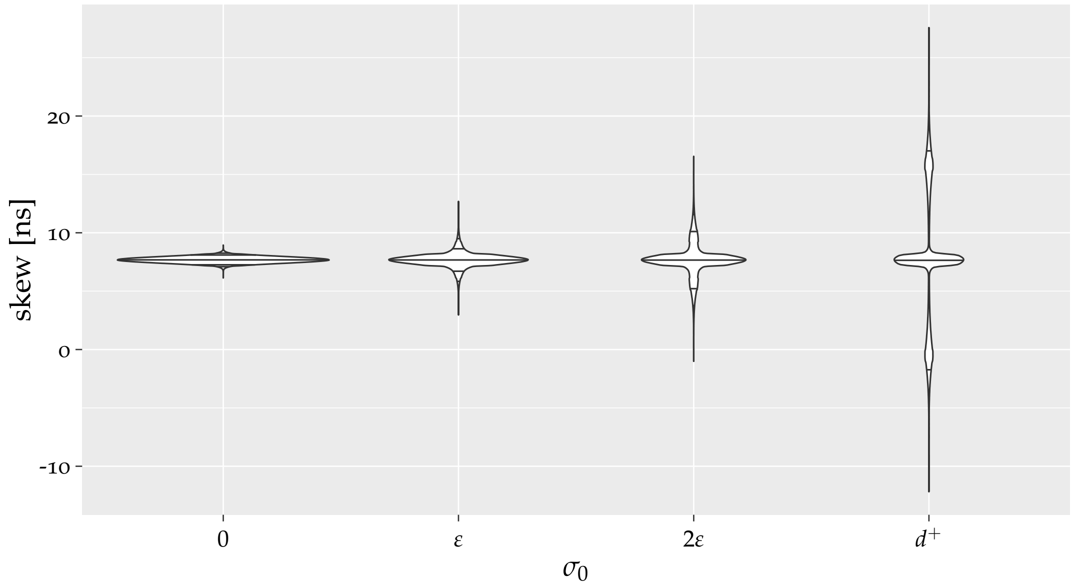
(b) inter-layer skew

Figure A.32: Skew in $\mathcal{G}_D$ with $\mathcal{A}^2_{\text{any}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-15, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



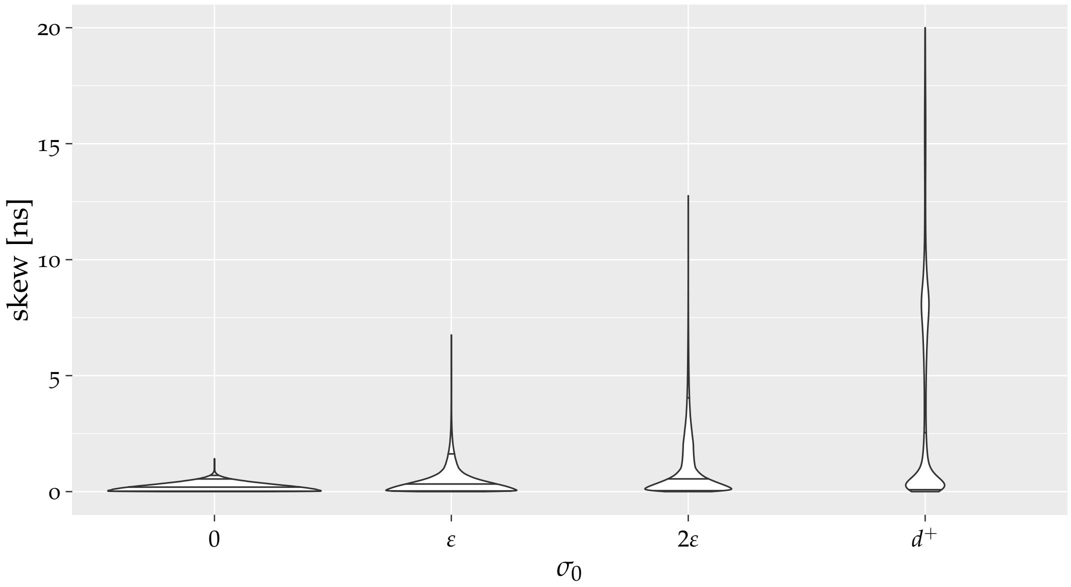
(a) intra-layer skew



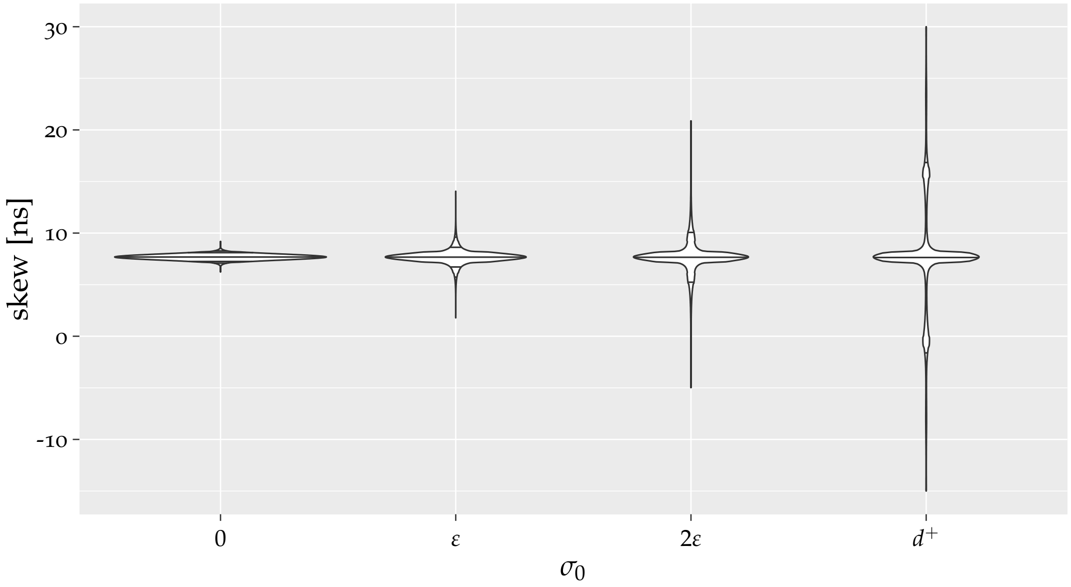
(b) inter-layer skew

Figure A.33: Skew in $\mathcal{G}_D$ with $\mathcal{A}^3_{\text{any}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skew



(b) inter-layer skew

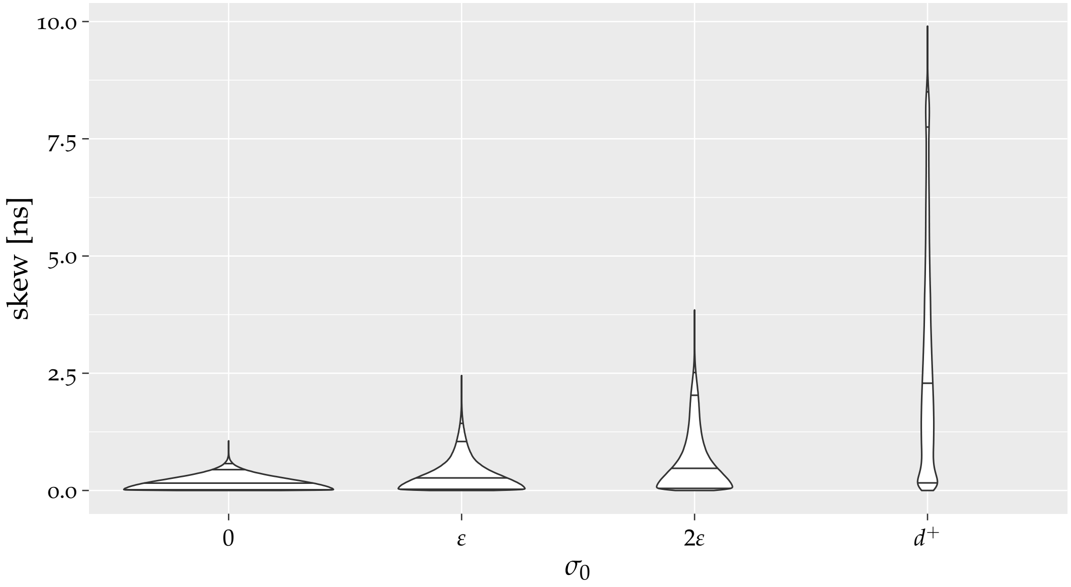
Figure A.34: Skew in $\mathcal{G}_D$ with $\mathcal{A}^3_{\text{any}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-15, 30]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

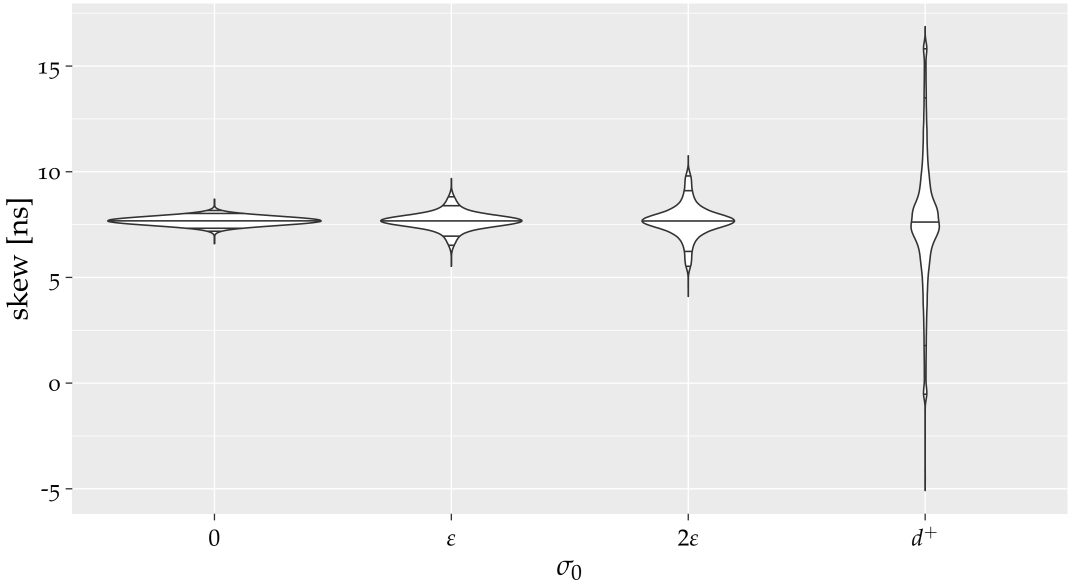
A.9 Plots of the Time-based Algorithms with $d = 0$ and $d = 1$ in $\mathcal{G}_D$

In this section we present the plots of the algorithms $\mathcal{A}^0_{\mathrm{avg},\,25\,\mathrm{ns}}$, $\mathcal{A}^0_{\mathrm{mid},\,25\,\mathrm{ns}}$, $\mathcal{A}^1_{\mathrm{avg},\,25\,\mathrm{ns}}$, and $\mathcal{A}^1_{\mathrm{mid},\,25\,\mathrm{ns}}$. We omit the common offset of 25 ns in the algorithm description. Further, recall that the provided inter-layer skews have been compensated by 25 ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



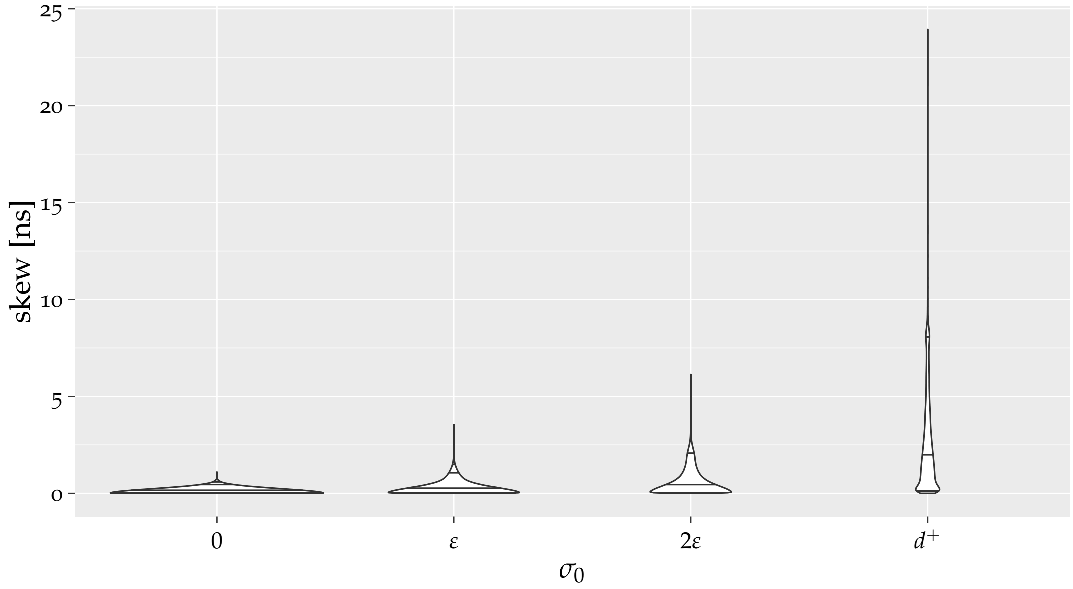
(a) intra-layer skew



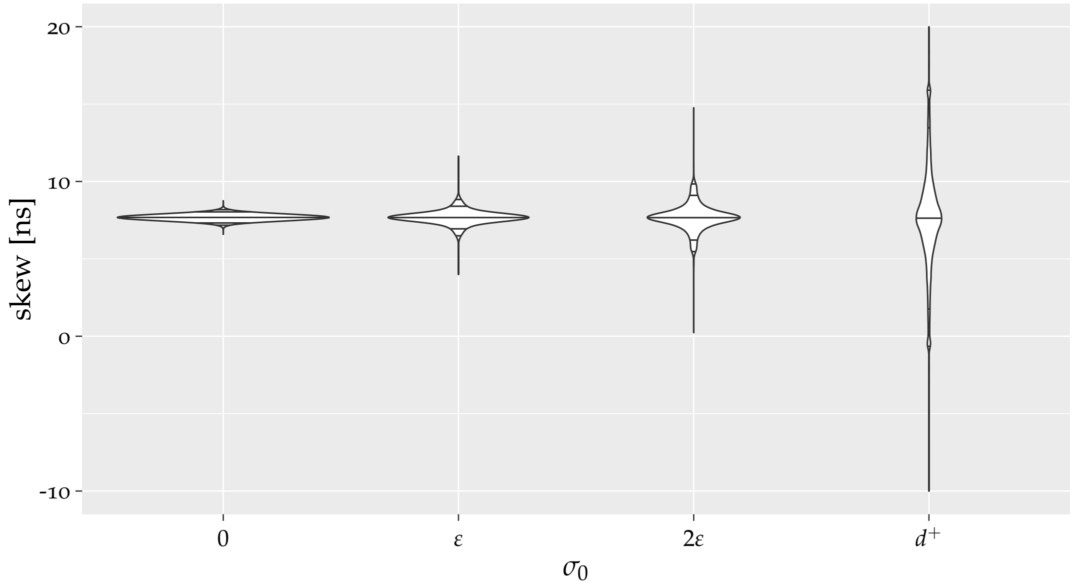
(b) inter-layer skew

Figure A.35: Skew in $\mathcal{G}_D$ with $\mathcal{A}^0_{\mathrm{mid},\,25\,\mathrm{ns}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



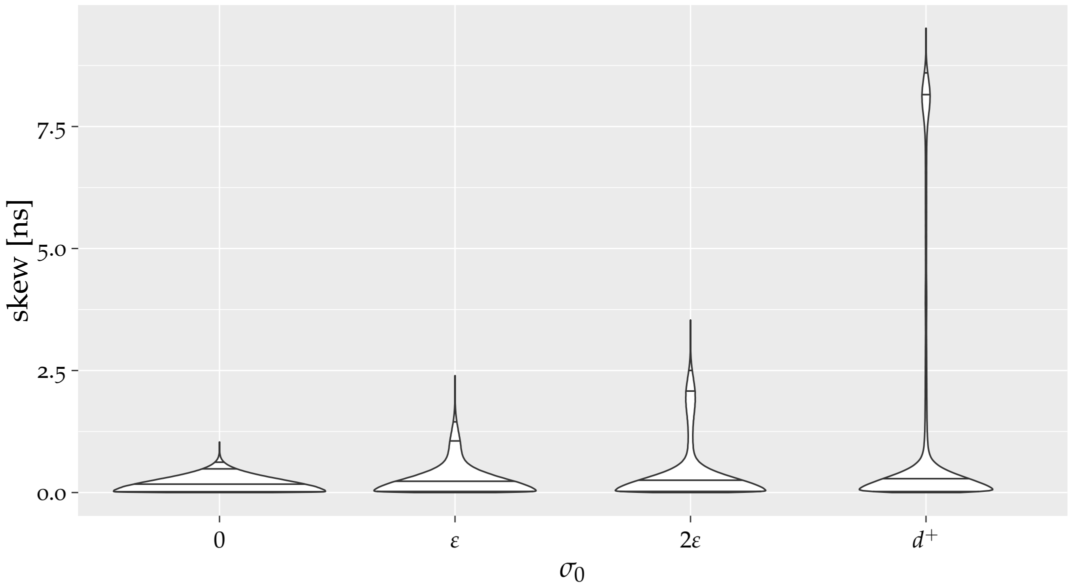
(a) intra-layer skew



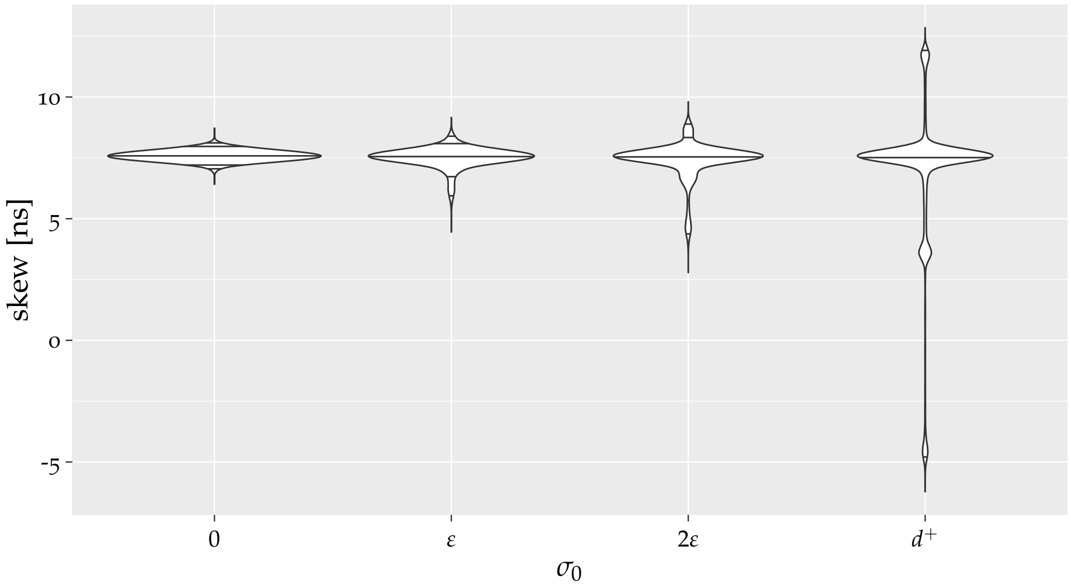
(b) inter-layer skew

Figure A.36: Skew in $\mathcal{G}_D$ with $\mathcal{A}^0_{\text{mid},25\,\text{ns}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



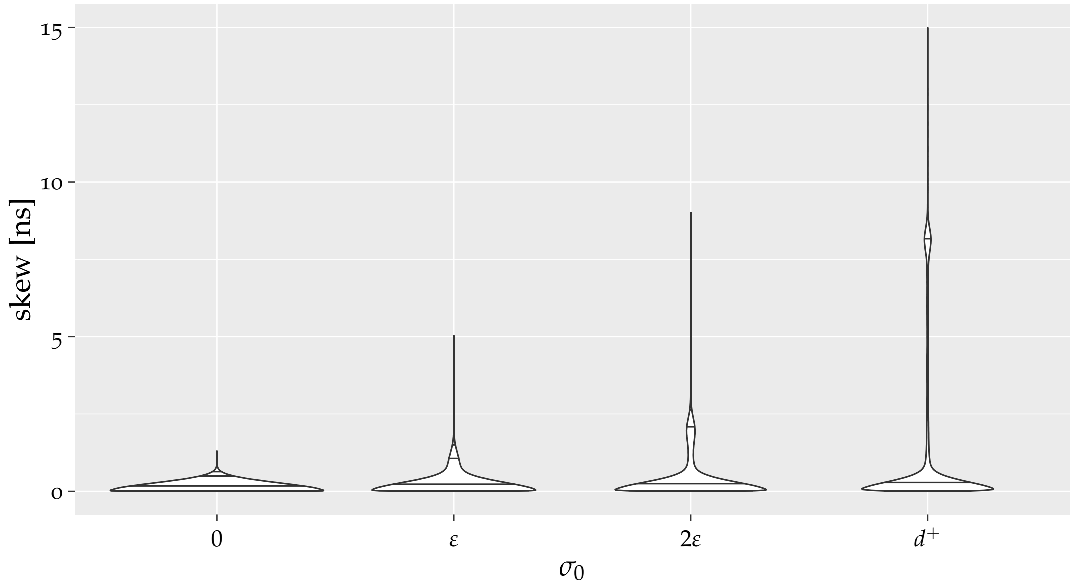
(a) intra-layer skew



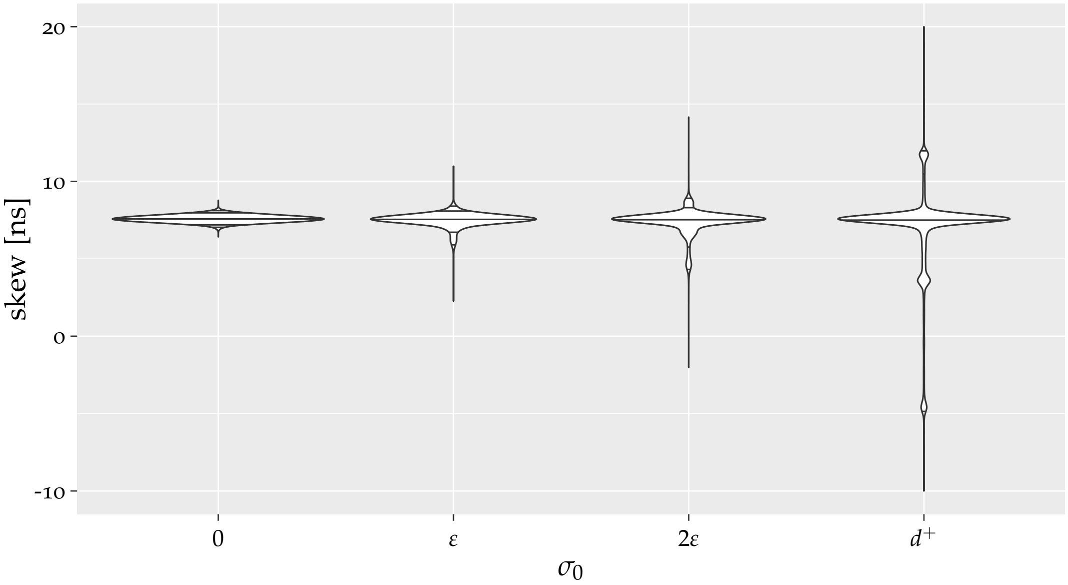
(b) inter-layer skew

Figure A.37: Skew in $\mathcal{G}_D$ with $\mathcal{A}^1_{\mathrm{mid},\,25\,\mathrm{ns}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



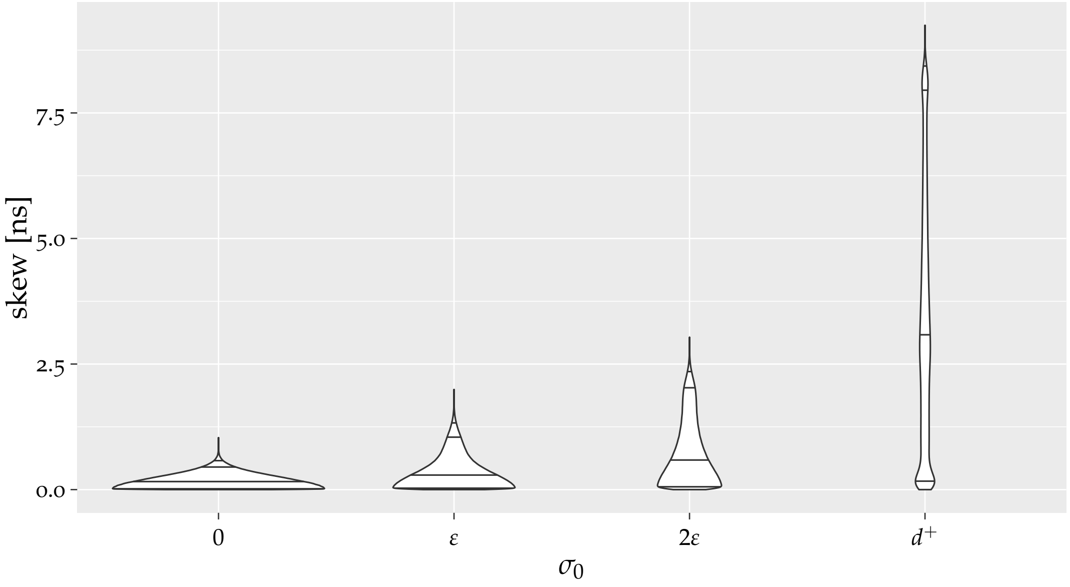
(a) intra-layer skew



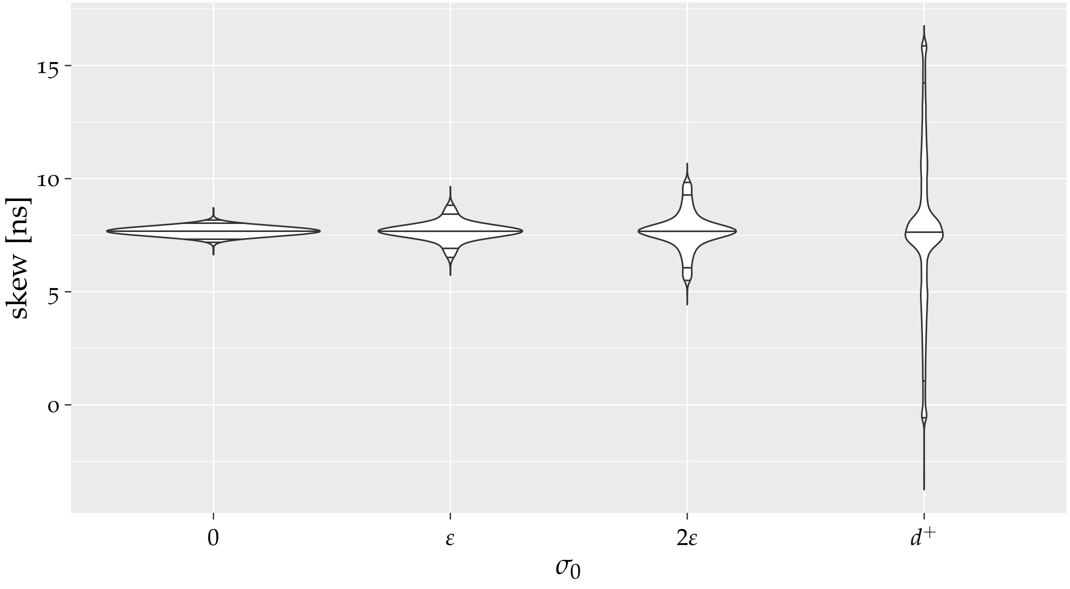
(b) inter-layer skew

Figure A.38: Skew in $\mathcal{G}_D$ with $\mathcal{A}^1_{\text{mid},25\,\text{ns}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



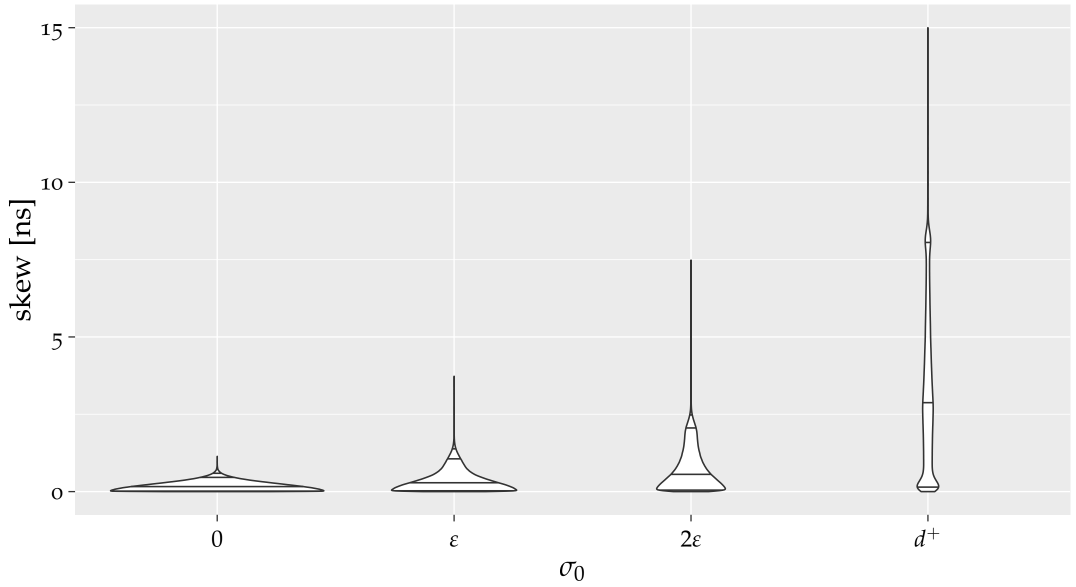
(a) intra-layer skew



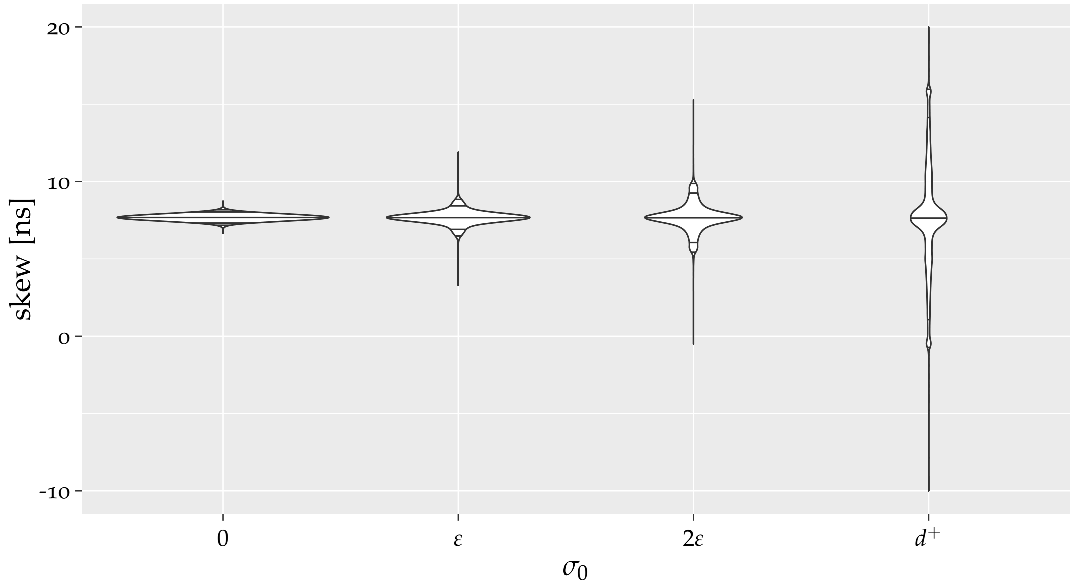
(b) inter-layer skew

Figure A.39: Skew in $\mathcal{G}_D$ with $\mathcal{A}^0_{\text{avg},\,25\,\text{ns}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



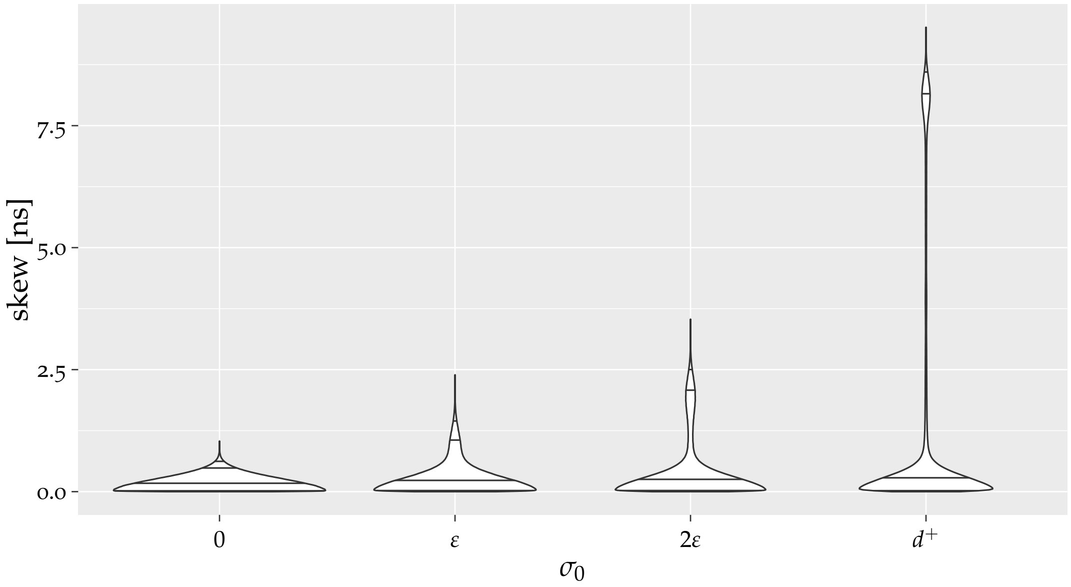
(a) intra-layer skew



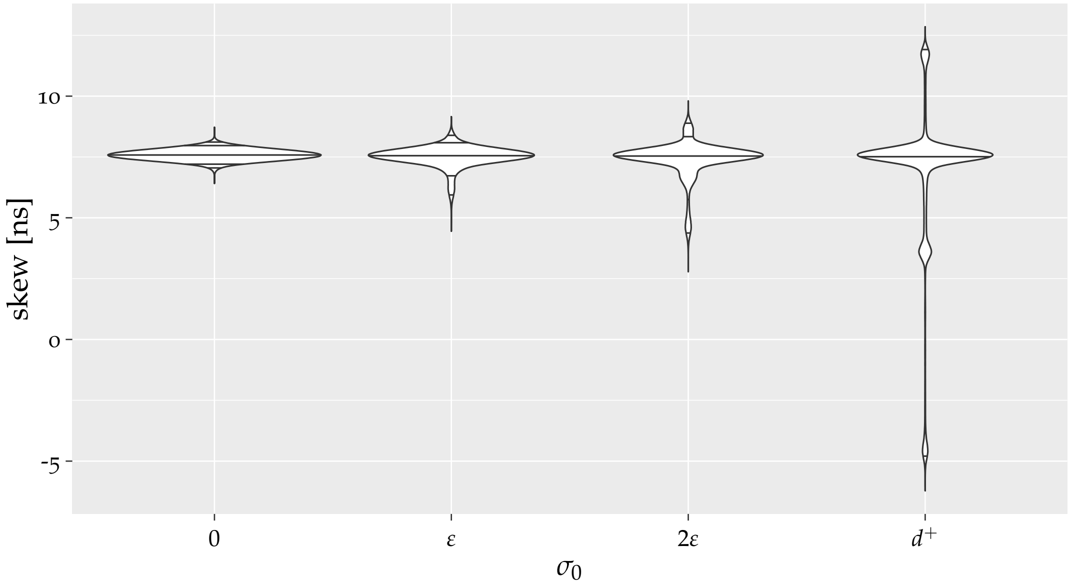
(b) inter-layer skew

Figure A.40: Skew in $\mathcal{G}_D$ with $\mathcal{A}^0_{\text{avg},25\,\text{ns}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



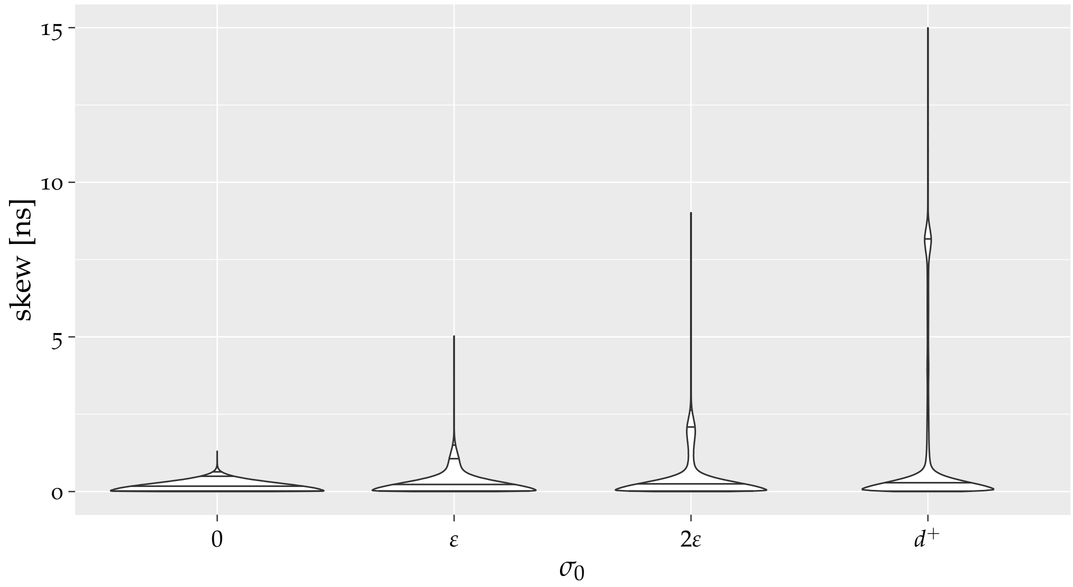
(a) intra-layer skew



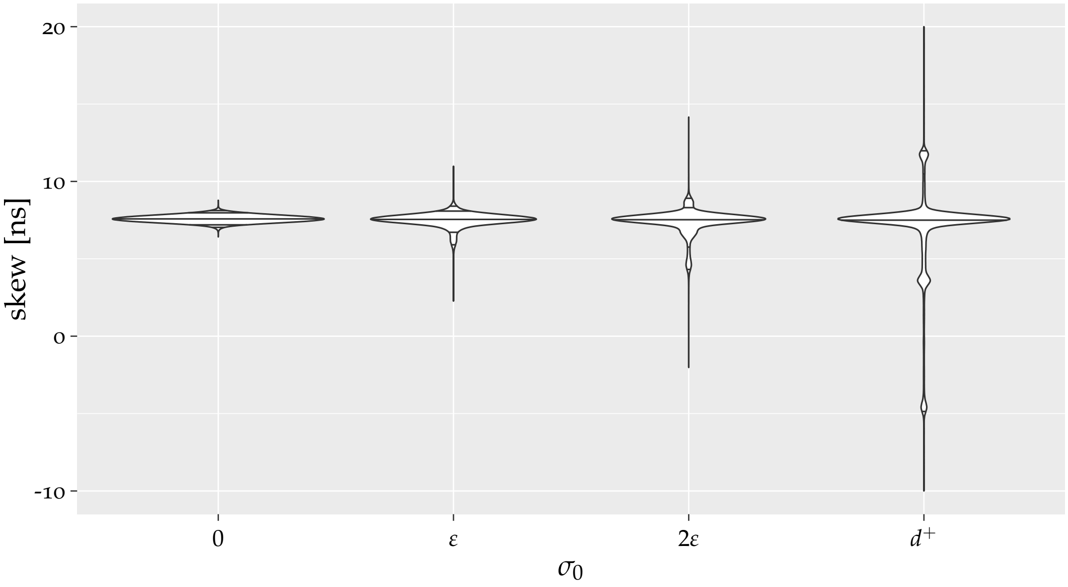
(b) inter-layer skew

Figure A.41: Skew in $\mathcal{G}_D$ with $\mathcal{A}^1_{\text{avg}, 25\,\text{ns}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skew



(b) inter-layer skew

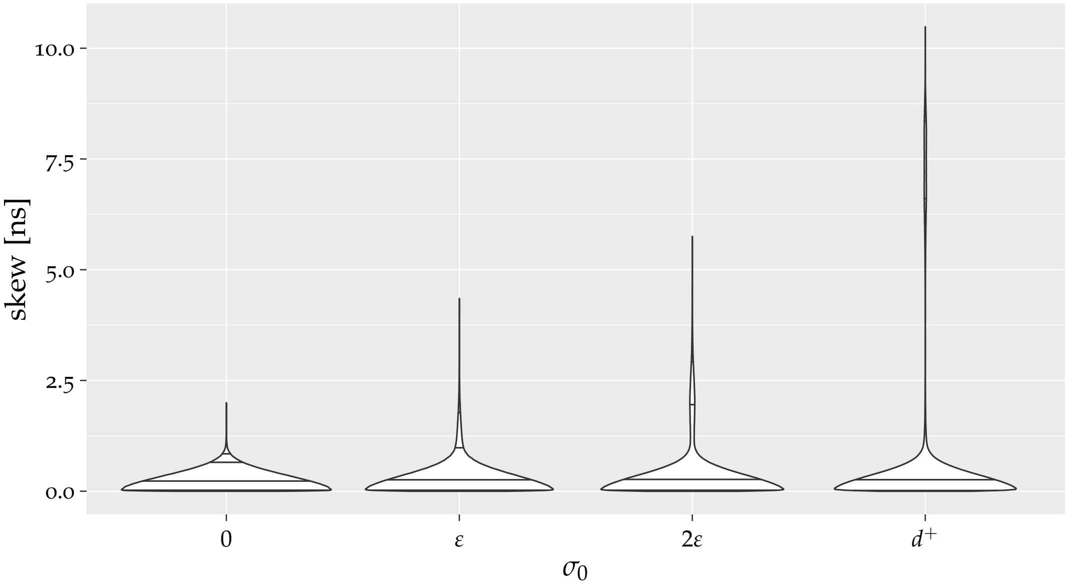
Figure A.42: Skew in $\mathcal{G}_D$ with $\mathcal{A}^1_{\mathrm{avg},\,25\,\mathrm{ns}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

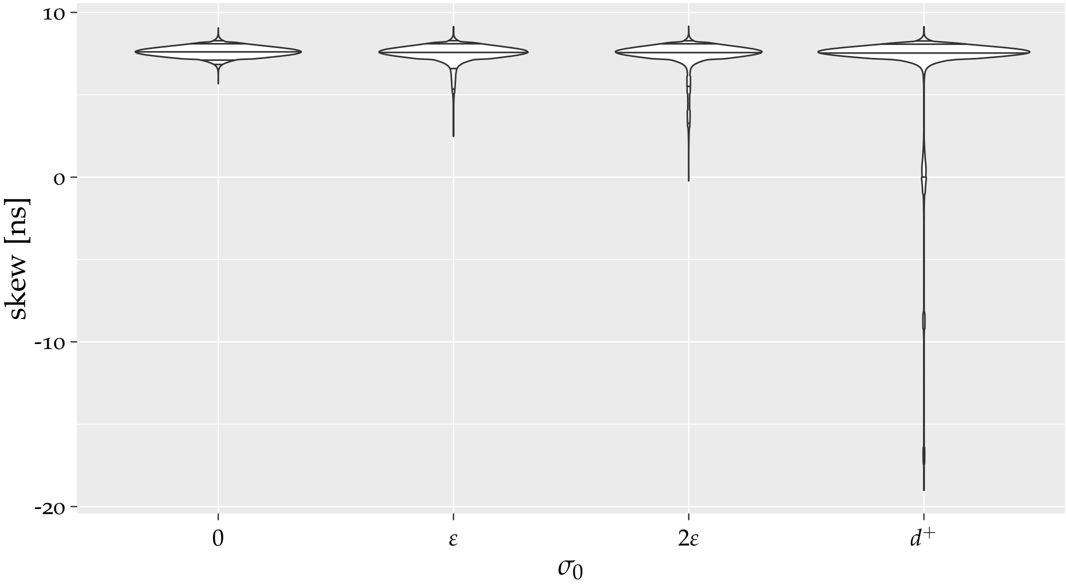
A.10 Plots of the Pulse-based Algorithms in $\mathcal{G}_E$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

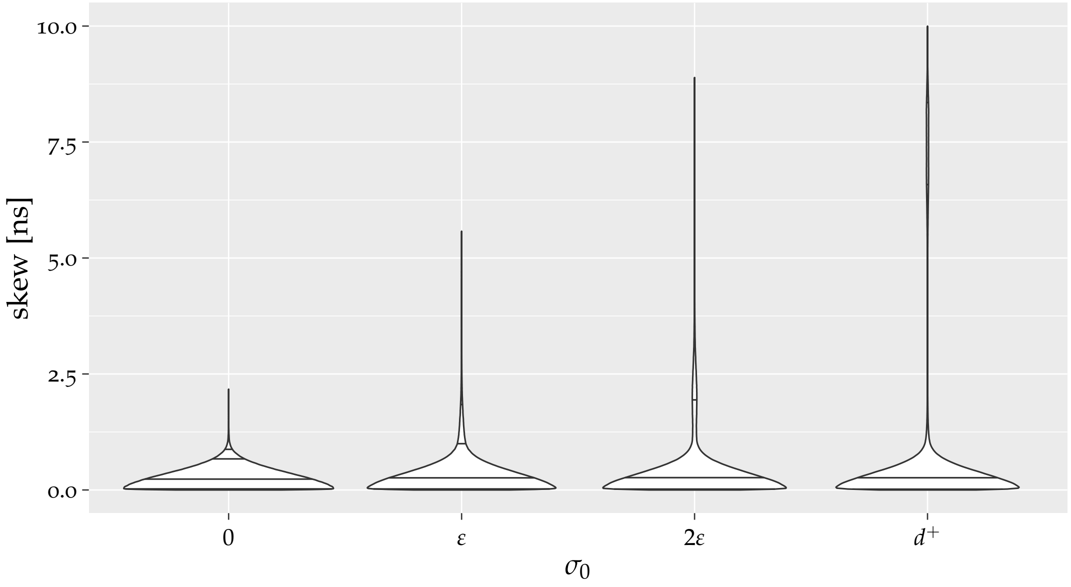


(a) intra-layer skew

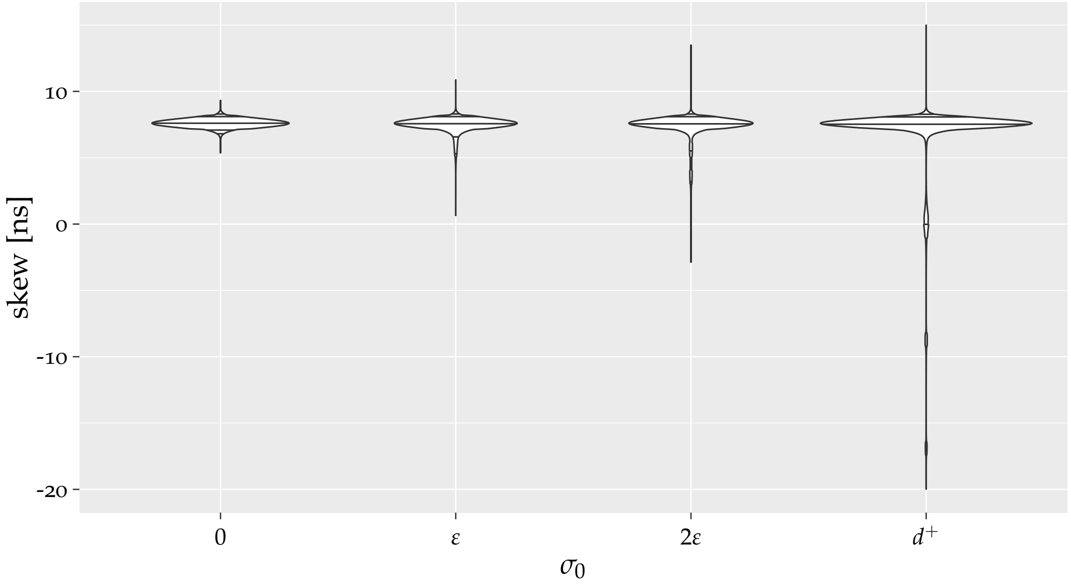


(b) inter-layer skew

Figure A.43: Skew in $\mathcal{G}_E$ with $\mathcal{A}^2_{\text{adj}}$ without faults.



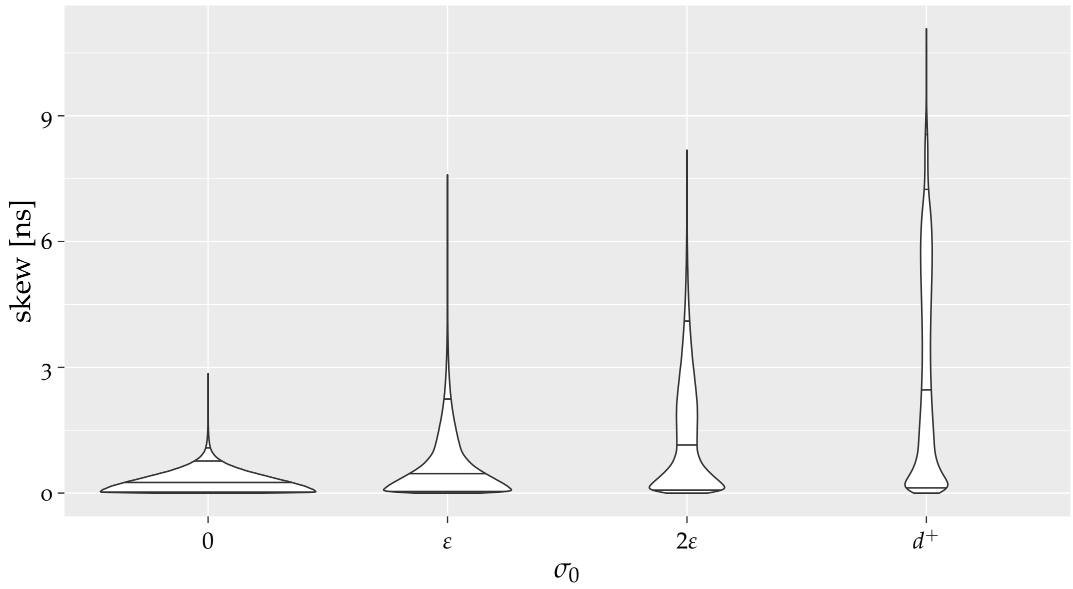
(a) intra-layer skew



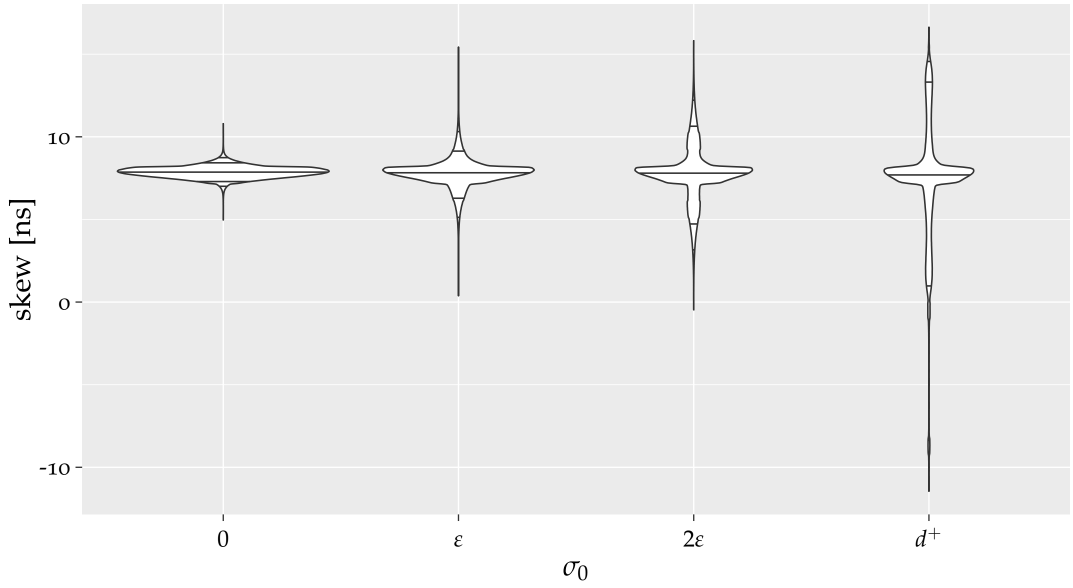
(b) inter-layer skew

Figure A.44: Skew in $\mathcal{G}_E$ with $\mathcal{A}^2_{\text{adj}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 10 ns, while the inter-layer skews have been limited to be within $[-20, 15]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



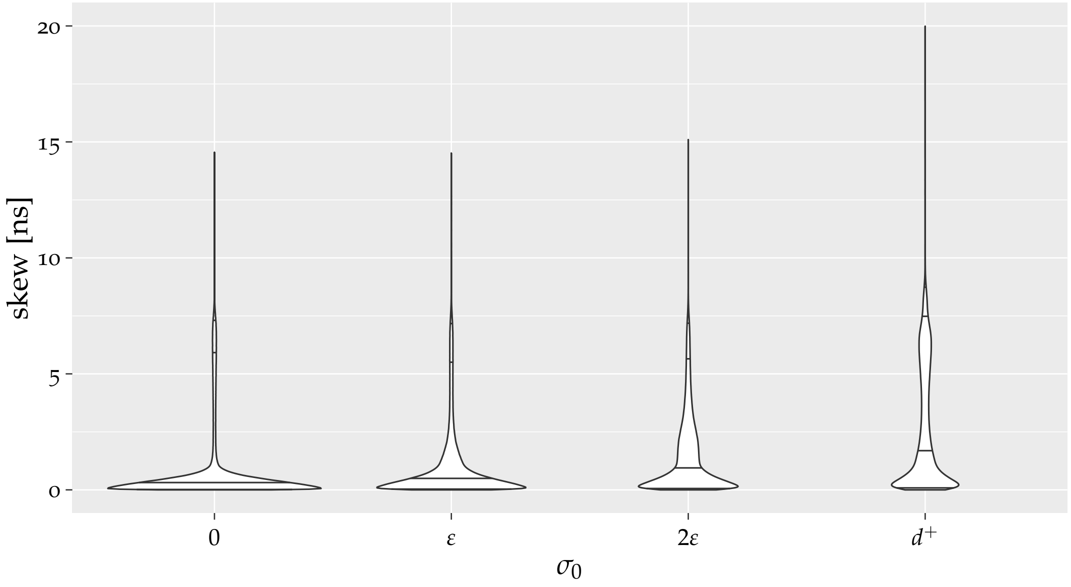
(a) intra-layer skew



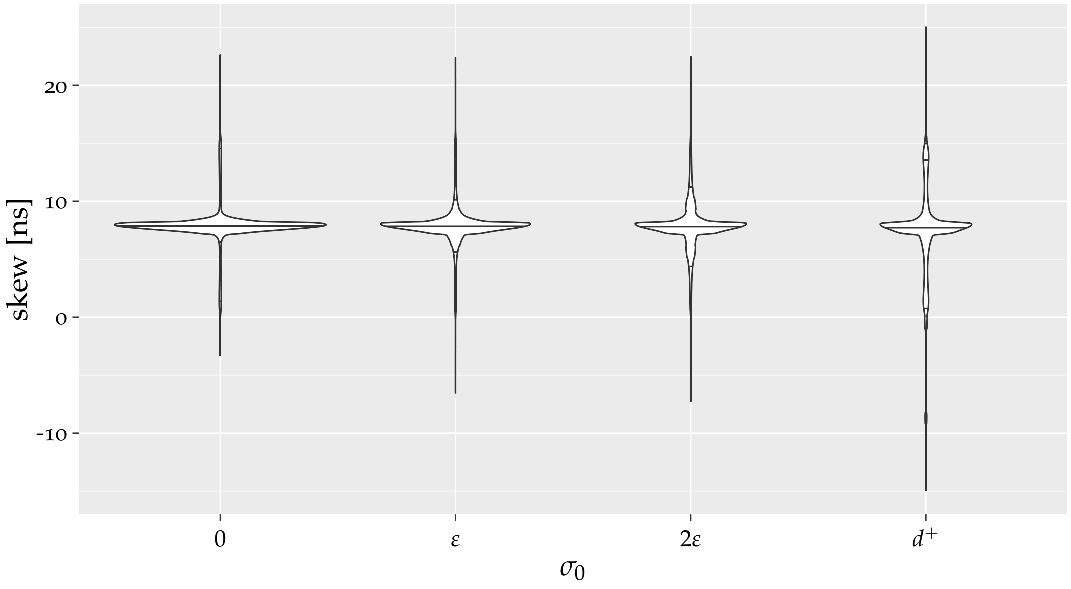
(b) inter-layer skew

Figure A.45: Skew in $\mathcal{G}_E$ with $\mathcal{A}^3_{\text{adj}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



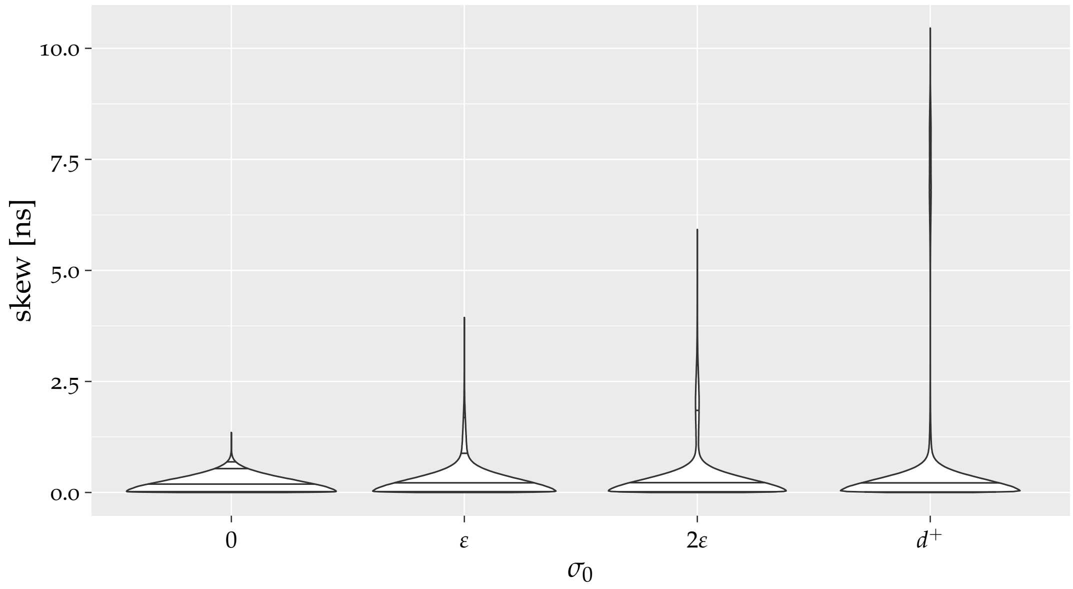
(a) intra-layer skew



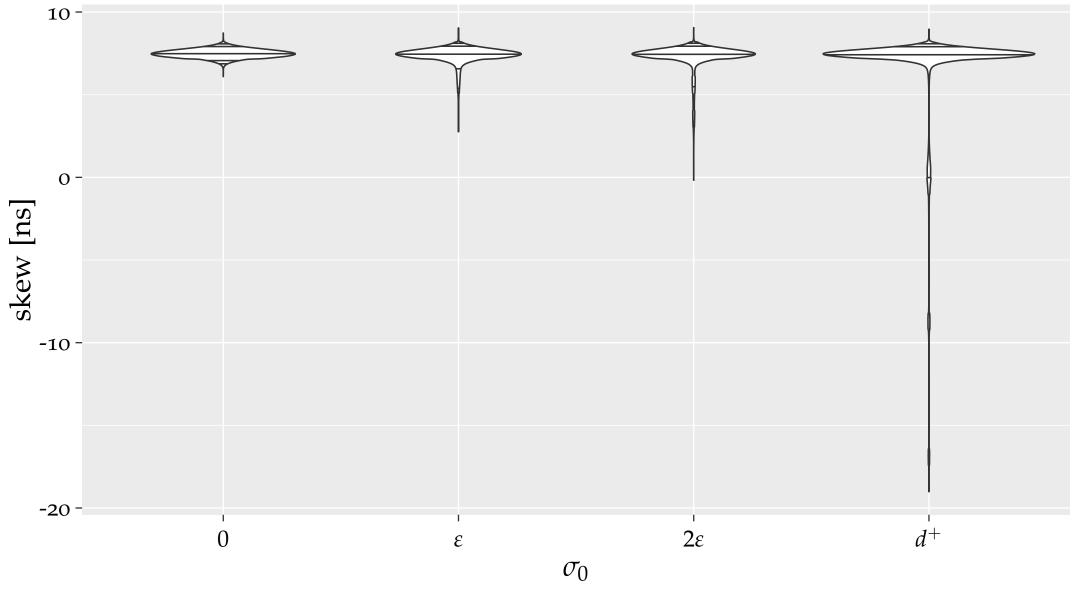
(b) inter-layer skew

Figure A.46: Skew in $\mathcal{G}_E$ with $\mathcal{A}^3_{\text{adj}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-15, 25]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



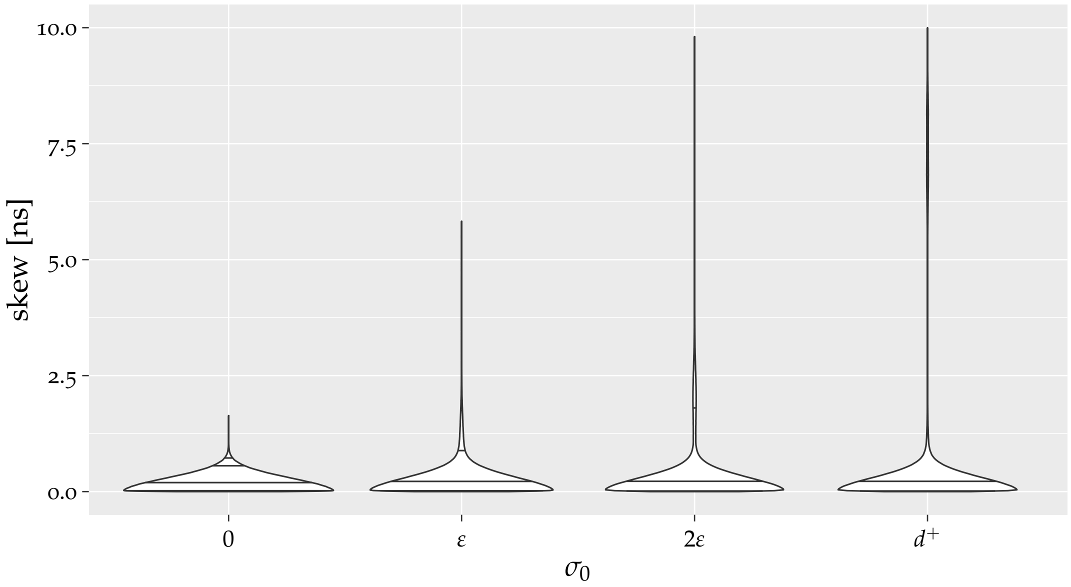
(a) intra-layer skew



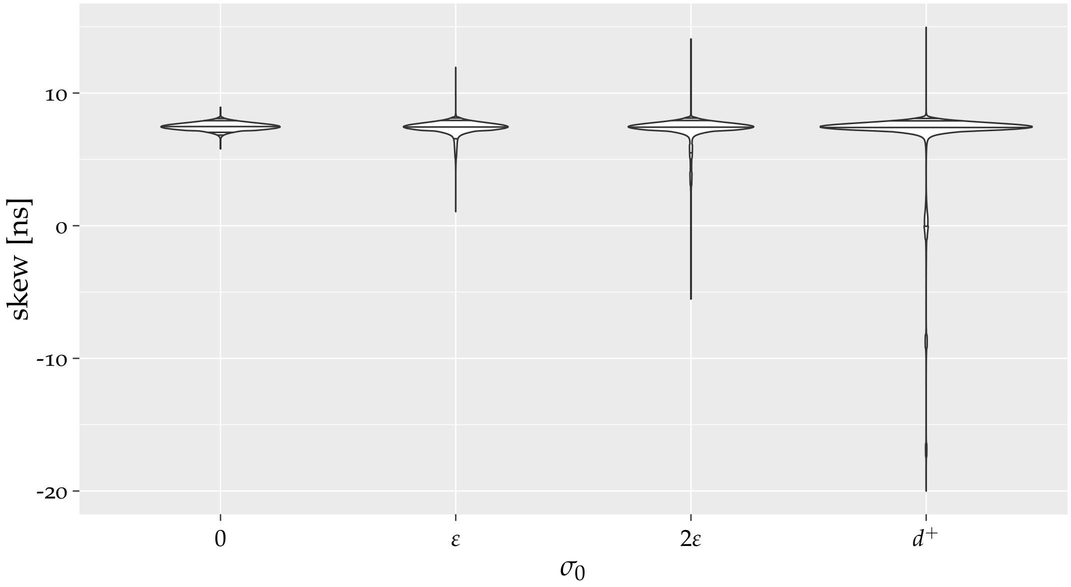
(b) inter-layer skew

Figure A.47: Skew in $\mathcal{G}_E$ with $\mathcal{A}^2_{\text{any}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



(a) intra-layer skew

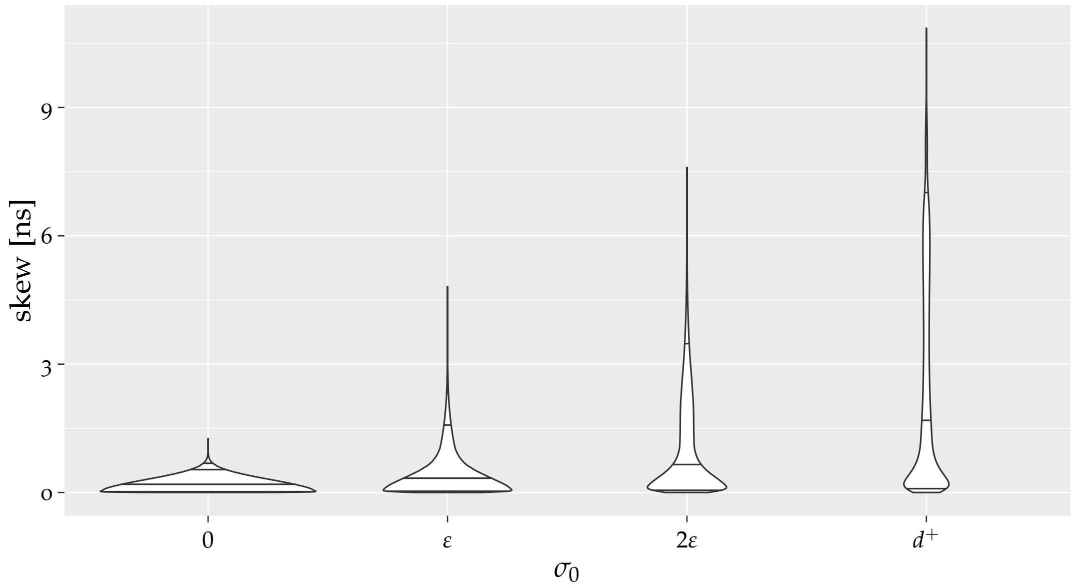


(b) inter-layer skew

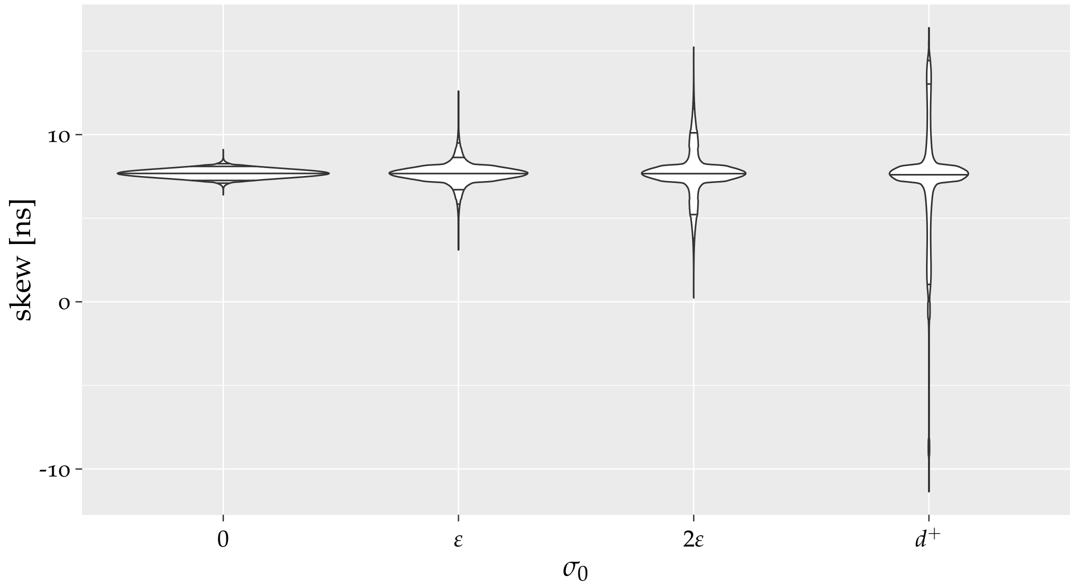
Figure A.48: Skew in $\mathcal{G}_E$ with $\mathcal{A}^2_{\text{any}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 10 ns, while the inter-layer skews have been limited to be within $[-20, 15]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

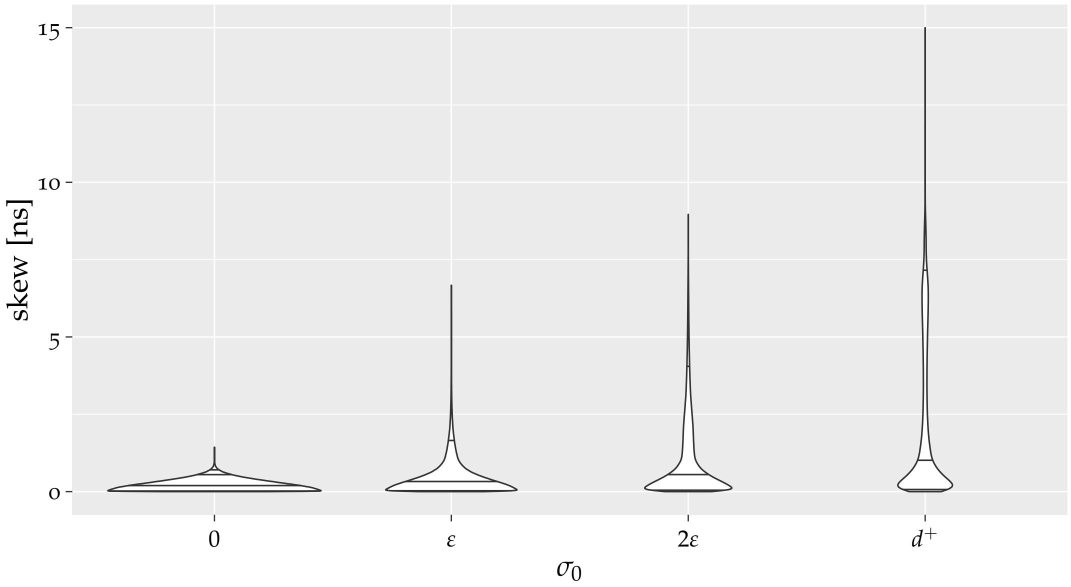


(a) intra-layer skew

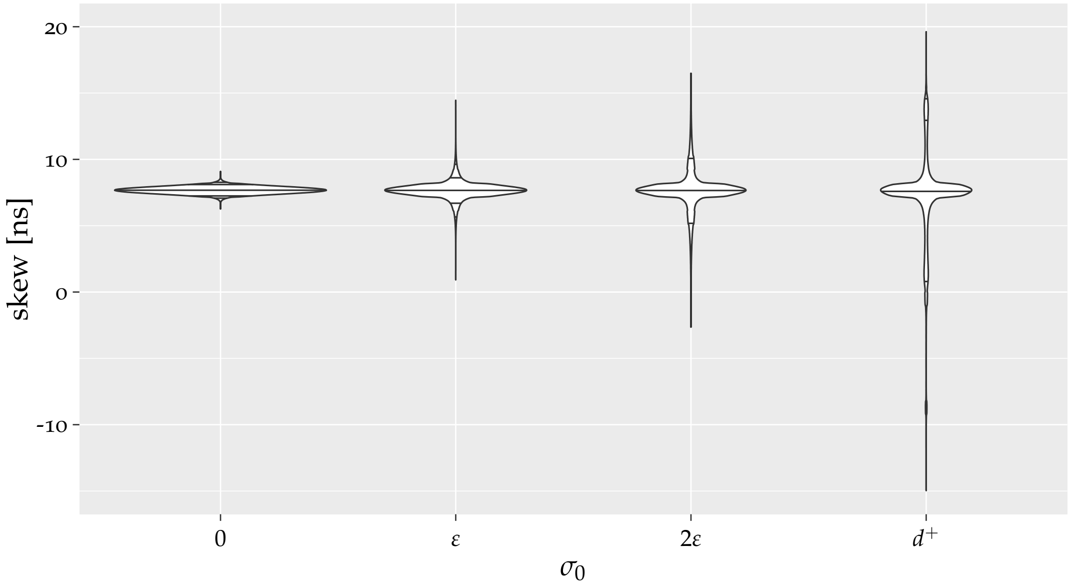


(b) inter-layer skew

Figure A.49: Skew in $\mathcal{G}_E$ with $\mathcal{A}^3_{\text{any}}$ without faults.



(a) intra-layer skew



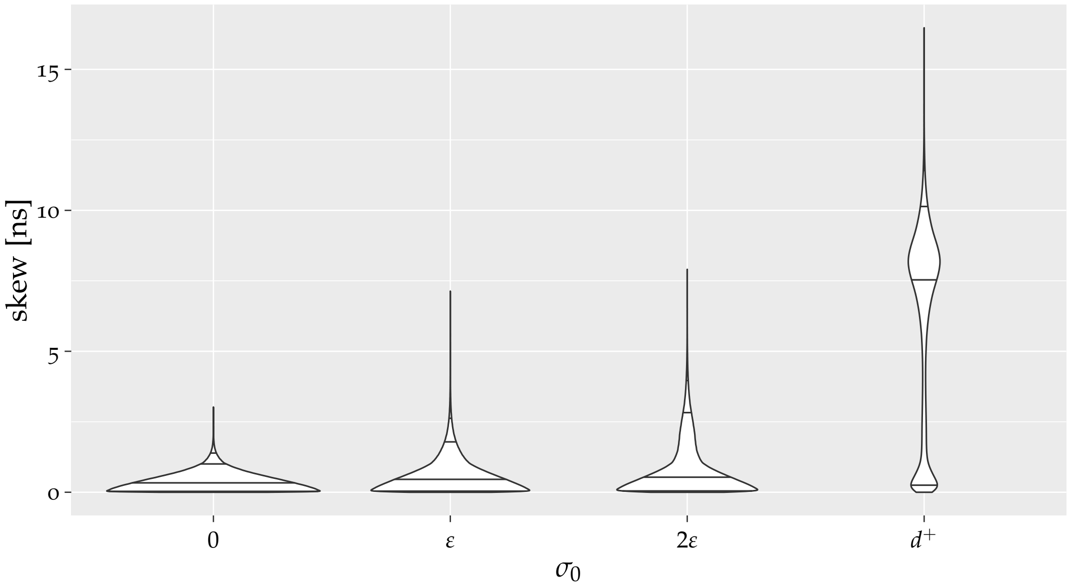
(b) inter-layer skew

Figure A.50: Skew in $\mathcal{G}_E$ with $\mathcal{A}^3_{\text{any}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-15, 20]$ ns.

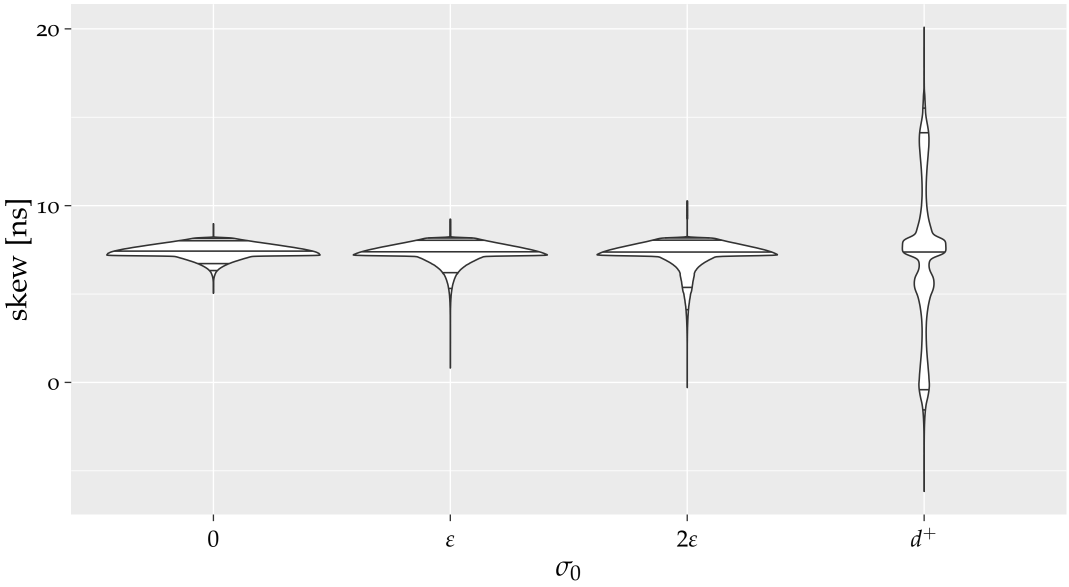
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

A.11 Plots of the Pulse-based Algorithms in $\mathcal{G}_F$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



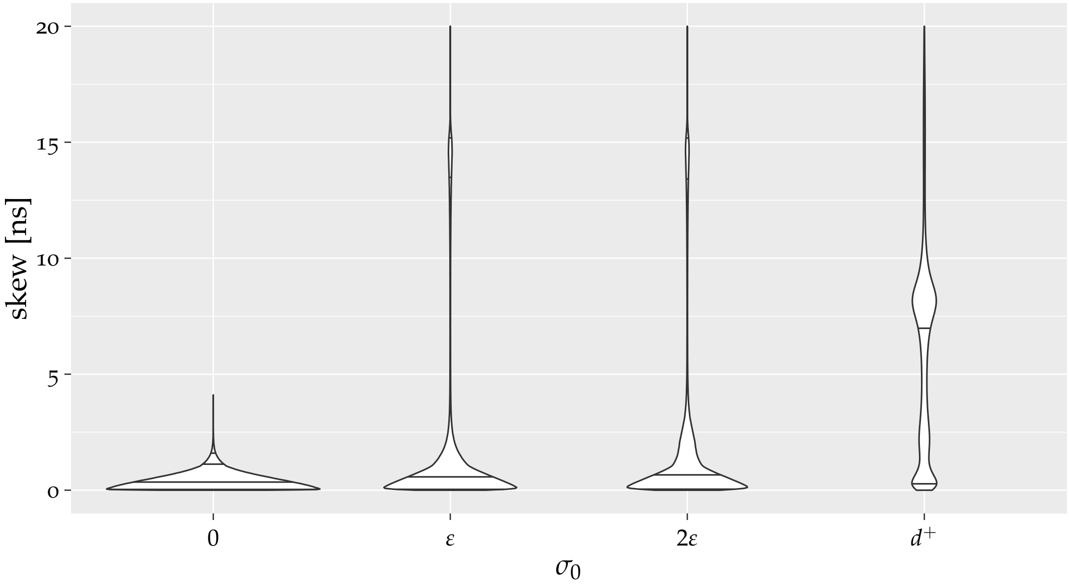
(a) intra-layer skews



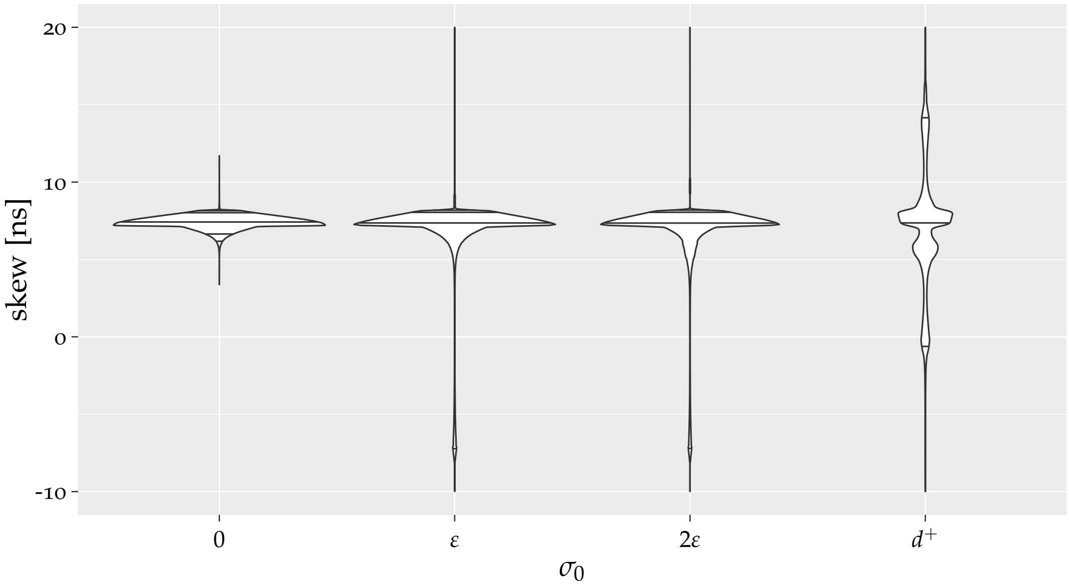
(b) inter-layer skews

Figure A.51: Skew in $\mathcal{G}_F^{13,14}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



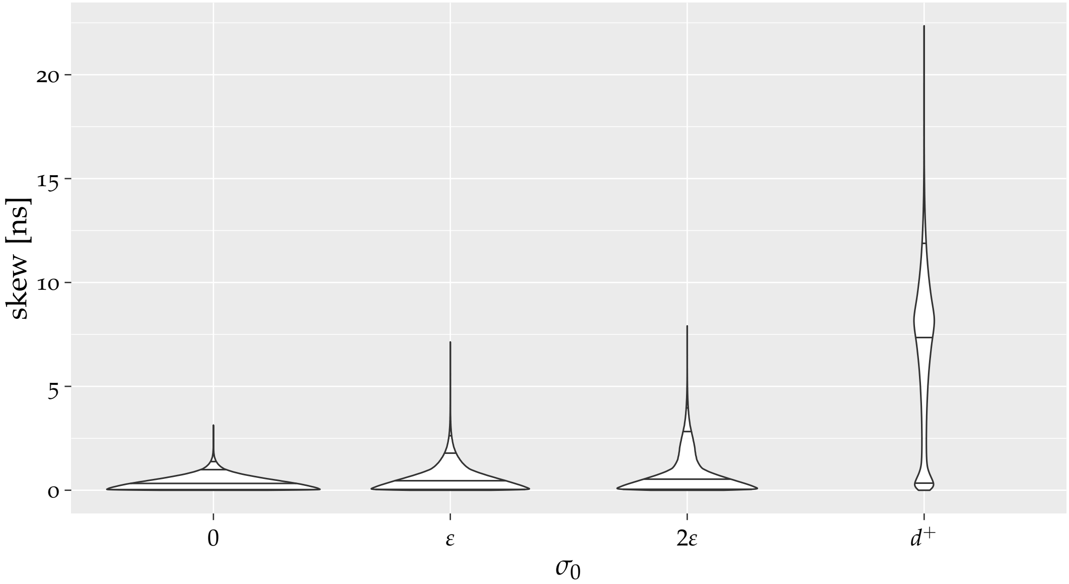
(a) intra-layer skews



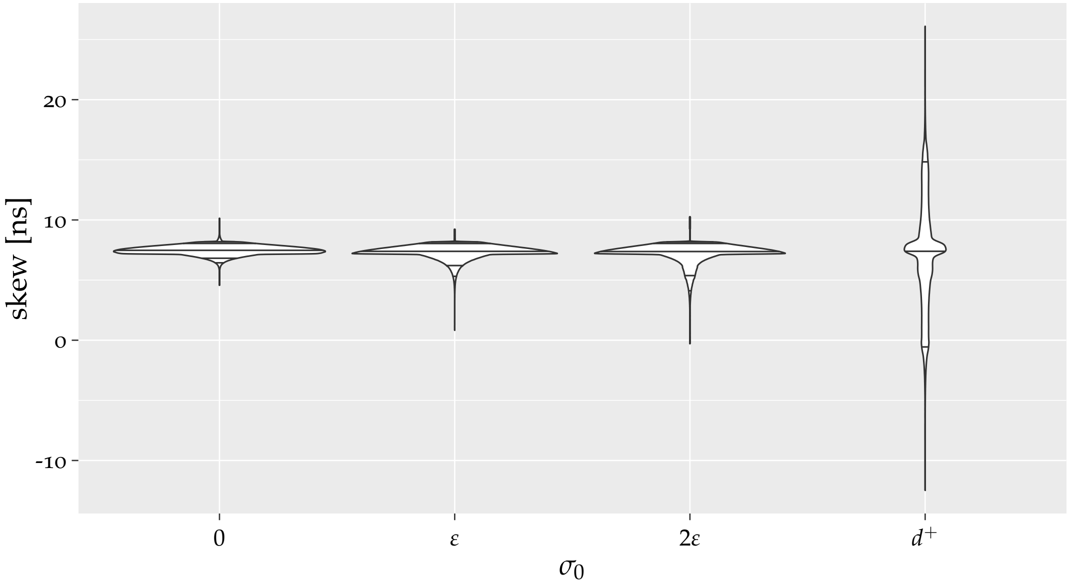
(b) inter-layer skews

Figure A.52: Skew in $\mathcal{G}_F^{13,14}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



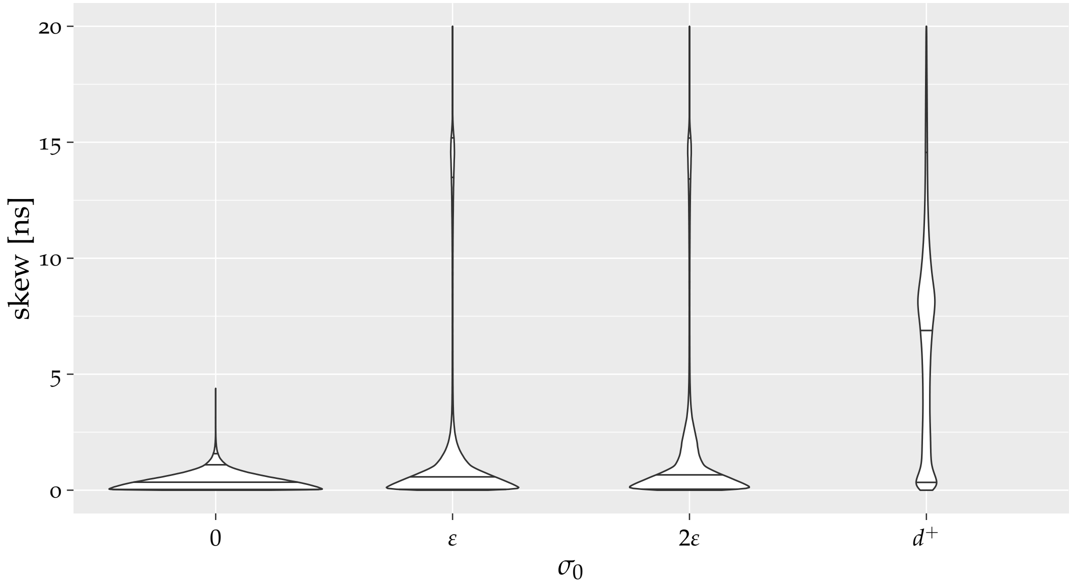
(a) intra-layer skews



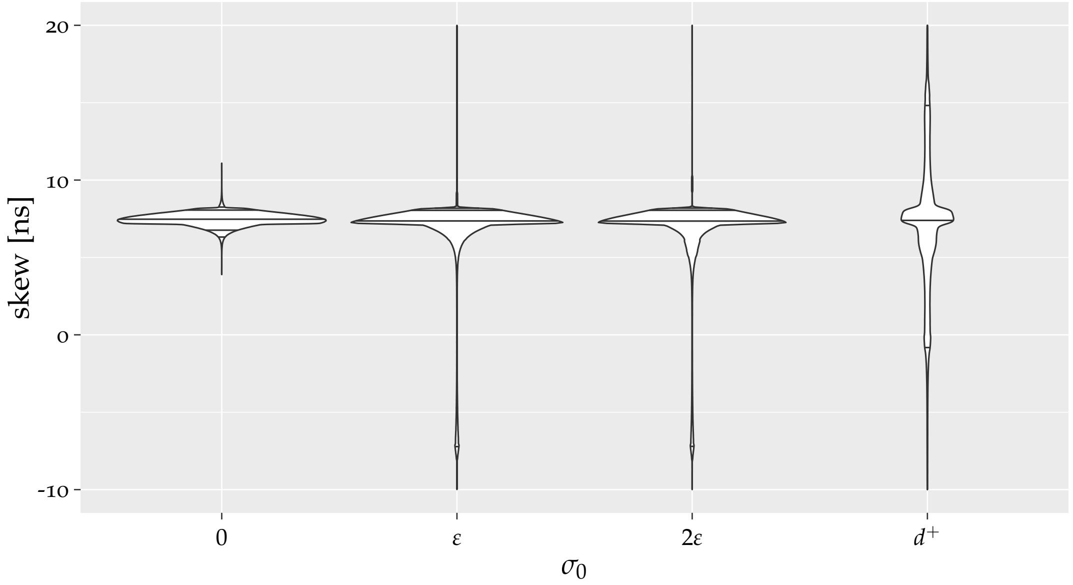
(b) inter-layer skews

Figure A.53: Skew in $\mathcal{G}_F^{13,15}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



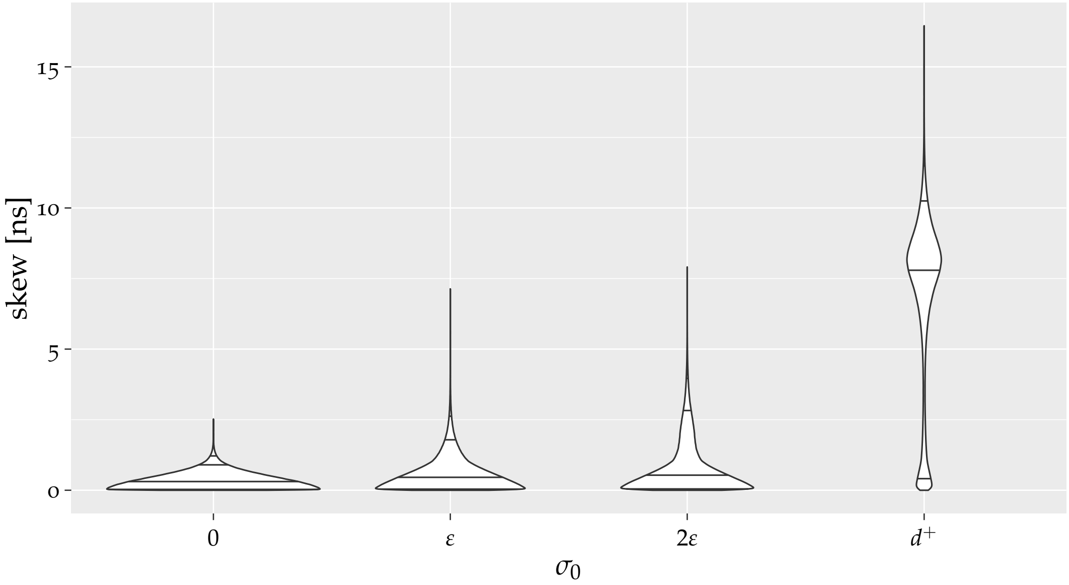
(a) intra-layer skews



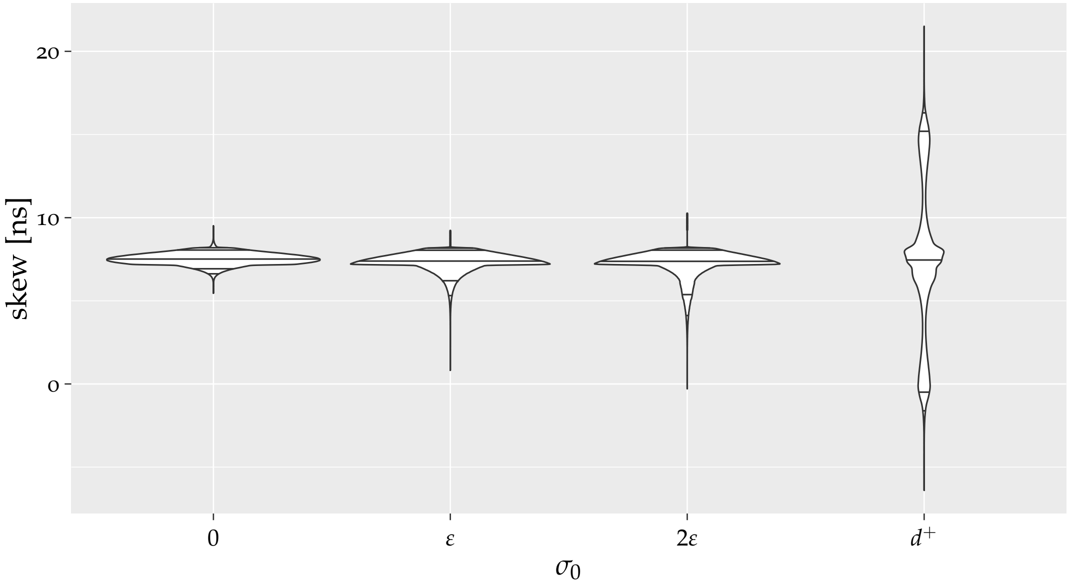
(b) inter-layer skews

Figure A.54: Skew in $\mathcal{G}_F^{13,15}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



(a) intra-layer skews

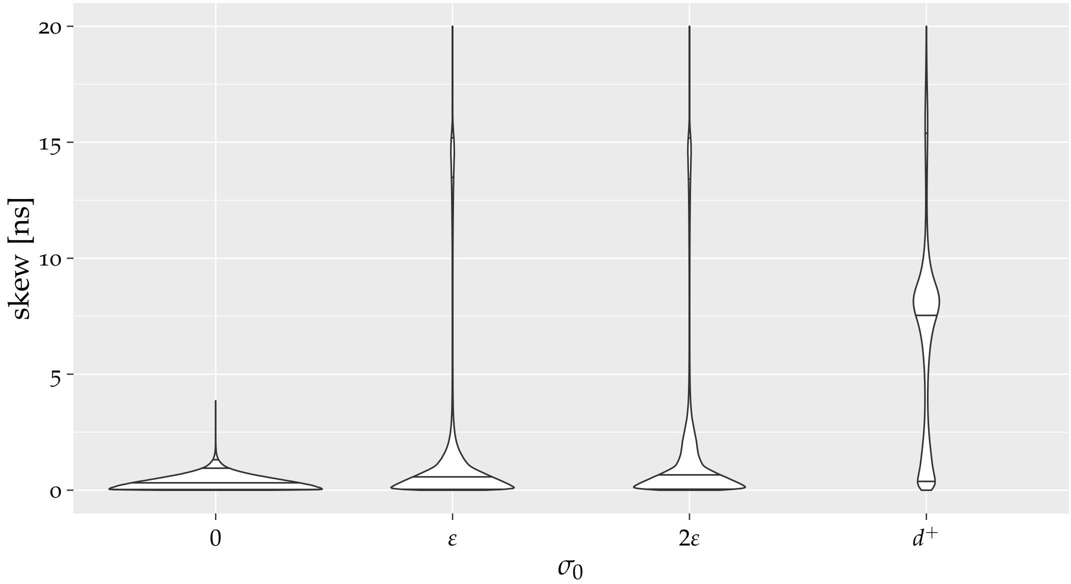


(b) inter-layer skews

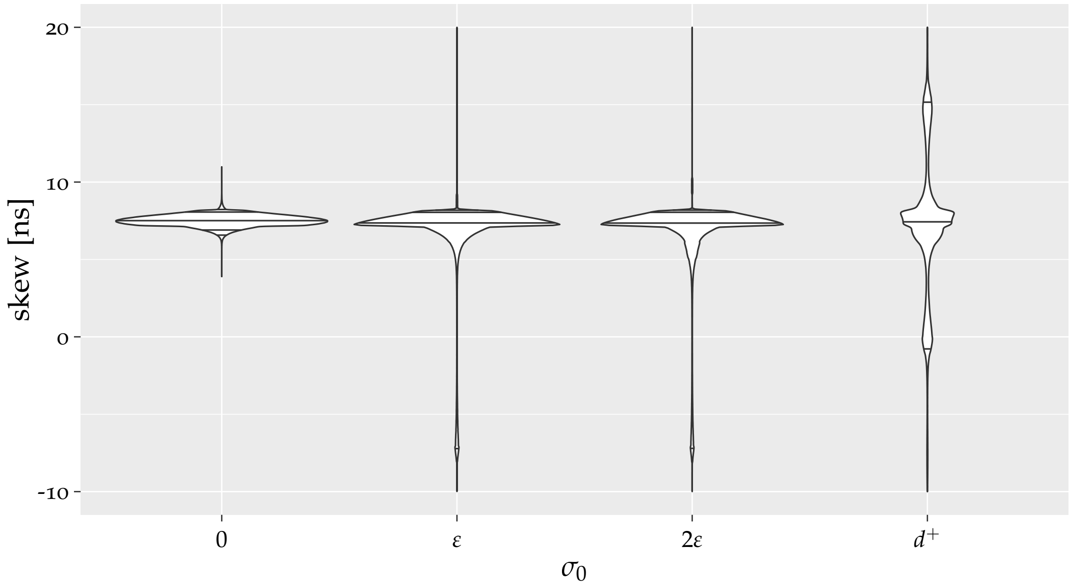
Figure A.55: Skew in $\mathcal{G}_F^{14,15}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

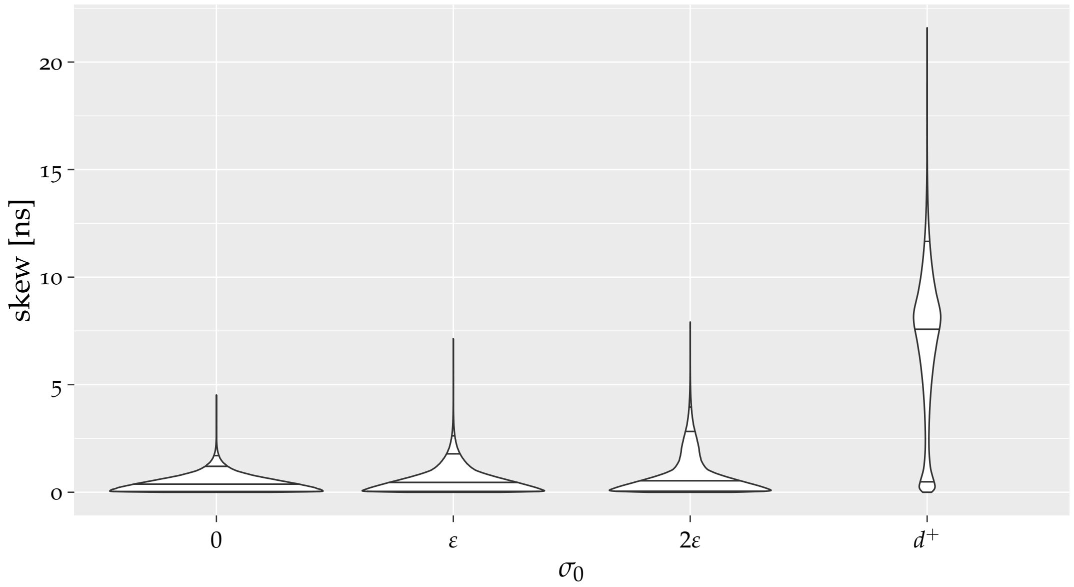


(a) intra-layer skews

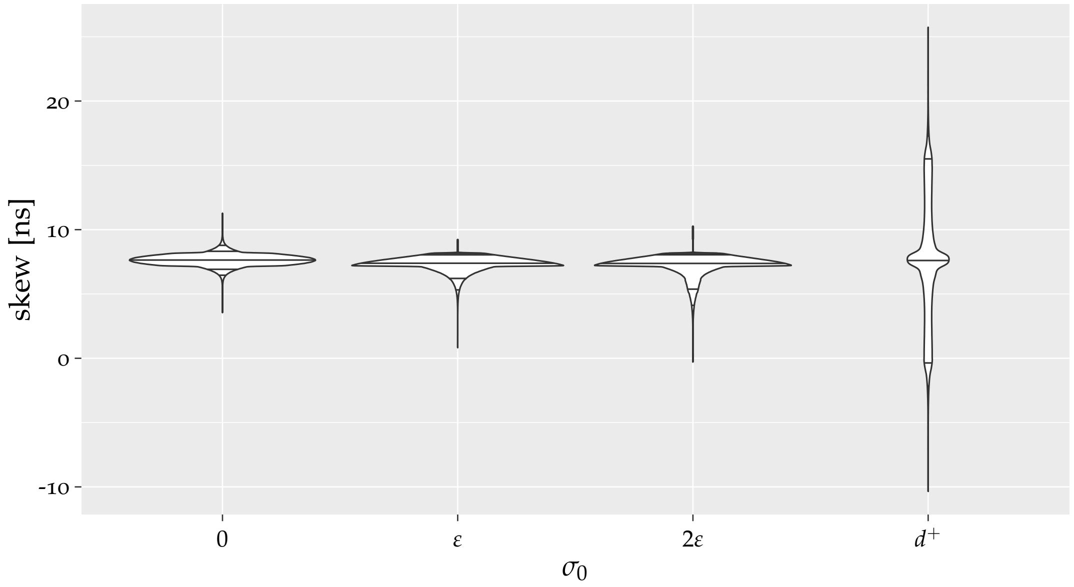


(b) inter-layer skews

Figure A.56: Skew in $\mathcal{G}_F^{14,15}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.



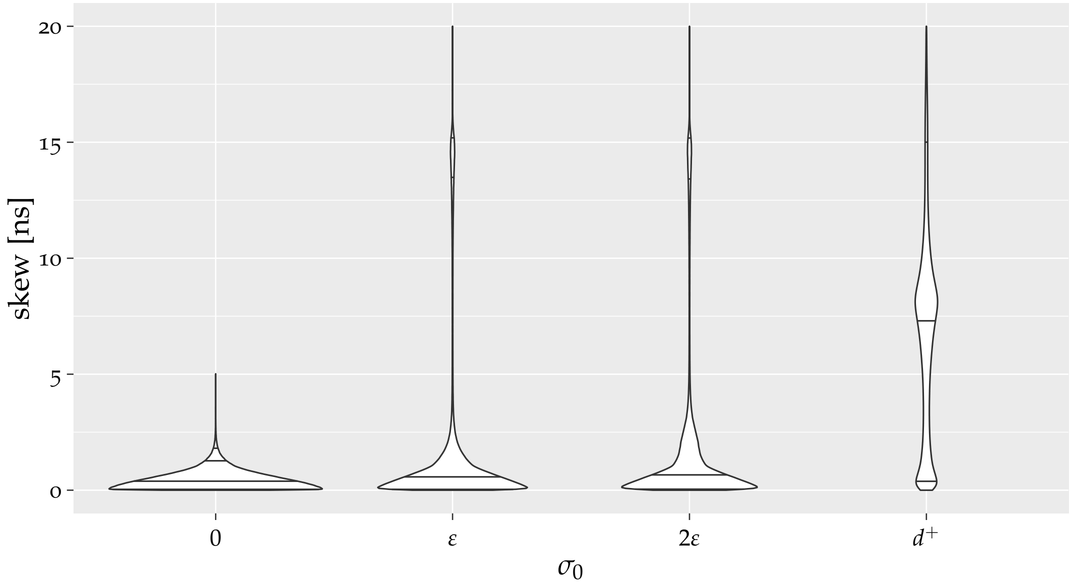
(a) intra-layer skews



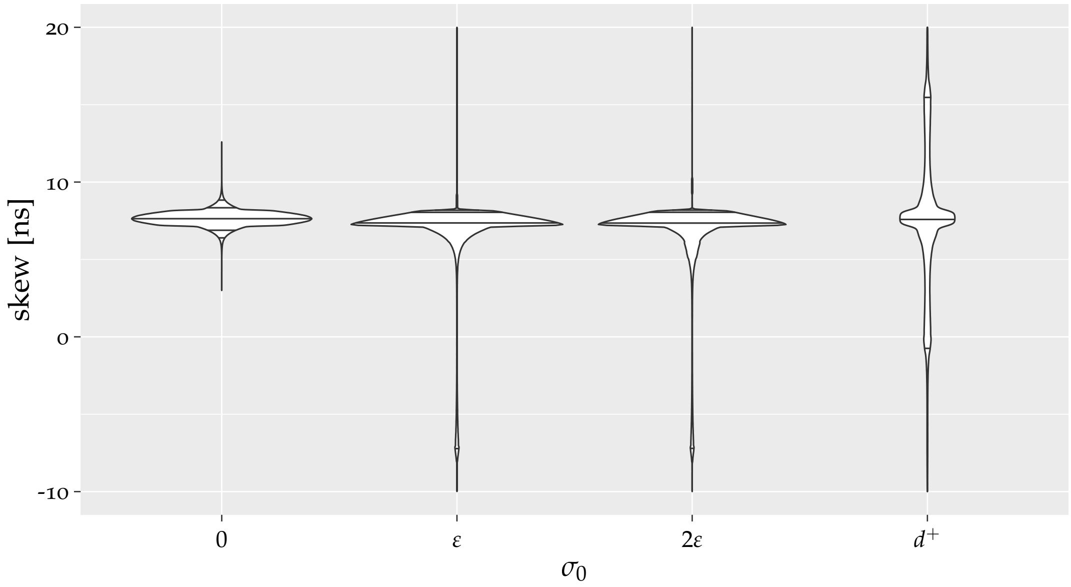
(b) inter-layer skews

Figure A.57: Skew in $\mathcal{G}_F^{14,16}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



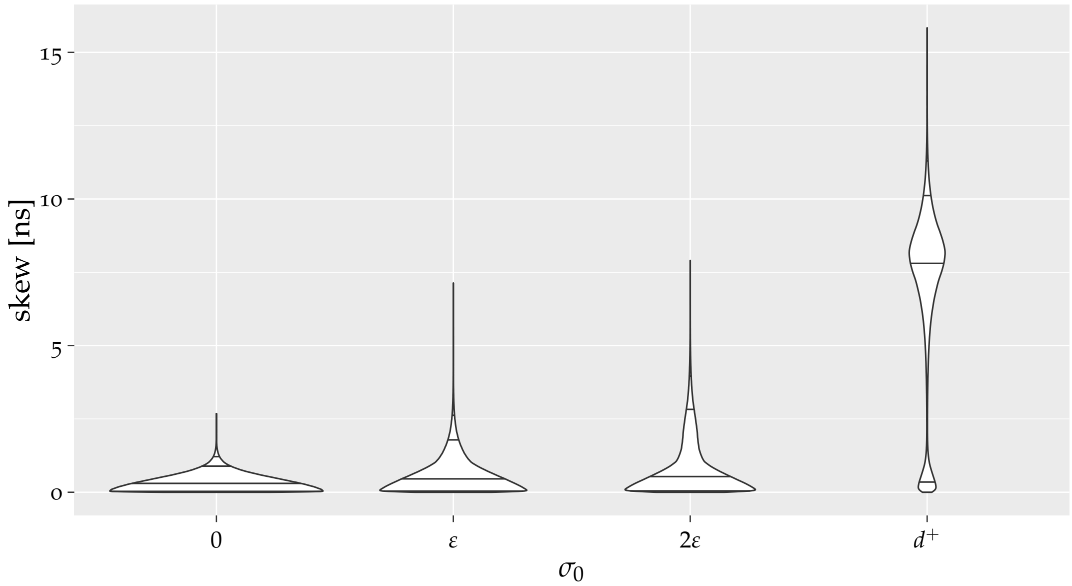
(a) intra-layer skews



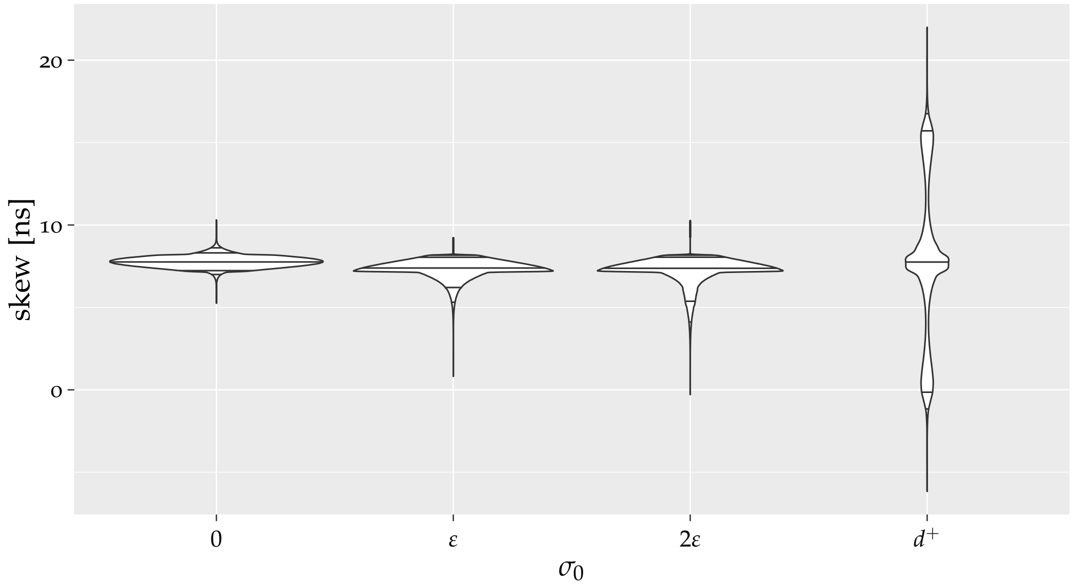
(b) inter-layer skews

Figure A.58: Skew in $\mathcal{G}_F^{14,16}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub



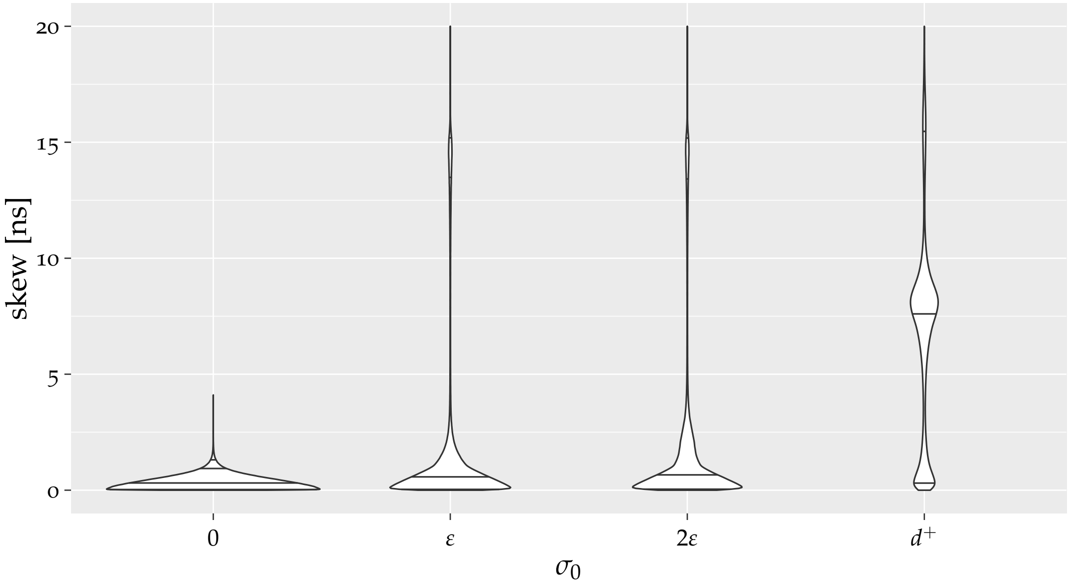
(a) intra-layer skews



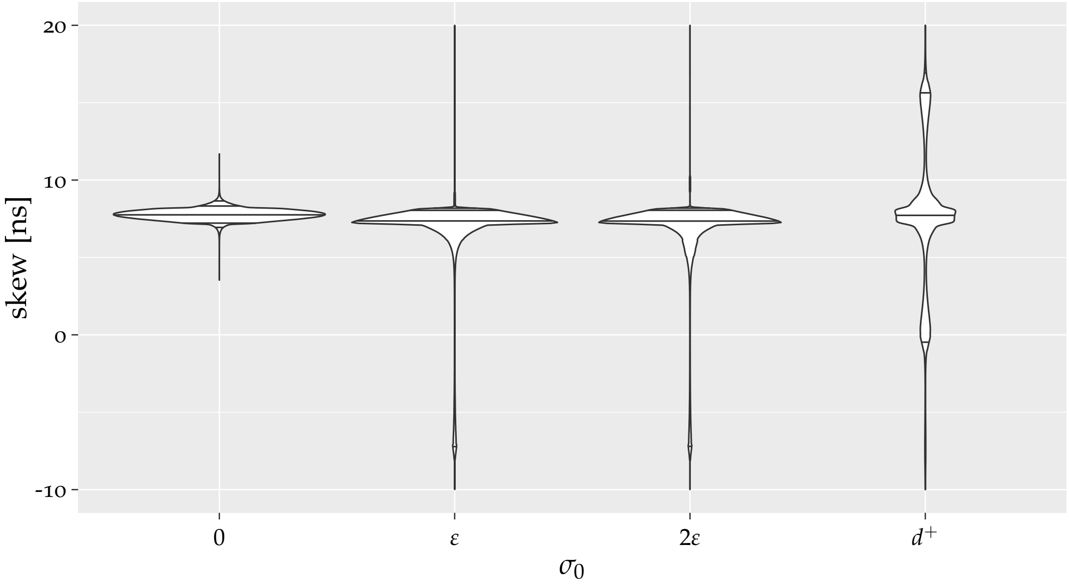
(b) inter-layer skews

Figure A.59: Skew in $\mathcal{G}_F^{15,16}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



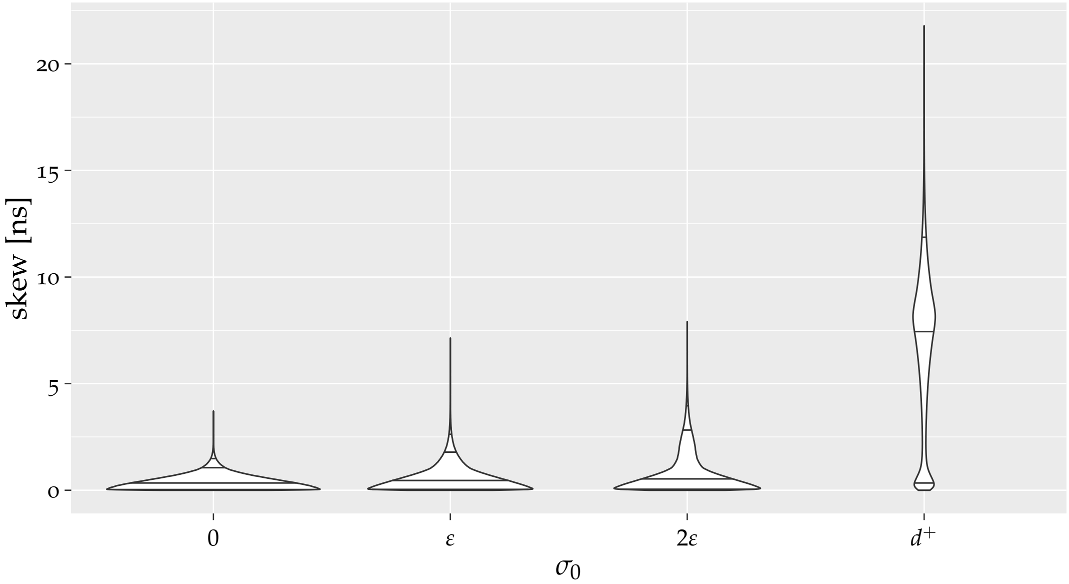
(a) intra-layer skews



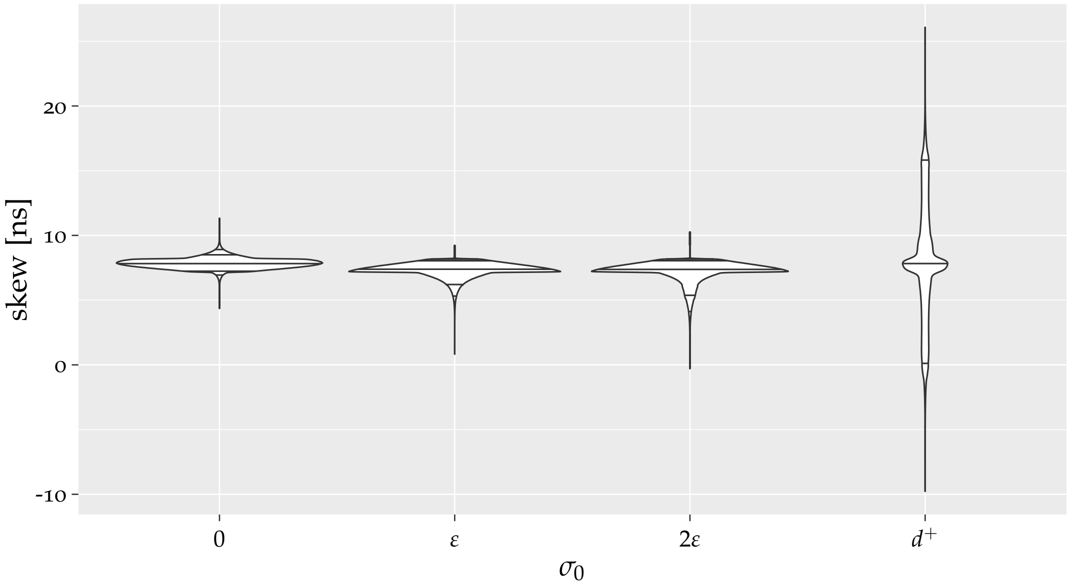
(b) inter-layer skews

Figure A.60: Skew in $\mathcal{G}_F^{15,16}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



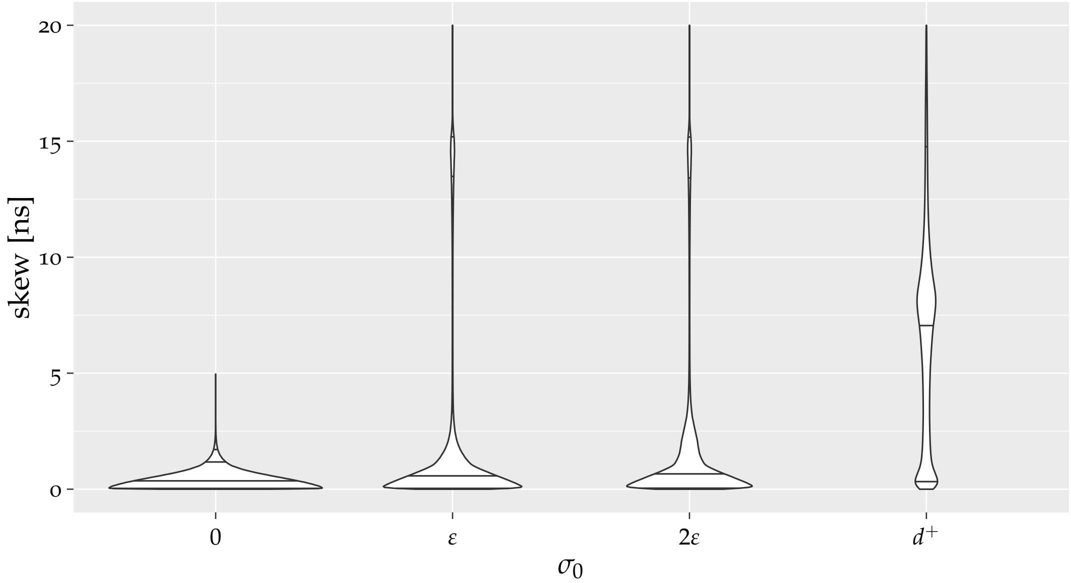
(a) intra-layer skews



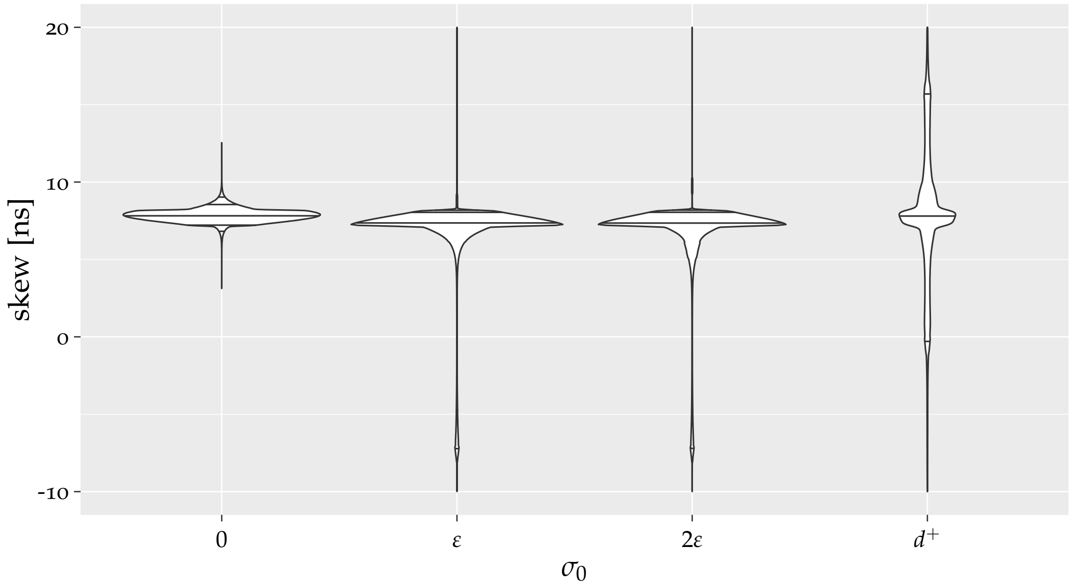
(b) inter-layer skews

Figure A.61: Skew in $\mathcal{G}_F^{15,17}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



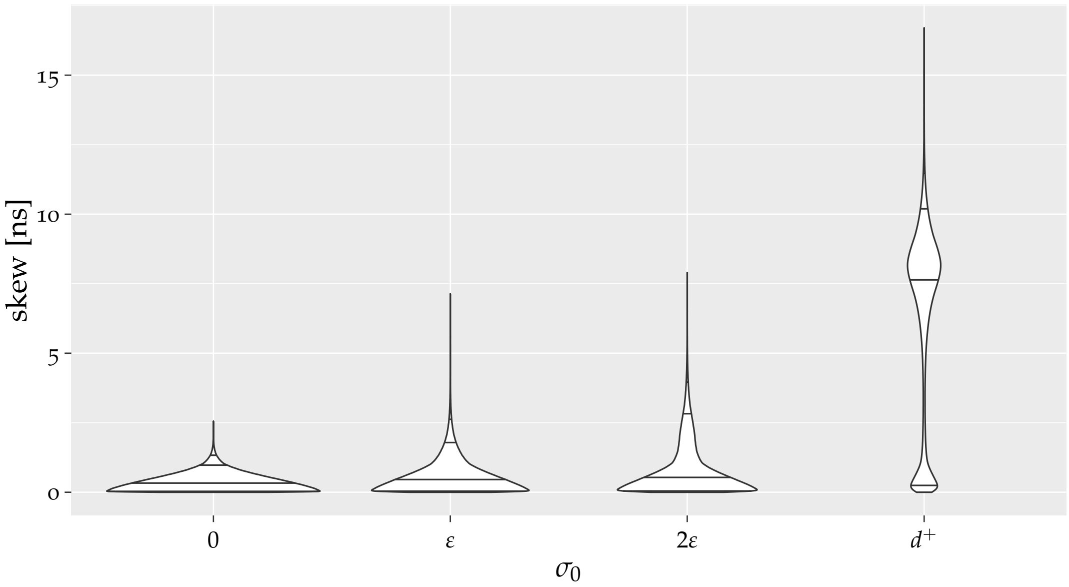
(a) intra-layer skews



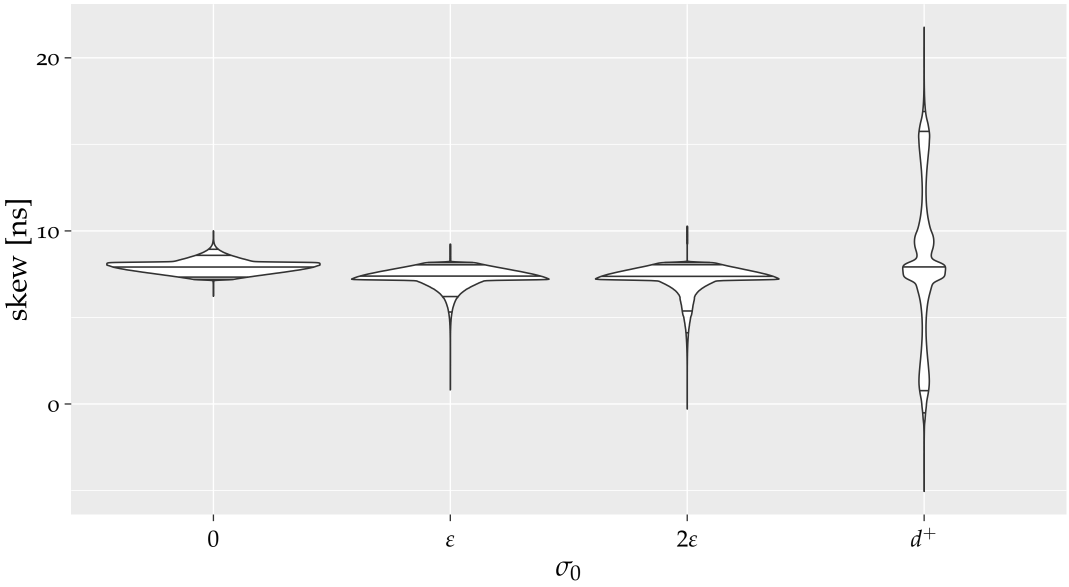
(b) inter-layer skews

Figure A.62: Skew in $\mathcal{G}_F^{15,17}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



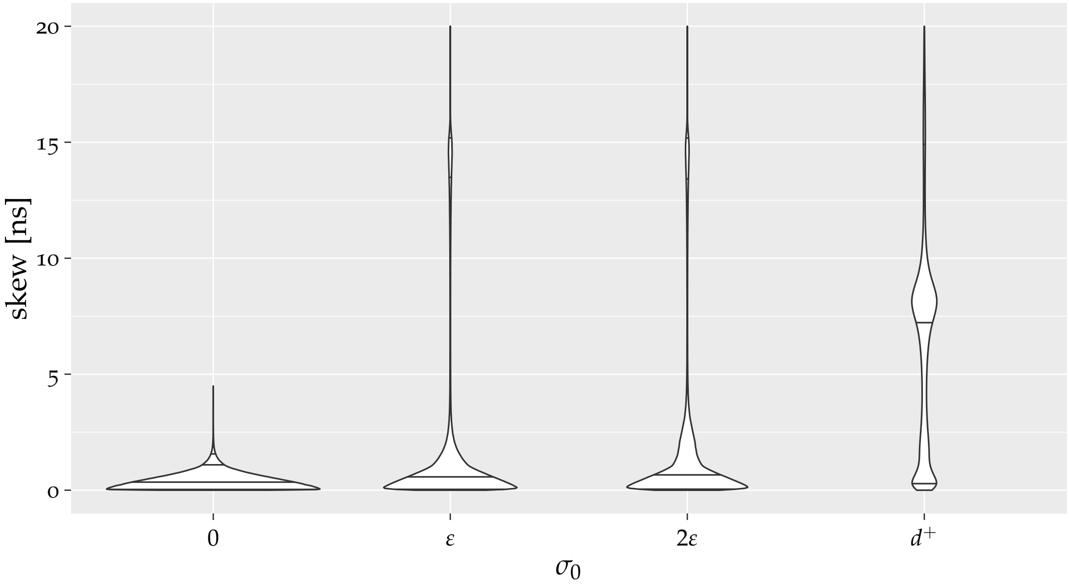
(a) intra-layer skews



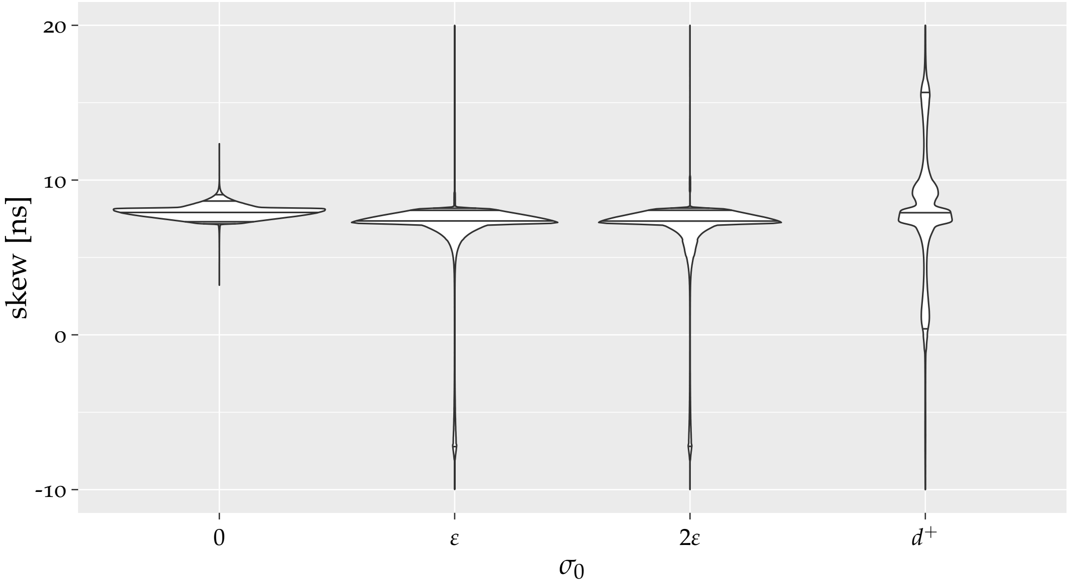
(b) inter-layer skews

Figure A.63: Skew in $\mathcal{G}_F^{16,17}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



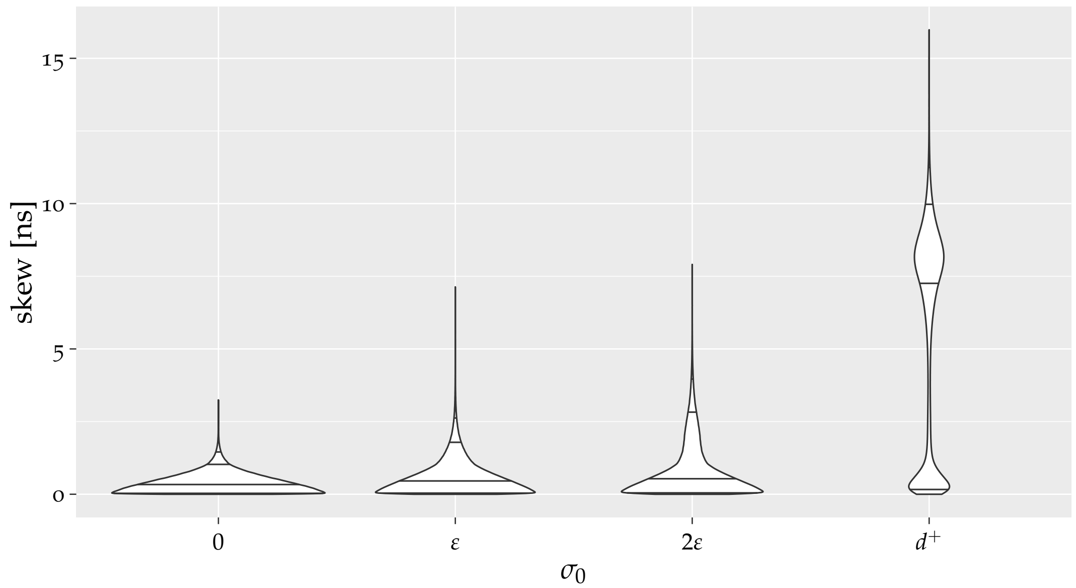
(a) intra-layer skews



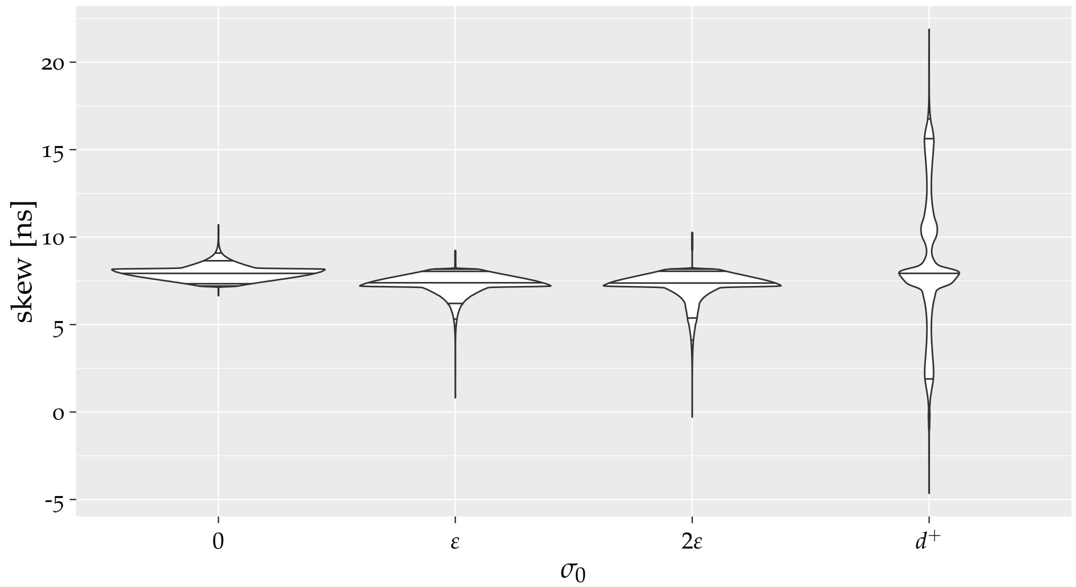
(b) inter-layer skews

Figure A.64: Skew in $\mathcal{G}_F^{16,17}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skews

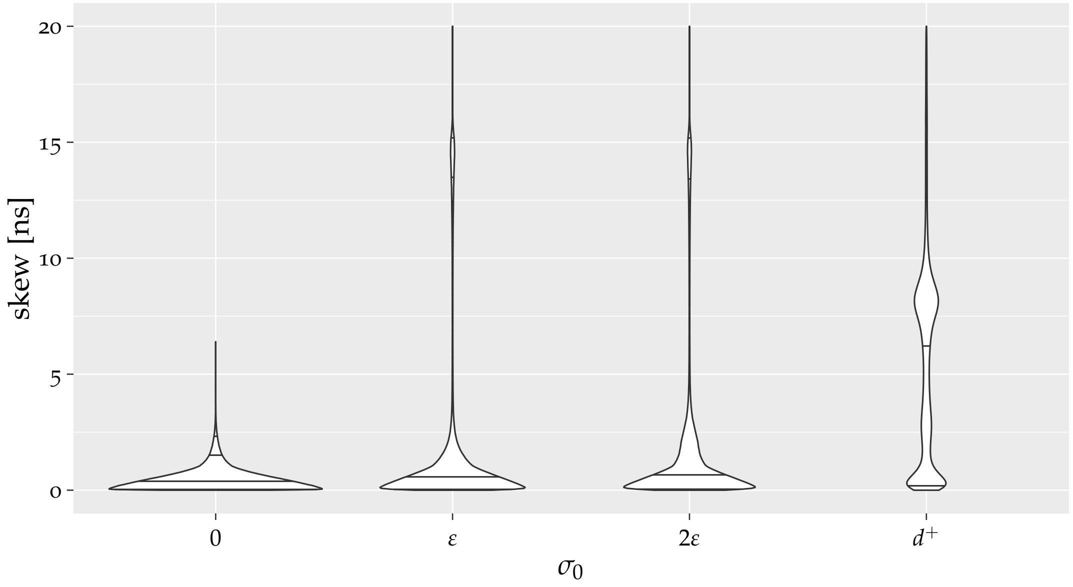


(b) inter-layer skews

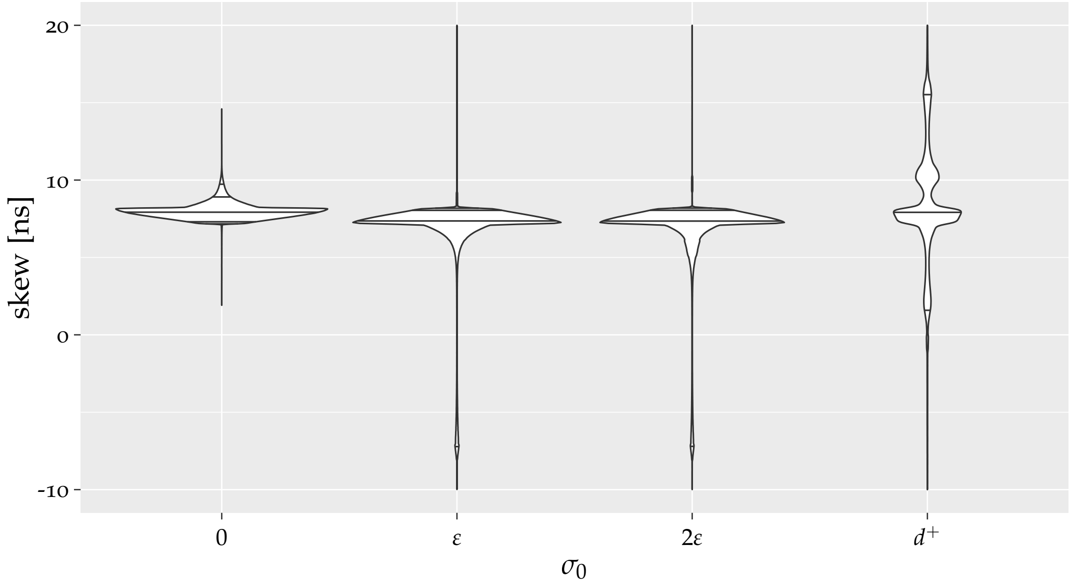
Figure A.65: Skew in $\mathcal{G}_F^{17,18}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skews



(b) inter-layer skews

Figure A.66: Skew in $\mathcal{G}_F^{17,18}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

A.12 Table for the Pulse-based Algorithms in $\mathcal{G}_F$

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.3: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_F^{13,14}$ topology.

			intra-layer							inter-layer							
σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
0	0	0	0.01	0.03	0.40	0.33	1.00	1.39	3.02	5.05	6.33	6.73	7.42	7.43	8.02	8.15	8.97
	15	0	0.01	0.03	0.43	0.35	1.12	1.60	4.10	3.38	6.19	6.65	7.41	7.43	8.03	8.16	11.70
	15	1	0.01	0.03	0.43	0.35	1.10	1.54	4.10	3.38	6.23	6.67	7.41	7.43	8.02	8.15	10.82
ε	0	0	0.01	0.05	0.65	0.53	1.68	2.13	3.71	4.36	5.81	6.27	7.34	7.40	8.07	8.19	9.62
	15	0	0.01	0.05	0.70	0.56	1.79	2.36	6.32	1.26	5.70	6.21	7.33	7.40	8.09	8.34	13.35
	15	1	0.01	0.05	0.69	0.56	1.77	2.31	6.32	1.26	5.73	6.22	7.33	7.40	8.08	8.27	12.74
2ε	0	0	0.01	0.06	1.09	0.84	2.64	3.16	5.24	2.82	5.13	5.63	7.23	7.40	8.19	8.85	11.12
	15	0	0.01	0.07	1.19	0.95	2.86	4.34	11.27	−3.45	4.89	5.58	7.22	7.40	8.24	9.39	17.44
	15	1	0.01	0.07	1.18	0.95	2.84	4.26	11.27	−3.43	4.92	5.58	7.21	7.39	8.22	9.30	17.44
d^+	0	0	0.04	0.20	6.20	7.52	10.13	11.40	16.47	−6.16	−1.51	−0.38	6.98	7.40	14.15	15.54	20.08
	15	0	0.04	0.21	6.27	6.96	15.09	19.49	37.79	−26.65	−7.18	−0.66	7.03	7.41	14.54	21.09	42.81
	15	1	0.04	0.21	6.26	6.98	14.96	19.43	37.79	−26.65	−7.00	−0.65	7.02	7.40	14.52	20.85	40.49

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.4: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_F^{13,15}$ topology.

			intra-layer							inter-layer							
σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
0	0	0	0.01	0.03	0.40	0.33	0.99	1.38	3.13	4.57	6.44	6.82	7.47	7.48	8.05	8.19	10.14
	15	0	0.01	0.03	0.42	0.35	1.10	1.58	4.38	3.91	6.33	6.77	7.47	7.48	8.07	8.26	11.09
	15	1	0.01	0.03	0.42	0.34	1.07	1.54	4.38	3.91	6.36	6.78	7.47	7.48	8.06	8.22	10.84
ε	0	0	0.01	0.06	0.77	0.63	1.95	2.60	5.54	2.55	5.66	6.28	7.40	7.45	8.20	8.91	12.30
	15	0	0.01	0.06	0.80	0.64	2.08	2.91	6.95	0.75	5.54	6.23	7.40	7.45	8.25	9.04	13.63
	15	1	0.01	0.06	0.80	0.64	2.06	2.87	6.95	1.17	5.56	6.24	7.40	7.45	8.23	9.00	13.53
2ε	0	0	0.02	0.09	1.44	1.23	3.54	4.64	9.78	−0.50	4.36	5.39	7.32	7.44	9.03	10.09	14.76
	15	0	0.02	0.09	1.50	1.23	3.82	5.38	12.47	−5.28	4.05	5.35	7.33	7.44	9.10	10.38	19.73
	15	1	0.02	0.09	1.49	1.23	3.79	5.33	12.47	−5.28	4.09	5.36	7.33	7.44	9.09	10.30	19.73
d^+	0	0	0.05	0.28	6.62	7.34	11.88	14.12	22.35	−12.48	−2.62	−0.52	7.18	7.43	14.86	16.87	26.09
	15	0	0.05	0.26	6.65	6.89	15.27	20.11	36.77	−25.86	−6.93	−0.90	7.22	7.44	15.41	21.33	40.80
	15	1	0.05	0.27	6.65	6.90	15.15	20.04	36.77	−25.67	−6.75	−0.88	7.22	7.44	15.39	21.09	40.80

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.5: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\mathrm{op}}$ (in [ns]) in the $\mathcal{G}_F^{14,15}$ topology.

			intra-layer							inter-layer							
σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
0	0	0	0.01	0.03	0.36	0.31	0.90	1.22	2.51	5.46	6.63	6.93	7.51	7.52	8.06	8.19	9.51
	15	0	0.01	0.03	0.38	0.32	0.95	1.32	3.85	3.90	6.58	6.91	7.51	7.52	8.07	8.24	10.97
	15	1	0.01	0.03	0.38	0.32	0.94	1.29	3.72	3.90	6.60	6.92	7.51	7.51	8.06	8.20	10.75
ε	0	0	0.02	0.07	0.81	0.74	1.81	2.31	4.87	3.77	6.00	6.44	7.46	7.49	8.34	8.85	11.09
	15	0	0.01	0.07	0.83	0.73	1.95	2.73	6.25	1.33	5.84	6.42	7.47	7.49	8.37	9.02	13.23
	15	1	0.01	0.07	0.83	0.73	1.94	2.70	6.25	1.33	5.86	6.42	7.46	7.49	8.37	8.98	13.08
2ε	0	0	0.03	0.15	1.60	1.57	3.32	4.13	7.39	1.01	4.78	5.57	7.42	7.48	9.20	10.02	13.46
	15	0	0.03	0.13	1.63	1.48	3.76	5.28	11.83	-3.73	4.21	5.51	7.43	7.49	9.29	10.58	18.46
	15	1	0.03	0.13	1.62	1.48	3.73	5.25	11.83	-3.73	4.26	5.51	7.42	7.48	9.27	10.50	18.23
d^+	0	0	0.07	0.39	6.90	7.80	10.26	11.50	16.45	-6.40	-1.57	-0.46	7.37	7.48	15.22	16.31	21.50
	15	0	0.06	0.31	6.91	7.56	15.76	19.07	35.52	-25.43	-7.68	-0.80	7.40	7.49	15.62	22.47	40.20
	15	1	0.06	0.32	6.90	7.56	15.68	18.98	35.52	-25.43	-7.56	-0.78	7.39	7.48	15.60	22.31	40.20

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.6: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\mathrm{op}}$ (in [ns]) in the $\mathcal{G}_F^{14,16}$ topology.

			intra-layer							inter-layer							
σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
0	0	0	0.01	0.03	0.46	0.38	1.21	1.70	4.52	3.56	6.47	6.92	7.64	7.64	8.32	8.79	11.26
	15	0	0.01	0.04	0.48	0.39	1.27	1.81	5.01	3.01	6.41	6.90	7.64	7.64	8.35	8.84	12.58
	15	1	0.01	0.04	0.48	0.39	1.25	1.79	5.01	3.01	6.43	6.91	7.64	7.64	8.34	8.82	12.58
ε	0	0	0.01	0.07	0.92	0.74	2.36	3.21	6.57	1.34	5.56	6.37	7.63	7.64	8.86	9.68	14.09
	15	0	0.01	0.07	0.93	0.74	2.46	3.43	8.02	−0.82	5.44	6.35	7.63	7.64	8.89	9.80	15.84
	15	1	0.01	0.07	0.93	0.74	2.45	3.41	8.02	0.16	5.47	6.36	7.63	7.64	8.88	9.77	15.84
2ε	0	0	0.03	0.13	1.71	1.53	4.00	5.20	10.43	−2.01	4.24	5.43	7.62	7.64	9.79	10.98	17.46
	15	0	0.02	0.12	1.72	1.48	4.28	5.75	13.53	−6.13	3.89	5.38	7.62	7.64	9.85	11.33	19.06
	15	1	0.02	0.12	1.72	1.48	4.25	5.73	11.87	−3.80	3.94	5.39	7.62	7.64	9.84	11.28	19.06
d^+	0	0	0.08	0.44	6.98	7.59	11.67	13.73	21.59	−10.34	−2.11	−0.32	7.60	7.64	15.53	17.30	25.72
	15	0	0.06	0.31	6.97	7.31	15.50	19.59	38.82	−28.14	−7.02	−0.78	7.61	7.64	16.02	22.24	44.18
	15	1	0.06	0.32	6.97	7.33	15.40	19.49	38.82	−28.14	−6.85	−0.76	7.61	7.64	16.00	22.07	44.18

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.7: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_F^{15,16}$ topology.

			intra-layer							inter-layer							
σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
0	0	0	0.01	0.03	0.36	0.30	0.89	1.21	2.68	5.26	7.00	7.24	7.76	7.76	8.31	8.62	10.30
	15	0	0.01	0.03	0.37	0.31	0.94	1.31	4.10	3.54	6.95	7.23	7.76	7.76	8.33	8.66	11.68
	15	1	0.01	0.03	0.37	0.31	0.93	1.29	3.62	4.46	6.97	7.23	7.76	7.76	8.33	8.65	11.68
ε	0	0	0.02	0.08	0.86	0.77	1.95	2.53	4.79	3.71	6.22	6.78	7.79	7.78	8.84	9.39	11.78
	15	0	0.02	0.08	0.87	0.76	2.08	2.85	6.22	2.34	6.05	6.74	7.78	7.77	8.86	9.55	13.30
	15	1	0.02	0.08	0.87	0.76	2.07	2.83	6.22	2.34	6.09	6.75	7.79	7.78	8.86	9.53	13.30
2ε	0	0	0.03	0.17	1.68	1.65	3.44	4.26	8.26	1.41	5.06	5.86	7.80	7.78	9.77	10.59	14.10
	15	0	0.03	0.14	1.70	1.56	3.88	5.23	10.36	−2.21	4.44	5.76	7.80	7.78	9.85	11.19	17.75
	15	1	0.03	0.14	1.69	1.56	3.86	5.20	10.36	−2.21	4.51	5.77	7.80	7.78	9.84	11.13	17.75
d^+	0	0	0.06	0.33	7.02	7.81	10.13	11.31	15.83	−6.16	−1.11	−0.11	7.81	7.78	15.73	16.77	21.99
	15	0	0.05	0.26	7.02	7.62	15.77	18.81	37.23	−28.92	−7.39	−0.47	7.80	7.78	16.08	23.02	44.70
	15	1	0.05	0.26	7.02	7.62	15.71	18.72	37.23	−28.48	−7.27	−0.46	7.81	7.78	16.07	22.91	44.70

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.8: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_F^{15,17}$ topology.

			intra-layer							inter-layer							
σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
0	0	0	0.01	0.03	0.42	0.34	1.06	1.49	3.71	4.36	6.95	7.25	7.83	7.83	8.51	8.92	11.33
	15	0	0.01	0.03	0.45	0.36	1.17	1.71	4.96	3.14	6.83	7.23	7.84	7.83	8.56	9.04	12.54
	15	1	0.01	0.03	0.44	0.36	1.14	1.65	4.96	3.14	6.88	7.24	7.84	7.83	8.54	8.99	12.54
ε	0	0	0.01	0.06	0.83	0.67	2.13	2.90	6.51	2.25	6.15	6.87	7.88	7.85	9.08	9.77	13.49
	15	0	0.01	0.06	0.86	0.68	2.26	3.18	7.24	0.81	6.01	6.83	7.88	7.85	9.11	9.90	14.48
	15	1	0.01	0.06	0.85	0.68	2.24	3.15	7.24	1.61	6.05	6.84	7.88	7.85	9.10	9.87	14.48
2ε	0	0	0.02	0.11	1.55	1.35	3.78	4.96	10.20	−1.04	4.89	5.99	7.94	7.86	9.97	11.12	16.85
	15	0	0.02	0.10	1.59	1.31	4.05	5.57	12.16	−2.89	4.55	5.94	7.93	7.86	10.02	11.44	19.33
	15	1	0.02	0.10	1.58	1.32	4.02	5.53	12.16	−2.82	4.62	5.95	7.93	7.86	10.01	11.40	18.97
d^+	0	0	0.06	0.29	6.74	7.43	11.86	14.09	21.78	−9.77	−1.77	0.16	8.01	7.86	15.85	17.89	26.07
	15	0	0.05	0.26	6.76	7.06	15.41	20.03	37.97	−29.98	−6.35	−0.33	7.99	7.86	16.27	22.39	45.67
	15	1	0.05	0.26	6.76	7.08	15.30	19.95	37.97	−29.98	−6.12	−0.31	8.00	7.87	16.25	22.21	45.67

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.9: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_F^{16,17}$ topology.

			intra-layer							inter-layer							
σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
0	0	0	0.01	0.03	0.39	0.33	0.97	1.33	2.55	6.23	7.21	7.34	7.92	7.92	8.59	8.94	10.00
	15	0	0.01	0.03	0.43	0.35	1.09	1.57	4.49	3.21	7.19	7.32	7.93	7.92	8.65	9.06	12.35
	15	1	0.01	0.03	0.42	0.34	1.05	1.48	4.49	4.56	7.20	7.33	7.93	7.92	8.63	9.02	12.35
ε	0	0	0.01	0.05	0.68	0.58	1.65	2.06	3.58	5.30	6.98	7.25	8.00	7.95	9.01	9.42	11.00
	15	0	0.01	0.05	0.72	0.61	1.77	2.36	6.08	2.11	6.81	7.23	8.00	7.95	9.06	9.53	13.76
	15	1	0.01	0.05	0.72	0.60	1.75	2.31	6.08	3.05	6.88	7.24	8.00	7.95	9.05	9.51	13.76
2ε	0	0	0.02	0.08	1.24	1.12	2.82	3.44	6.59	2.72	6.04	6.86	8.10	7.95	9.70	10.31	13.79
	15	0	0.02	0.08	1.32	1.14	3.15	4.77	11.35	−2.12	5.56	6.78	8.09	7.95	9.75	10.71	18.45
	15	1	0.02	0.08	1.32	1.14	3.12	4.73	11.35	−1.91	5.64	6.80	8.10	7.96	9.74	10.66	18.45
d^+	0	0	0.04	0.21	6.45	7.64	10.19	11.47	16.70	−5.05	−0.48	0.80	8.23	7.95	15.77	16.91	21.76
	15	0	0.04	0.22	6.49	7.22	15.44	19.57	37.78	−29.18	−6.30	0.38	8.21	7.95	16.09	22.84	45.23
	15	1	0.04	0.22	6.49	7.24	15.33	19.50	37.78	−28.22	−6.08	0.40	8.22	7.96	16.08	22.69	45.23

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

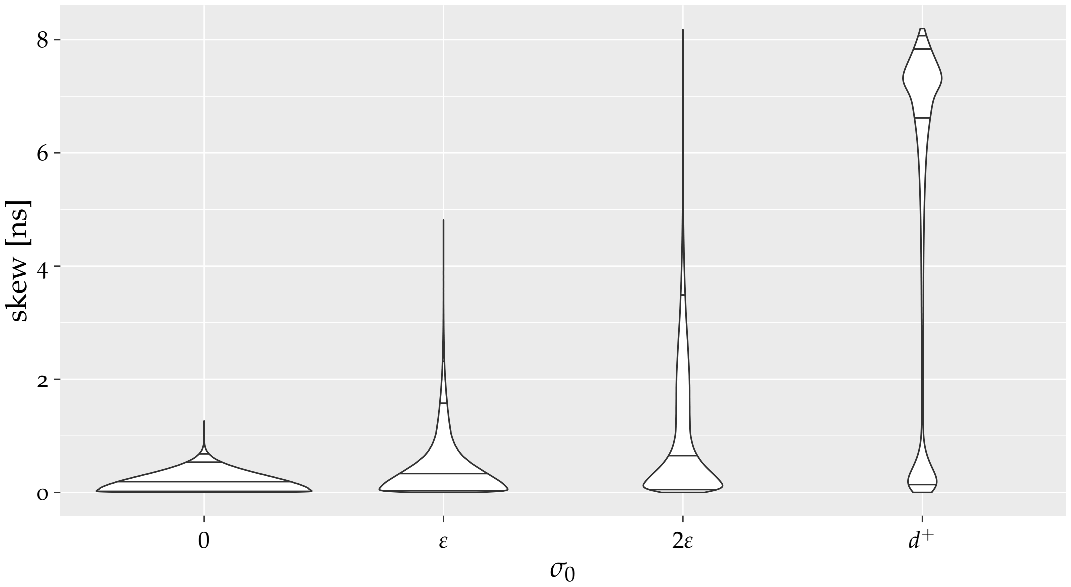
Table A.10: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_F^{17,18}$ topology.

			intra-layer							inter-layer							
σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
0	0	0	0.01	0.03	0.41	0.33	1.03	1.46	3.24	6.65	7.21	7.34	7.94	7.93	8.65	9.09	10.71
	15	0	0.01	0.03	0.52	0.39	1.51	2.33	6.40	1.92	7.19	7.32	7.98	7.94	8.92	9.74	14.59
	15	1	0.01	0.03	0.49	0.38	1.39	2.10	6.00	3.68	7.20	7.33	7.98	7.94	8.86	9.58	14.10
ε	0	0	0.01	0.04	0.62	0.48	1.74	2.32	3.74	6.13	7.18	7.30	8.03	7.96	9.14	9.79	11.38
	15	0	0.01	0.05	0.72	0.54	2.00	2.69	7.59	0.42	7.12	7.28	8.06	7.97	9.36	10.03	15.81
	15	1	0.01	0.05	0.70	0.53	1.94	2.58	7.59	2.65	7.17	7.29	8.06	7.97	9.31	9.96	15.30
2ε	0	0	0.01	0.05	0.91	0.60	2.56	3.04	4.74	5.09	7.00	7.26	8.13	7.97	9.83	10.36	11.84
	15	0	0.01	0.06	1.03	0.72	2.71	3.62	11.90	−2.98	6.52	7.24	8.17	7.97	9.93	10.50	19.90
	15	1	0.01	0.06	1.02	0.71	2.68	3.49	11.19	−1.96	6.75	7.25	8.17	7.98	9.91	10.47	19.07
d^+	0	0	0.02	0.12	5.57	7.19	9.95	11.22	15.98	−4.65	0.39	1.92	8.53	7.95	15.65	16.77	21.87
	15	0	0.03	0.14	5.71	6.06	14.36	19.62	47.01	−36.17	−4.56	1.59	8.52	7.96	15.89	22.01	54.14
	15	1	0.03	0.14	5.70	6.11	14.20	19.56	47.01	−36.17	−4.26	1.61	8.53	7.97	15.89	21.83	54.14

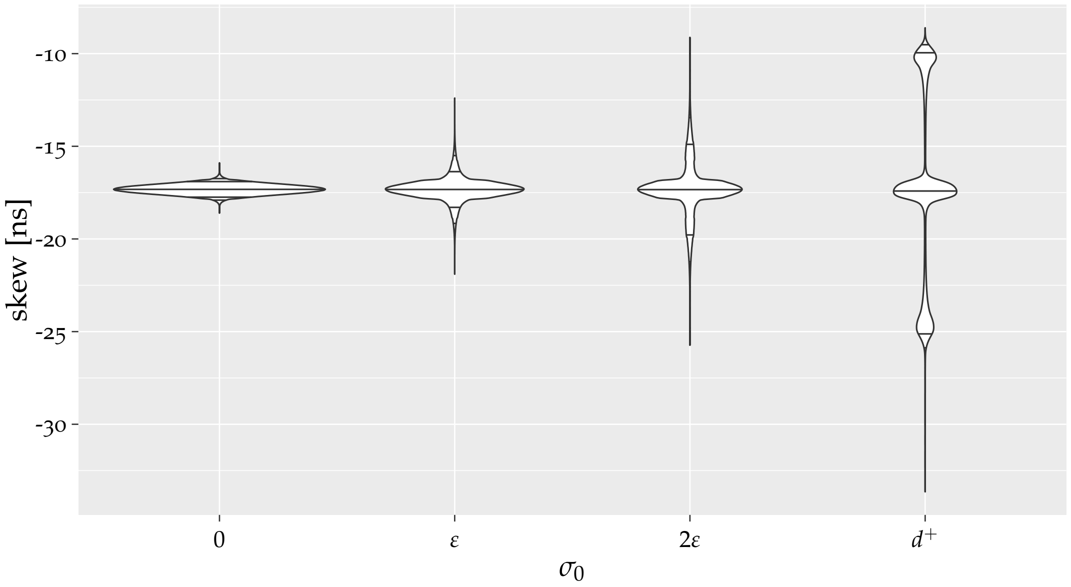
A.13 Plots with the 2–local fault Model in $\mathcal{G}_C$

In this section we also present the plots of the algorithms $\mathcal{A}^{3,2f}_{\text{adj}}$, $\mathcal{A}^{3,2f}_{\text{any}}$,$\mathcal{A}^{0,2f}_{\text{avg},\,25\,\text{ns}}$, $\mathcal{A}^{0,2f}_{\text{mid},\,25\,\text{ns}}$, $\mathcal{A}^{1,2f}_{\text{avg},\,25\,\text{ns}}$, $\mathcal{A}^{1,2f}_{\text{mid},\,25\,\text{ns}}$, $\mathcal{A}^{2,2f}_{\text{avg},\,25\,\text{ns}}$, and $\mathcal{A}^{2,2f}_{\text{mid},\,25\,\text{ns}}$. We omit the common offset of 25 ns in the algorithm description. Further, recall that the provided inter-layer skews have been compensated by 25 ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



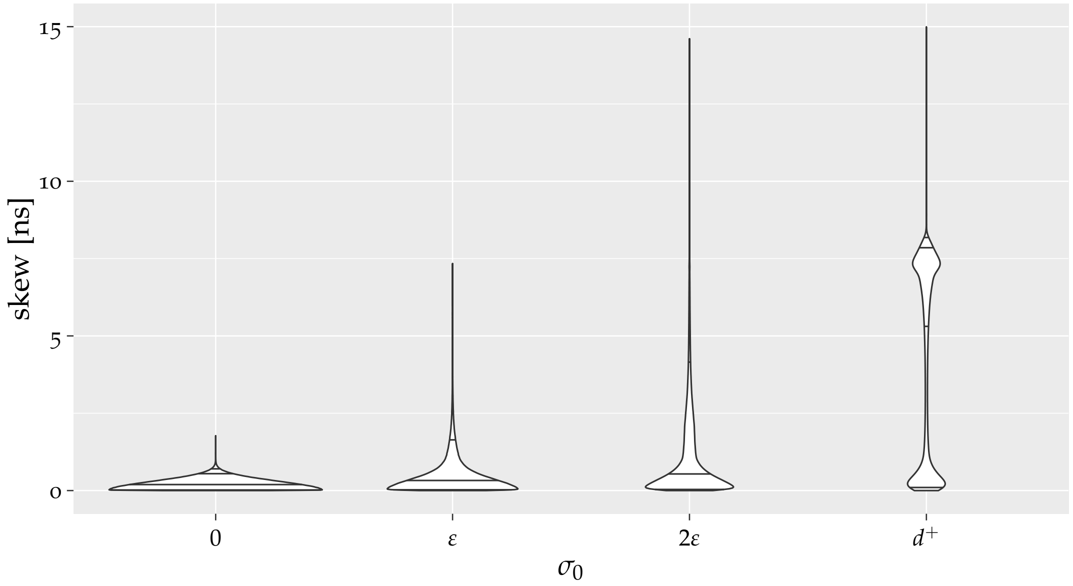
(a) intra-layer skews



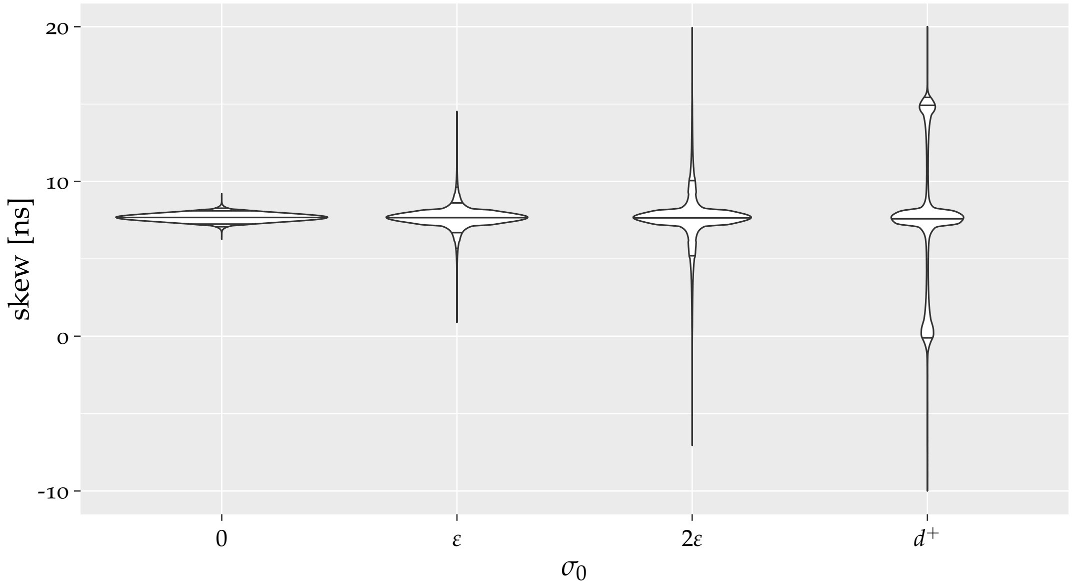
(b) inter-layer skews

Figure A.67: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{any}}^{3,2f}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



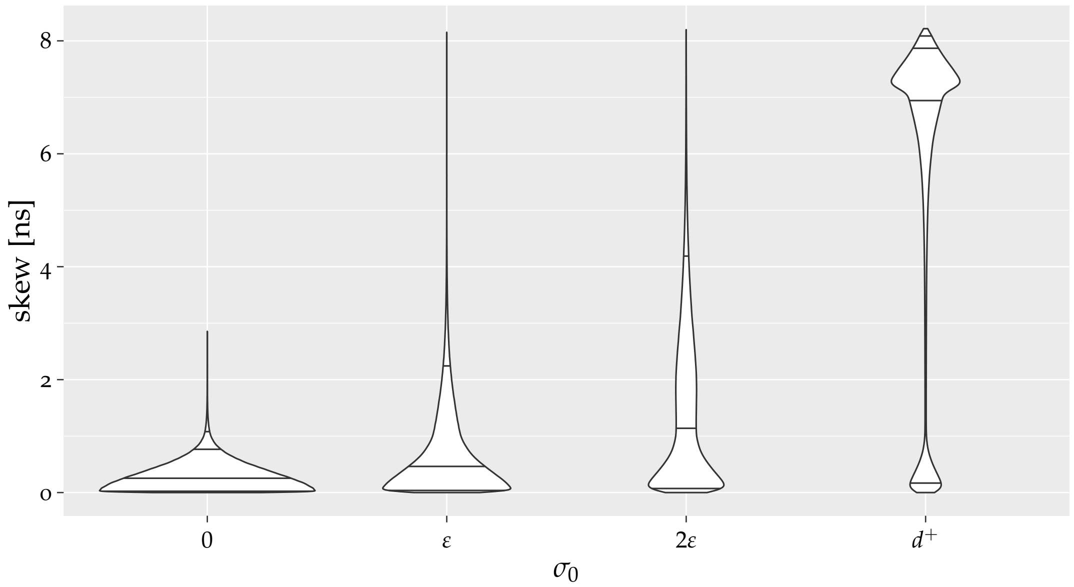
(a) intra-layer skews



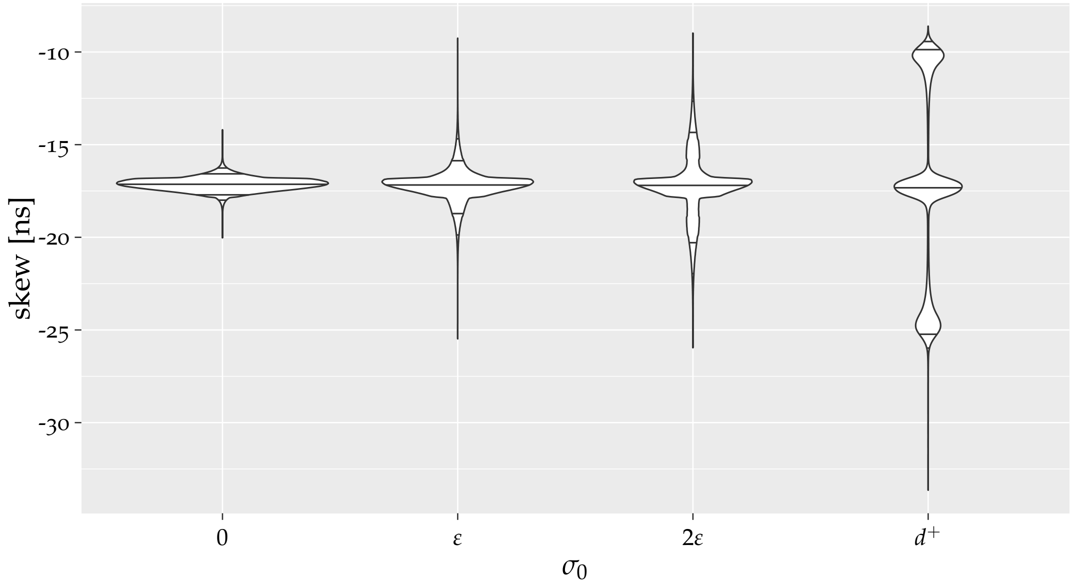
(b) inter-layer skews

Figure A.68: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{any}}^{3,2f}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-10, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



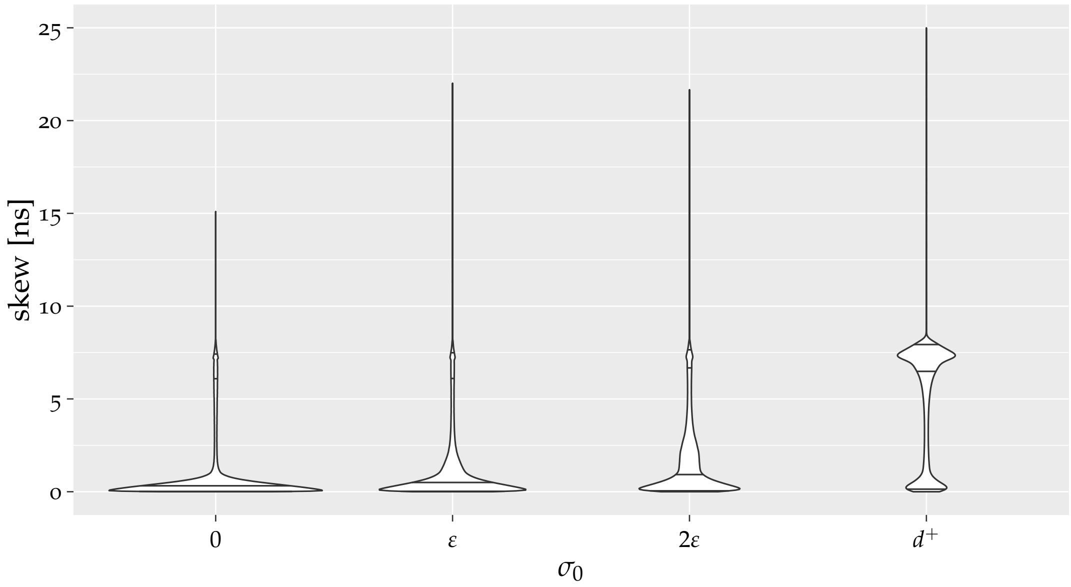
(a) intra-layer skews



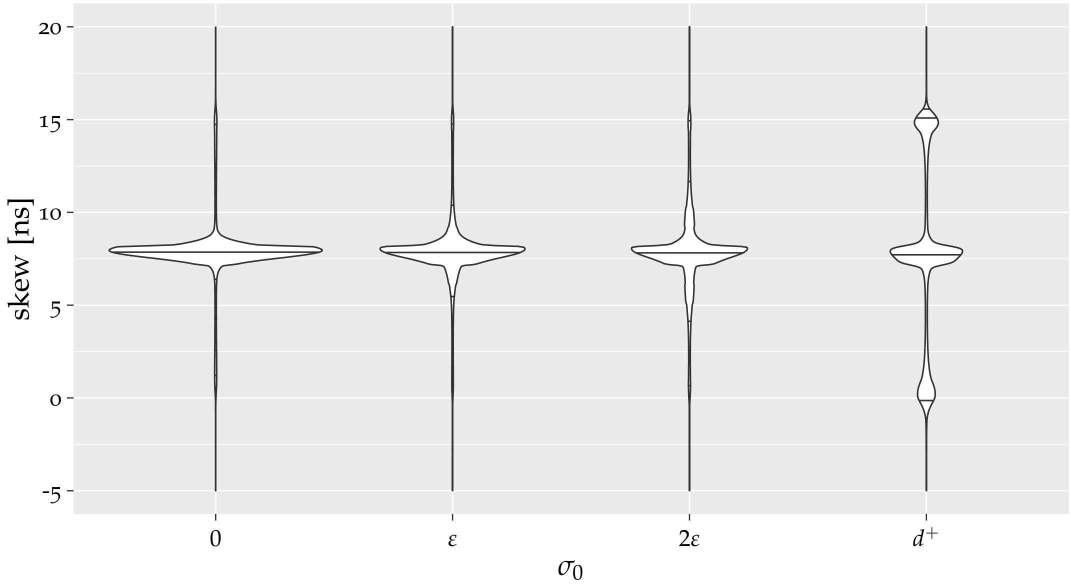
(b) inter-layer skews

Figure A.69: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{adj}}^{3,2f}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



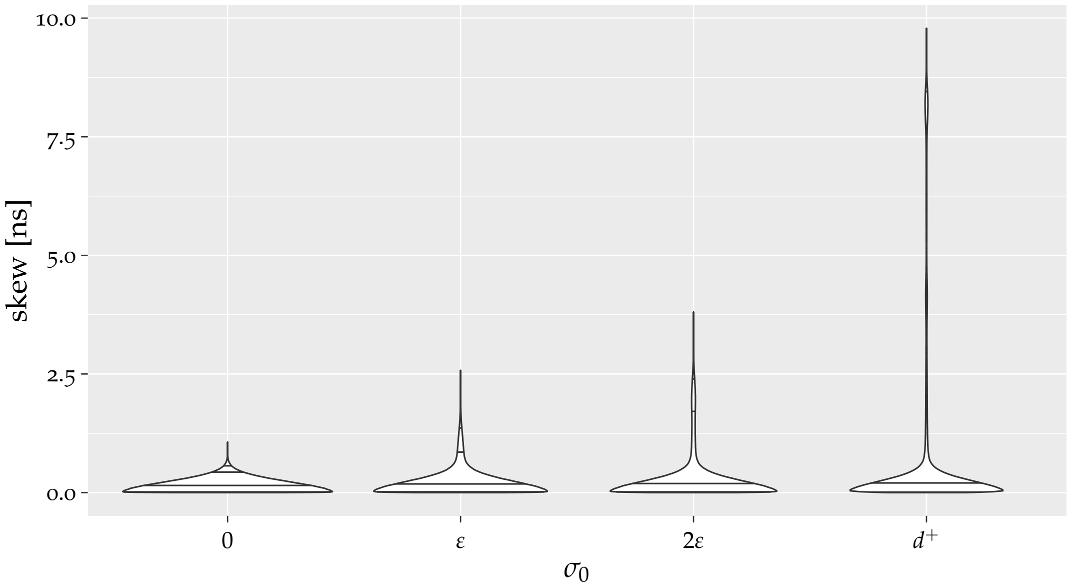
(a) intra-layer skews



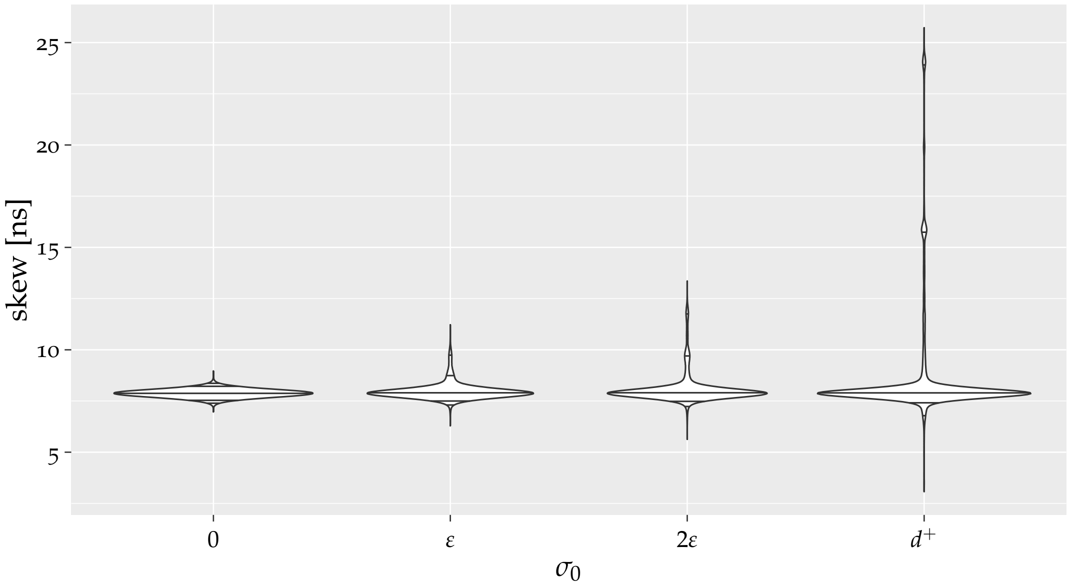
(b) inter-layer skews

Figure A.70: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{adj}}^{3,2f}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 25 ns, while the inter-layer skews have been limited to be within $[-5, 20]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



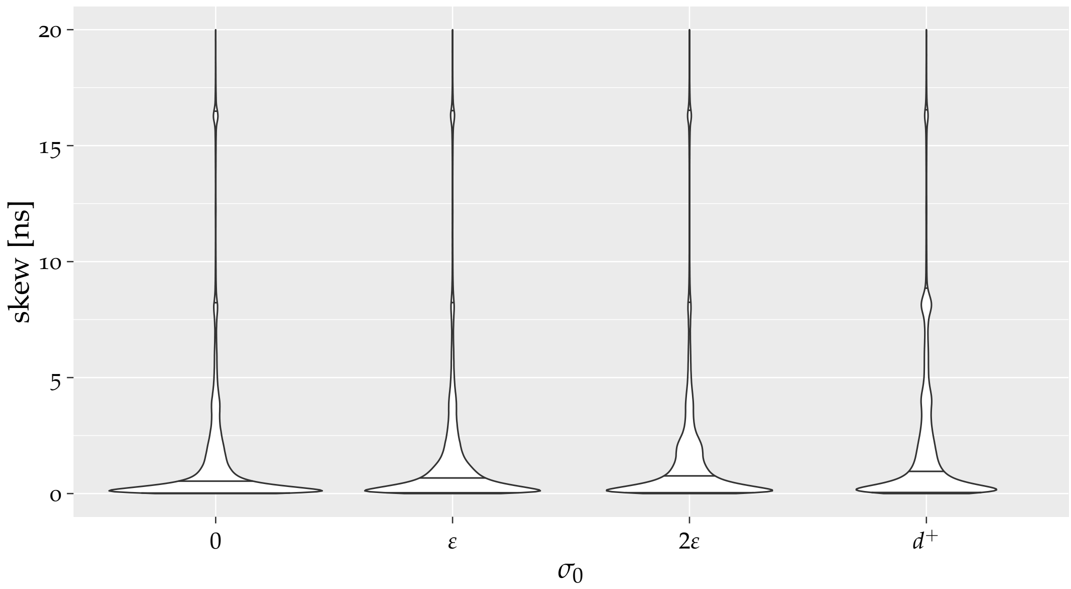
(a) intra-layer skews



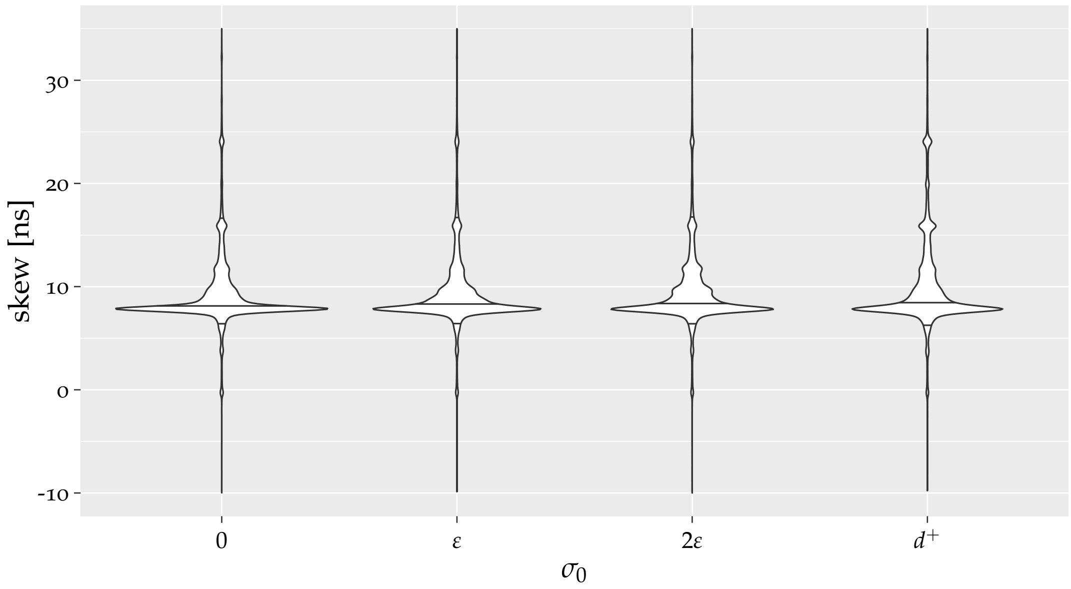
(b) inter-layer skews

Figure A.71: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{mid}}^{0,2f}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



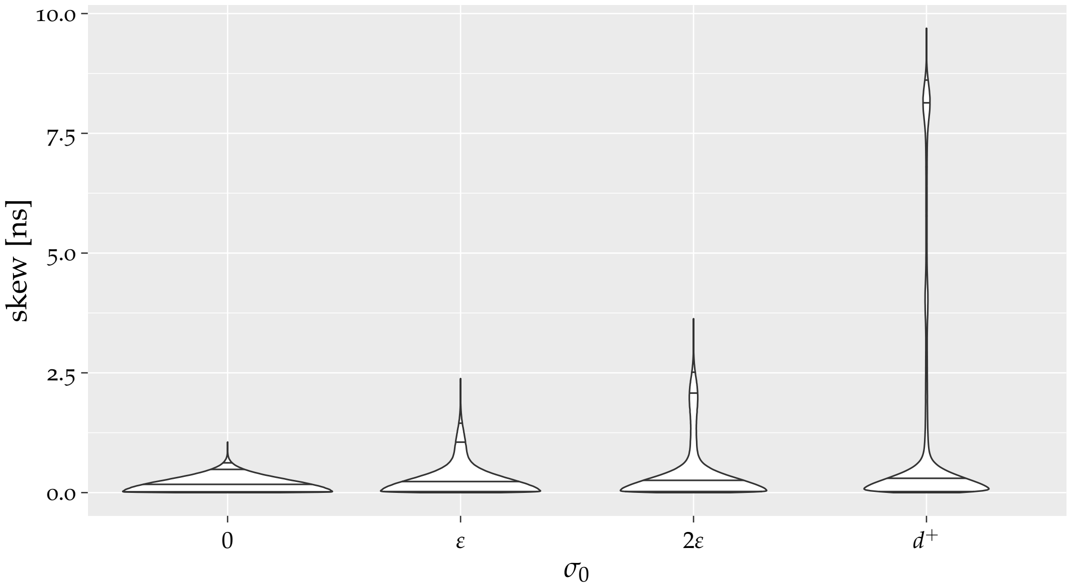
(a) intra-layer skews



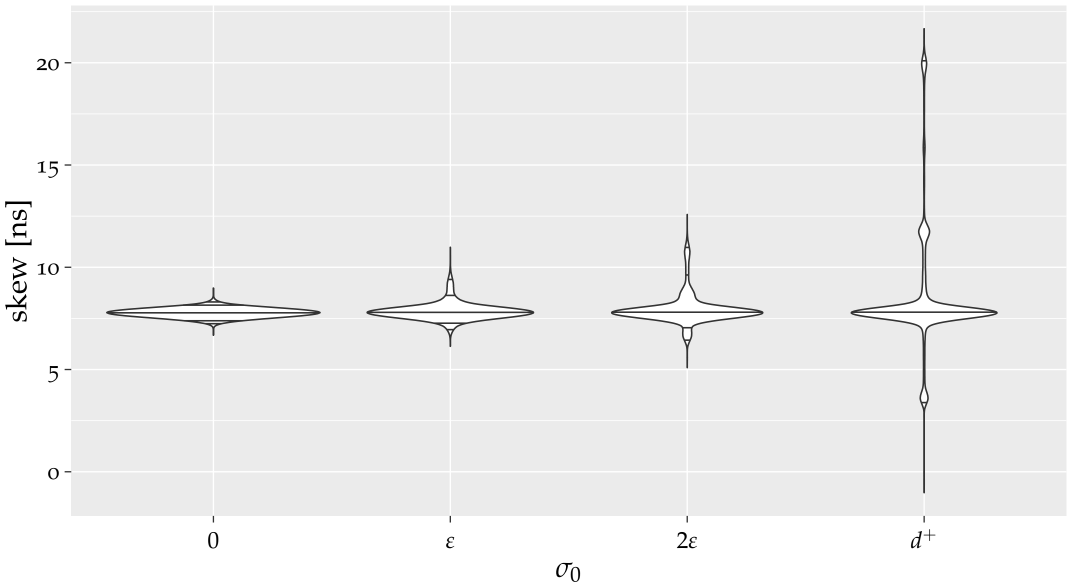
(b) inter-layer skews

Figure A.72: Skews in $\mathcal{G}_C$ with $\mathcal{A}^{0,2f}_{\text{mid}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-10, 35]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



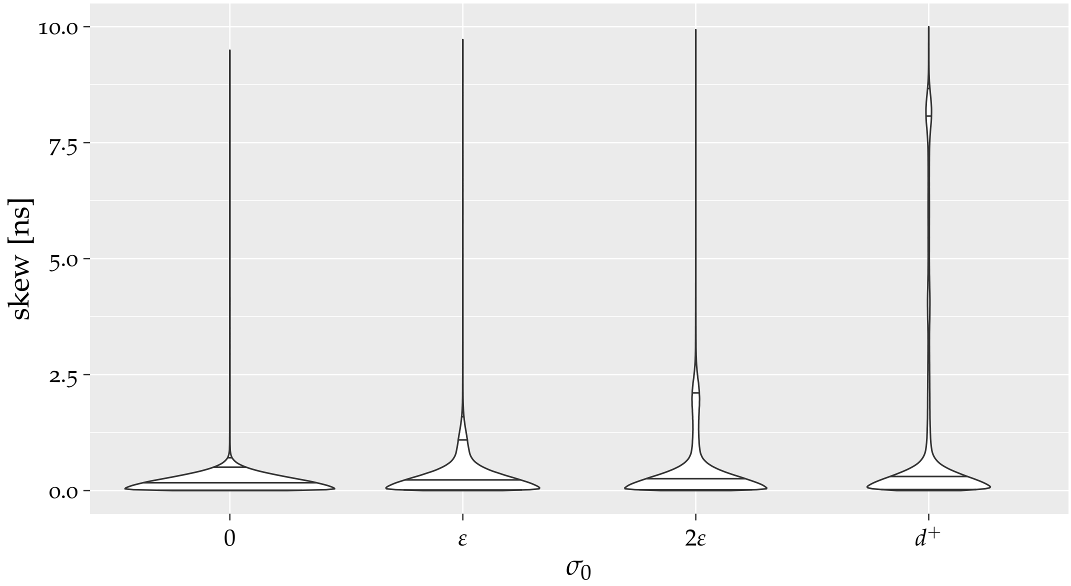
(a) intra-layer skews



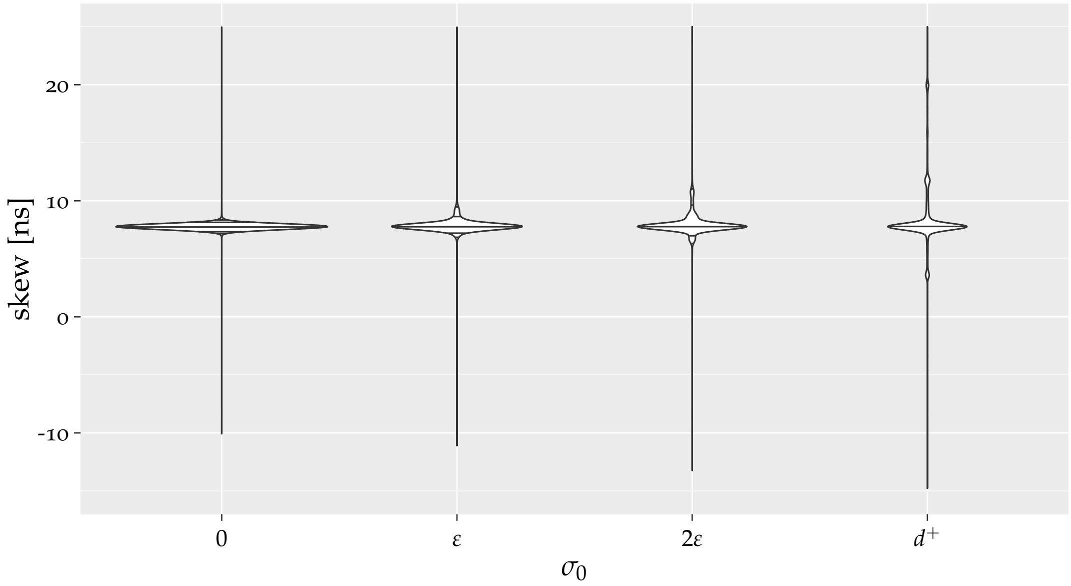
(b) inter-layer skews

Figure A.73: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{mid}}^{1,2f}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



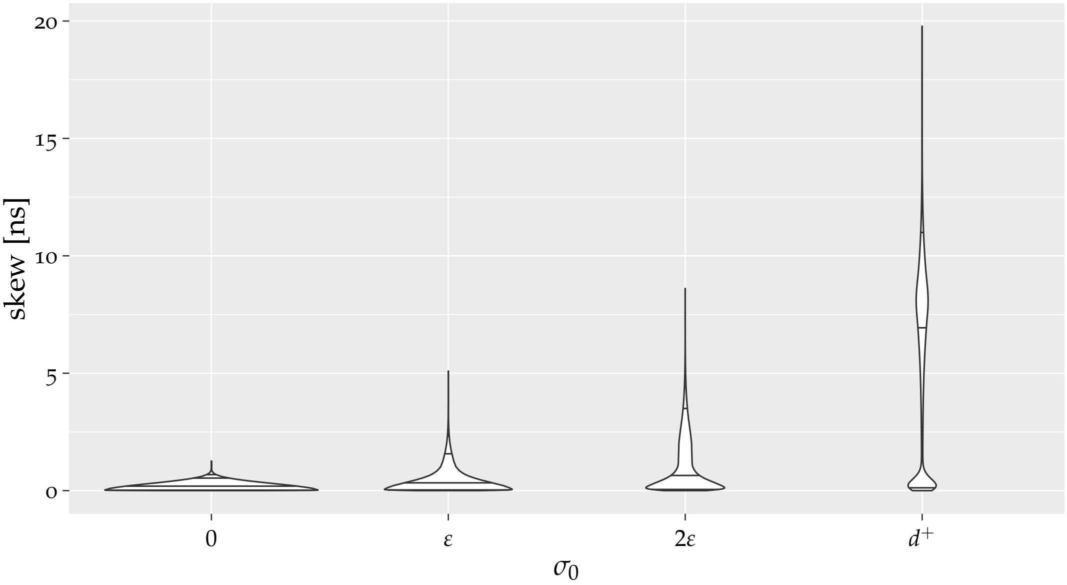
(a) intra-layer skews



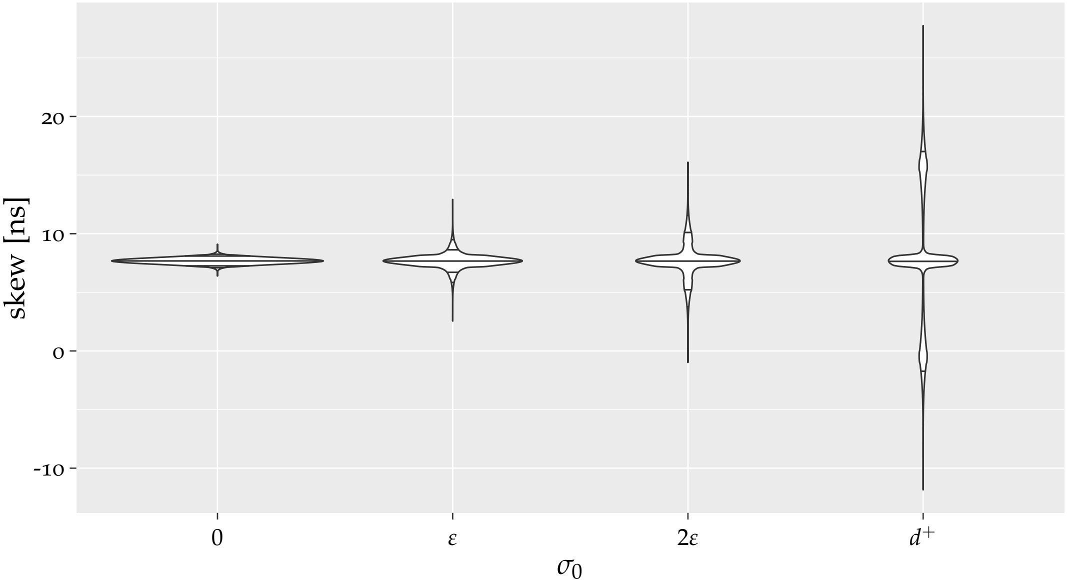
(b) inter-layer skews

Figure A.74: Skews in $\mathcal{G}_C$ with $\mathcal{A}^{1,2f}_{\text{mid}}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 10 ns, while the inter-layer skews have been limited to be within $[-15, 25]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



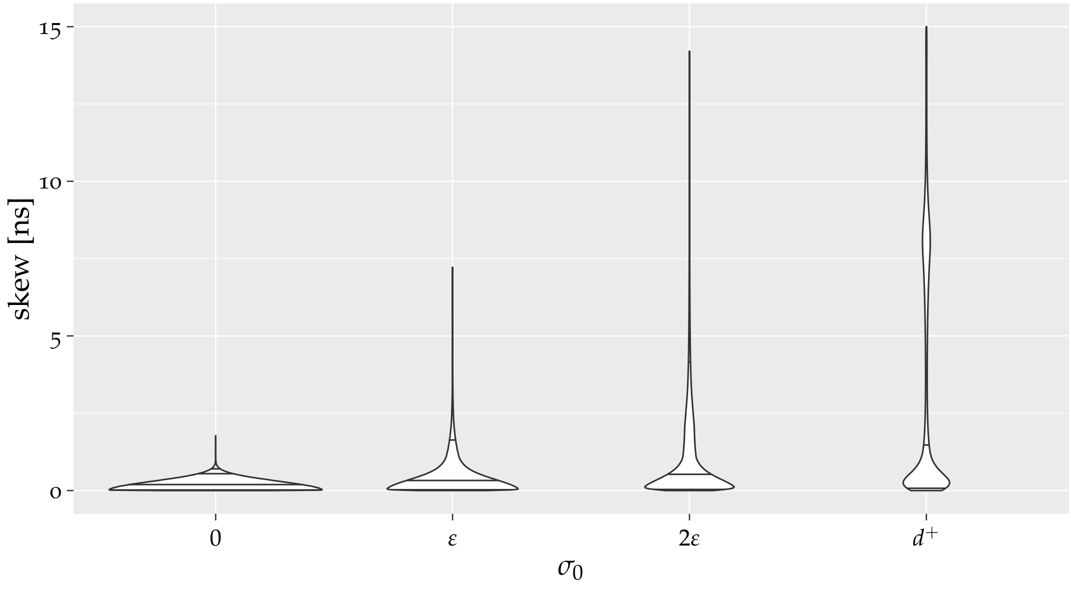
(a) intra-layer skews



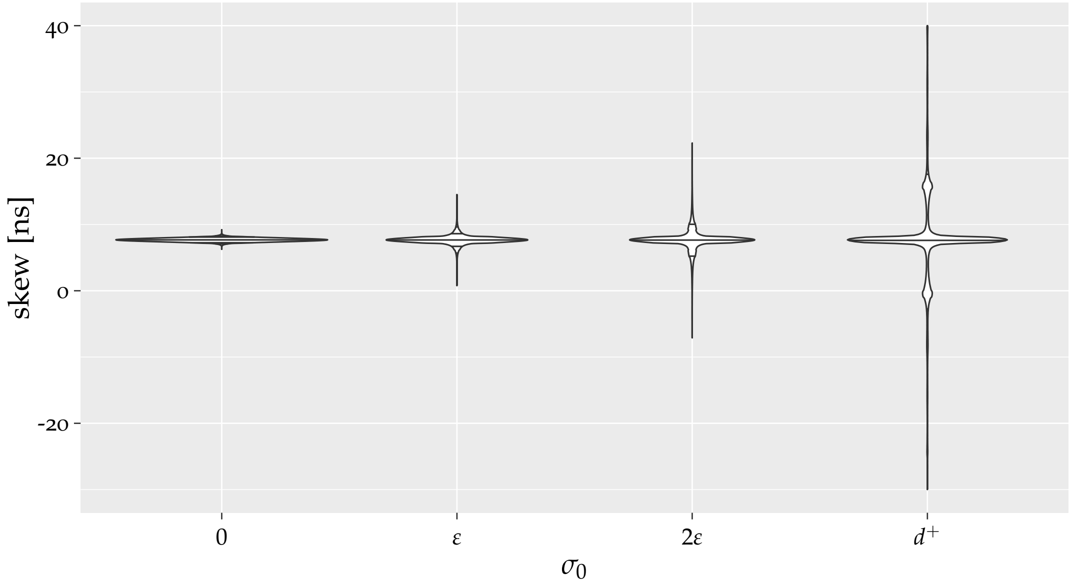
(b) inter-layer skews

Figure A.75: Skews in $\mathcal{G}_C$ with $\mathcal{A}^{2,2f}_{\text{mid}}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



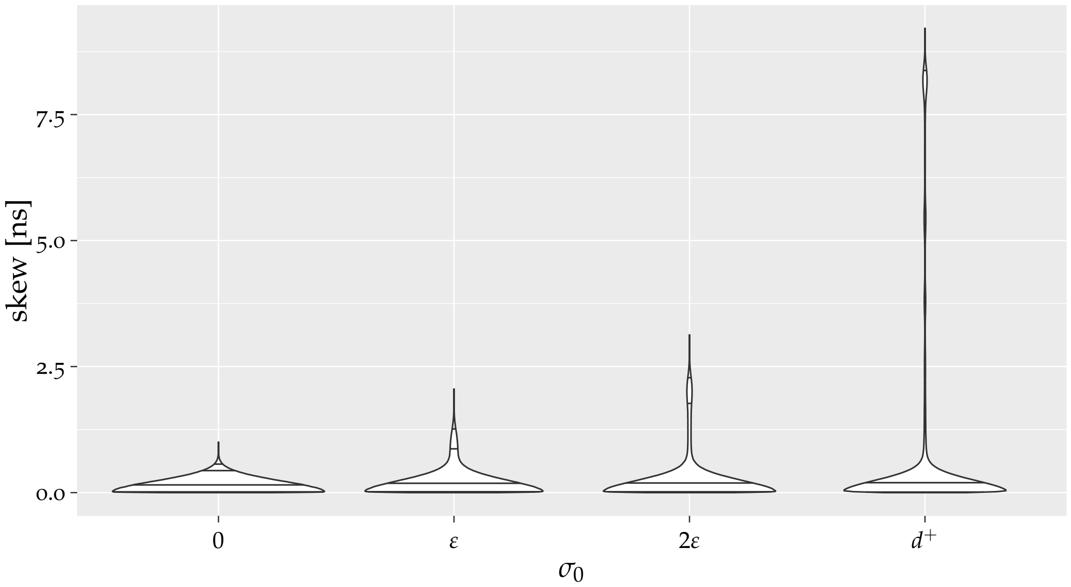
(a) intra-layer skews



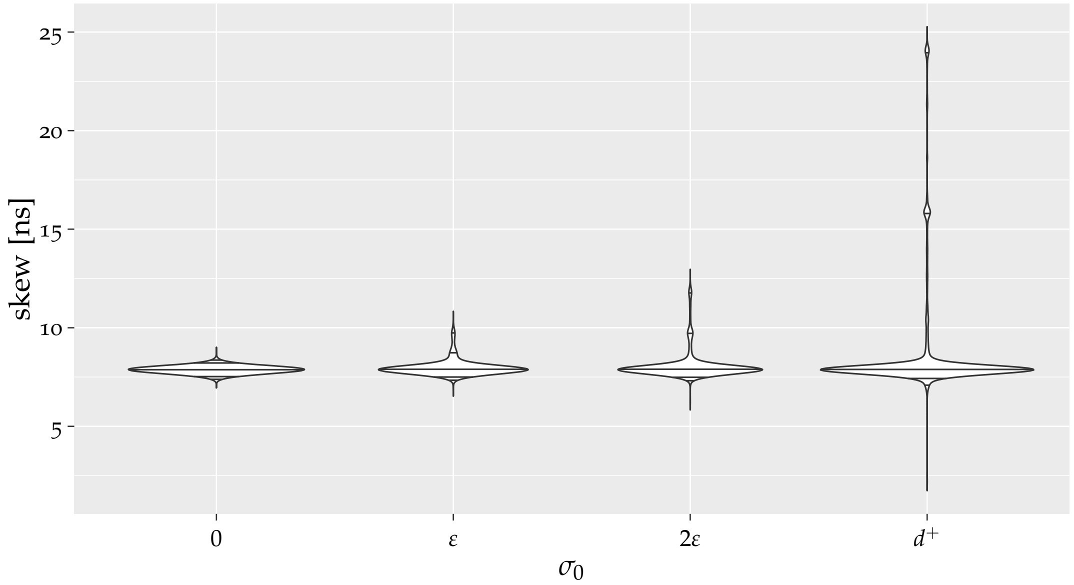
(b) inter-layer skews

Figure A.76: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{mid}}^{2,2f}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-30, 40]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



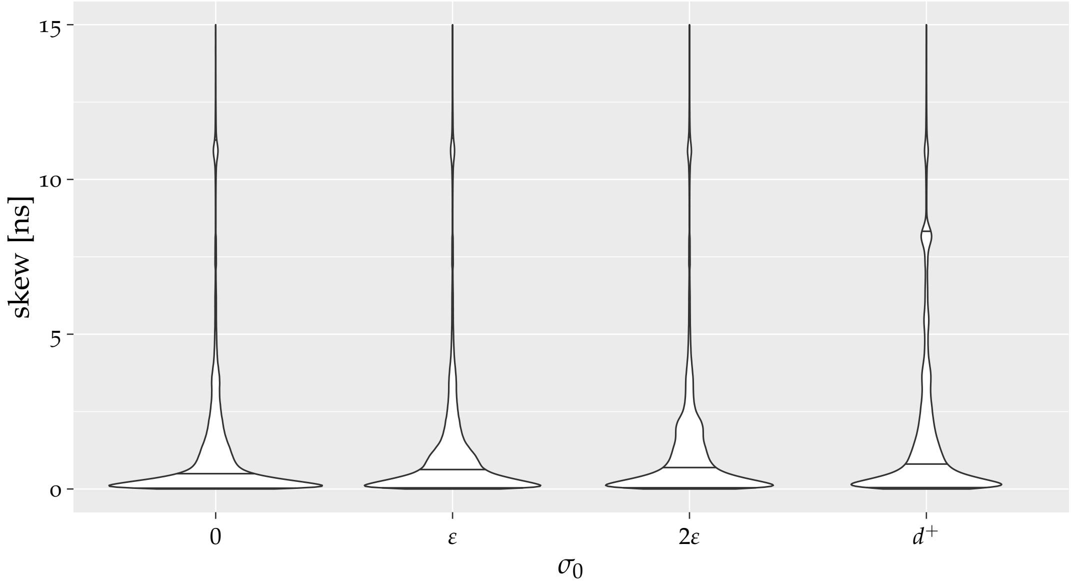
(a) intra-layer skews



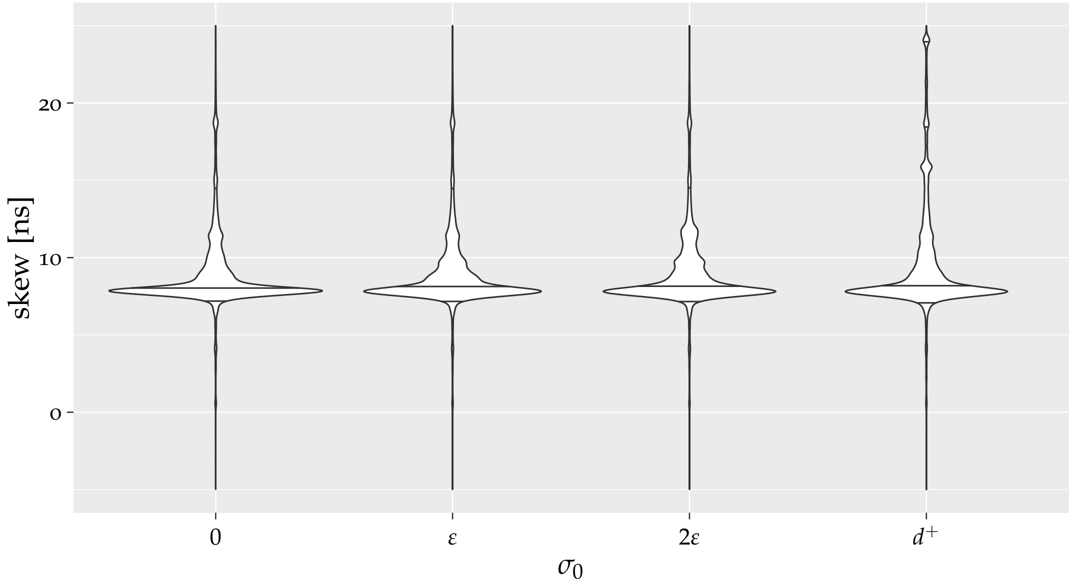
(b) inter-layer skews

Figure A.77: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{avg}}^{0,2f}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



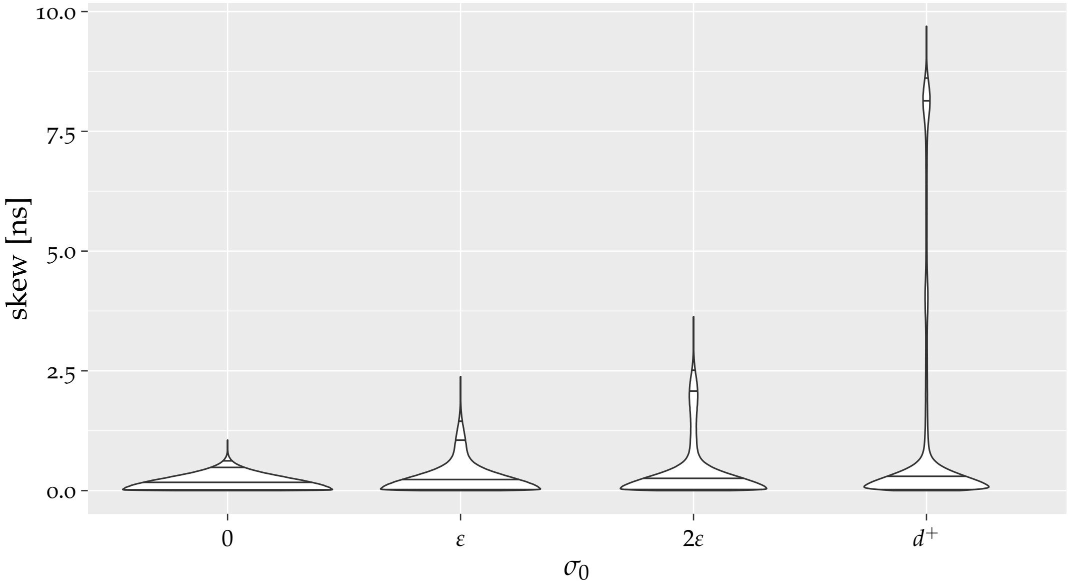
(a) intra-layer skews



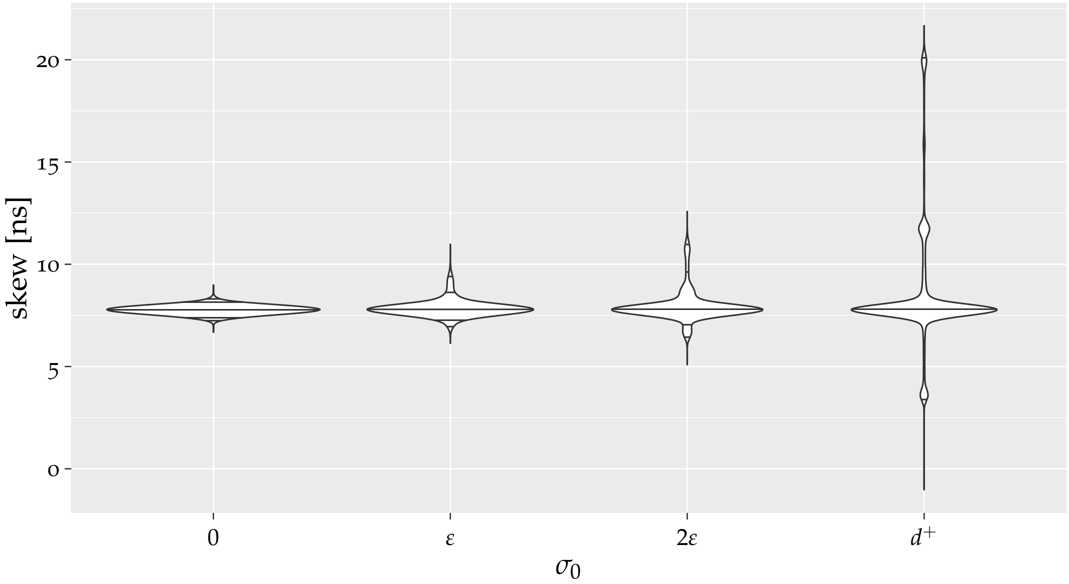
(b) inter-layer skews

Figure A.78: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{avg}}^{0,2f}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-5, 25]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



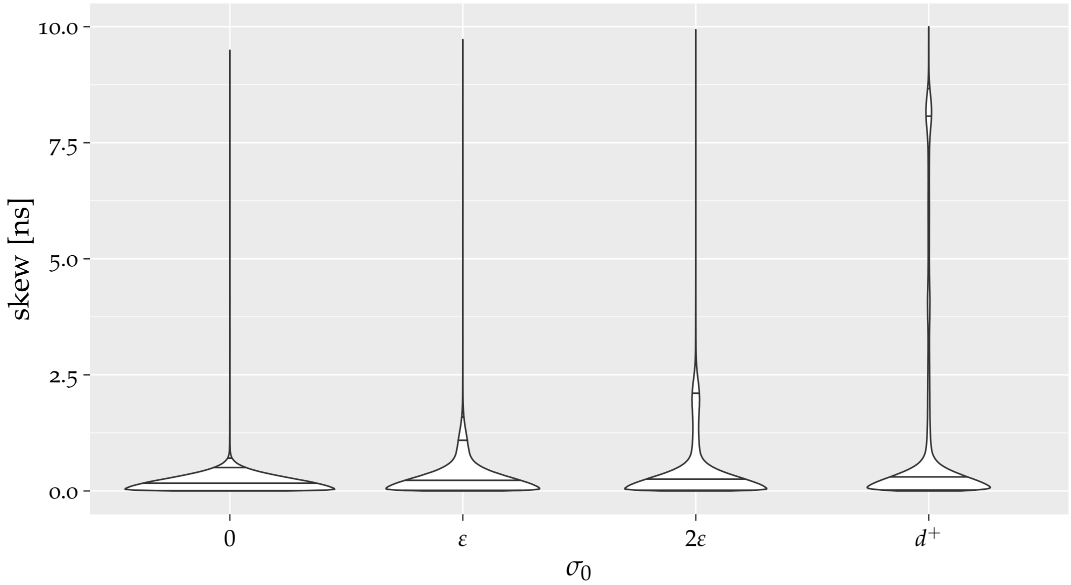
(a) intra-layer skews



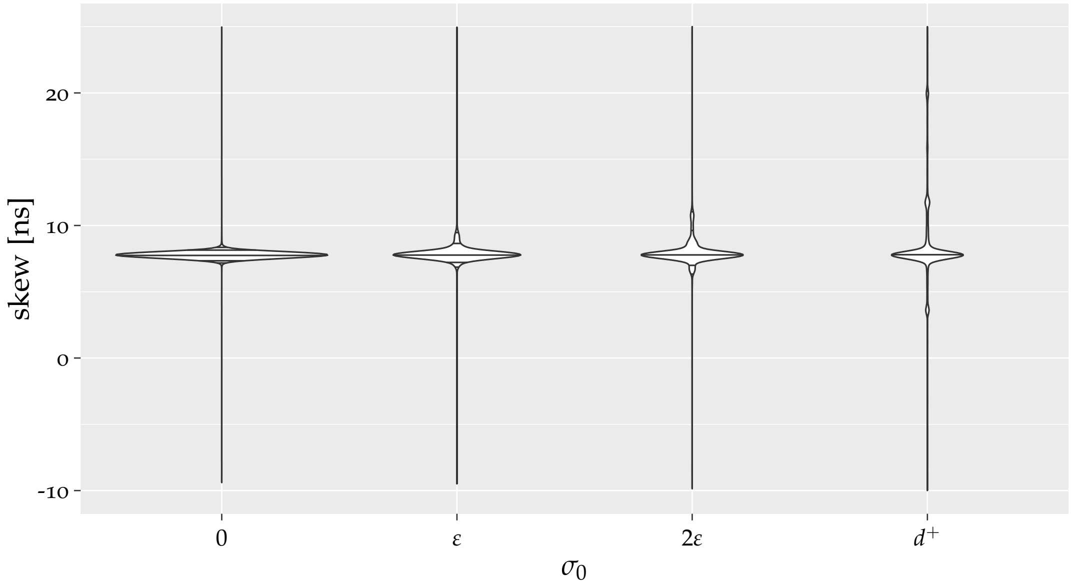
(b) inter-layer skews

Figure A.79: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{avg}}^{1,2f}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



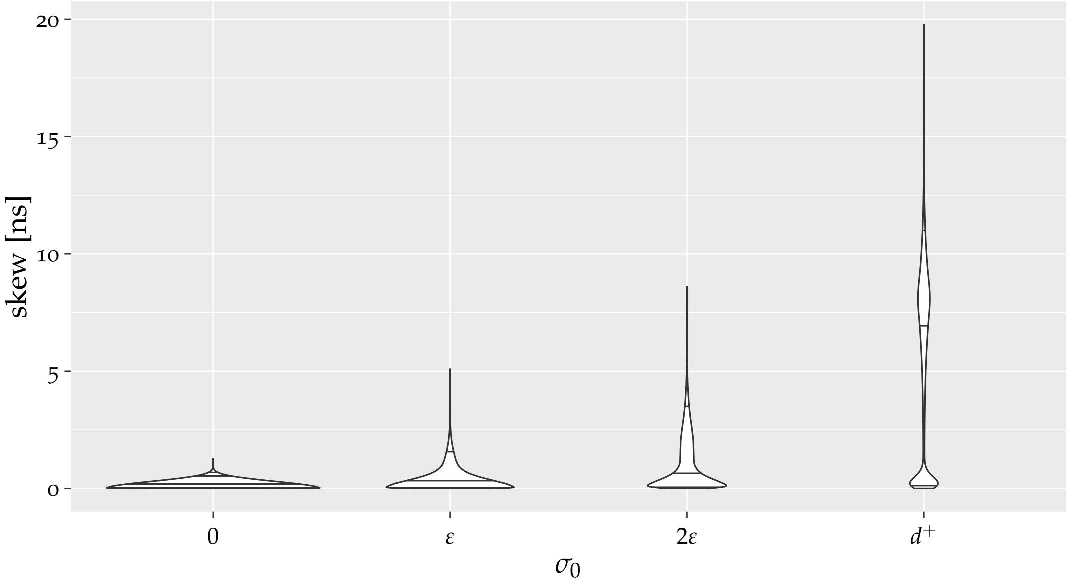
(a) intra-layer skews



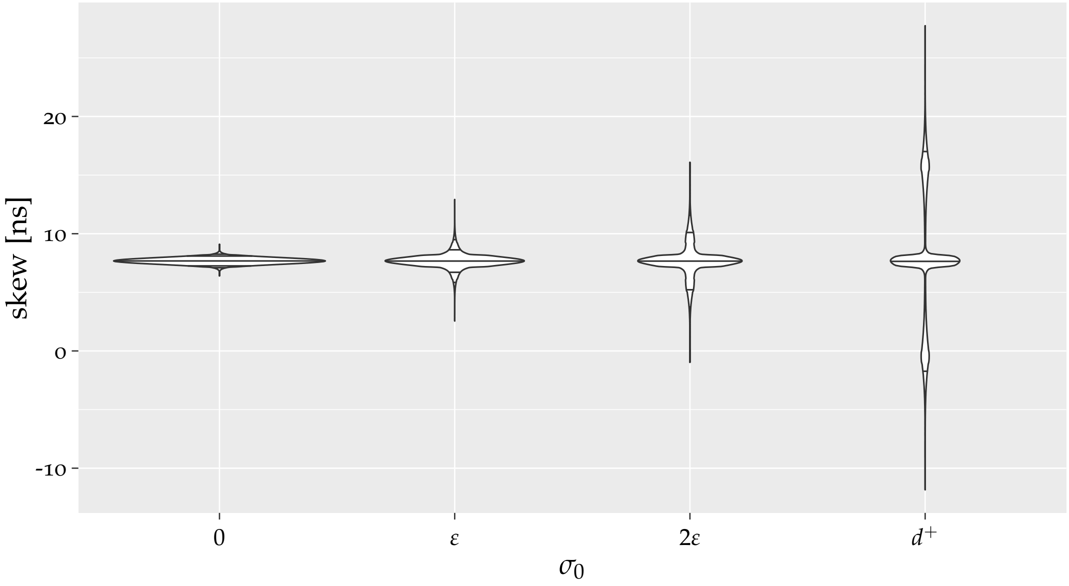
(b) inter-layer skews

Figure A.80: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{avg}}^{1,2f}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 10 ns, while the inter-layer skews have been limited to be within $[-10, 25]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skews

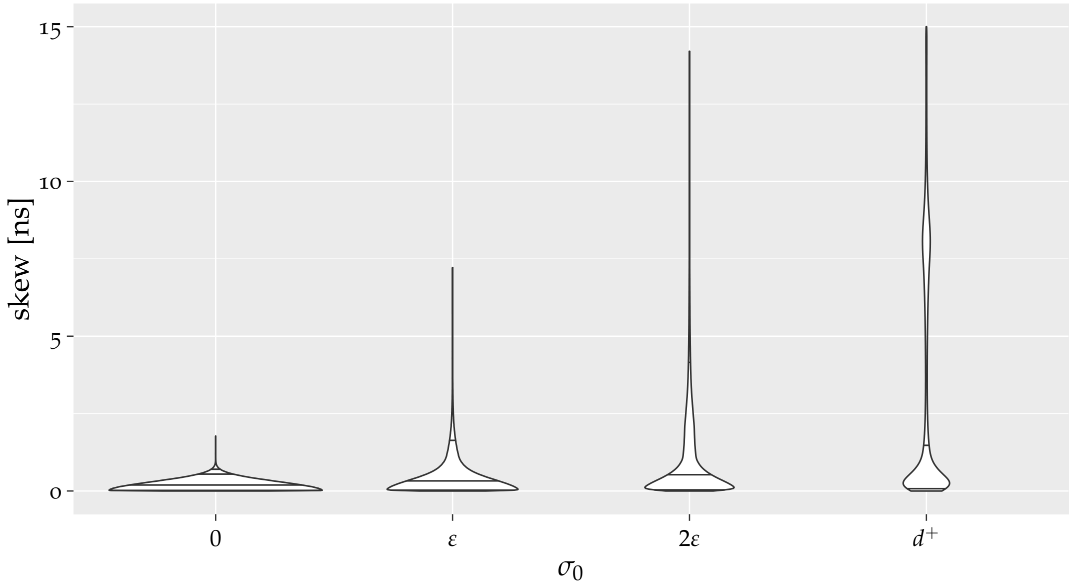


(b) inter-layer skews

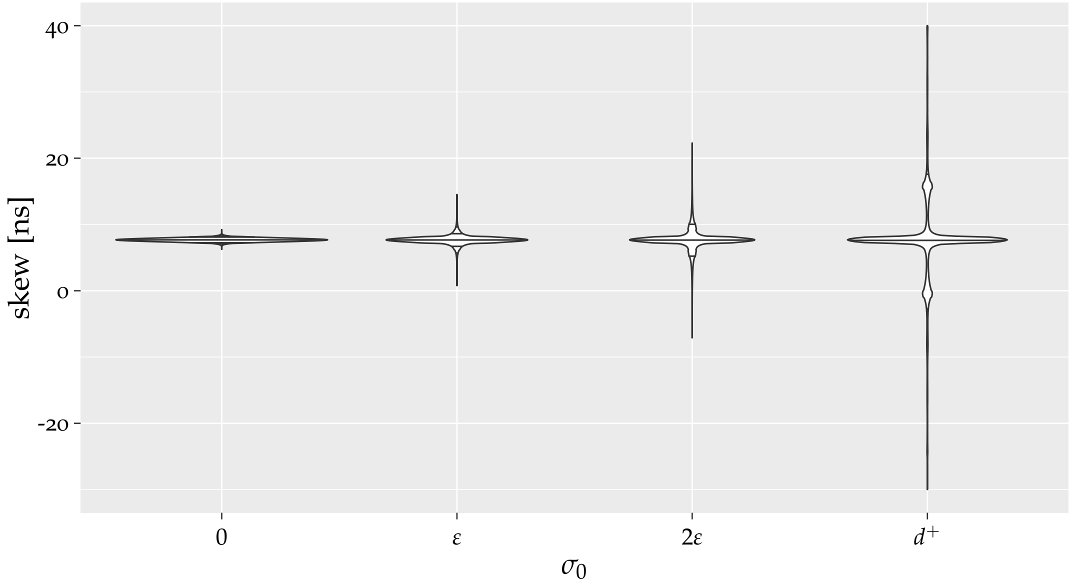
Figure A.81: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{avg}}^{2,2f}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



(a) intra-layer skews



(b) inter-layer skews

Figure A.82: Skews in $\mathcal{G}_C$ with $\mathcal{A}_{\text{avg}}^{2,2f}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 15 ns, while the inter-layer skews have been limited to be within $[-30, 40]$ ns.

A.14 Tables with the 2–local fault Model in $\mathcal{G}_C$

In this section we also present the tables of the algorithms $\mathcal{A}^{3,2f}_{\text{adj}}$, $\mathcal{A}^{3,2f}_{\text{any}}$, $\mathcal{A}^{0,2f}_{\text{avg},\,25\,\text{ns}}$, $\mathcal{A}^{0,2f}_{\text{mid},\,25\,\text{ns}}$, $\mathcal{A}^{1,2f}_{\text{avg},\,25\,\text{ns}}$, $\mathcal{A}^{1,2f}_{\text{mid},\,25\,\text{ns}}$, $\mathcal{A}^{1,2f}_{\text{avg},\,80\,\text{ns}}$, $\mathcal{A}^{1,2f}_{\text{mid},\,80\,\text{ns}}$, $\mathcal{A}^{2,2f}_{\text{avg},\,25\,\text{ns}}$, and $\mathcal{A}^{2,2f}_{\text{mid},\,25\,\text{ns}}$. We omit the common offset of 25 ns in the algorithm description. Further, recall that the provided inter-layer skews of the time-based algorithms have been compensated by their stated offset.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.11: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology under the 2–local fault model with pulse-based algorithms and input skews 0 and ε.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^{3,2f}_{\text{adj}}$	0	0	0	0.01	0.02	0.31	0.25	0.77	1.08	2.85	4.97	7.01	7.30	7.86	7.87	8.43	8.75	10.79
		15	0	0.01	0.03	1.03	0.33	6.09	7.43	15.10	−10.67	1.24	6.42	7.90	7.88	9.81	14.77	28.44
		15	1	0.01	0.03	0.97	0.32	5.76	7.19	15.10	−7.80	1.50	6.61	7.86	7.88	9.04	14.11	22.52
$\mathcal{A}^{3,2f}_{\text{any}}$	0	0	0	0.00	0.02	0.22	0.19	0.54	0.68	1.26	6.40	7.10	7.26	7.68	7.68	8.10	8.26	9.10
		15	0	0.00	0.02	0.23	0.19	0.55	0.70	1.77	6.27	7.08	7.25	7.68	7.68	8.11	8.27	9.20
		15	1	0.00	0.02	0.22	0.19	0.54	0.69	1.30	6.36	7.09	7.26	7.68	7.68	8.10	8.26	9.07
$\mathcal{A}^{3,2f}_{\text{adj}}$	ε	0	0	0.01	0.04	0.72	0.46	2.24	3.32	8.15	−0.47	5.14	6.30	7.80	7.84	9.15	10.33	15.74
		15	0	0.01	0.04	1.25	0.51	6.10	7.50	22.01	−16.45	1.12	5.49	7.86	7.86	10.41	14.79	36.80
		15	1	0.01	0.04	1.19	0.50	5.76	7.31	15.21	−7.77	1.36	5.63	7.82	7.86	9.92	14.21	20.73
$\mathcal{A}^{3,2f}_{\text{any}}$	ε	0	0	0.01	0.03	0.49	0.33	1.58	2.32	4.81	3.10	5.85	6.72	7.68	7.68	8.64	9.51	12.60
		15	0	0.01	0.03	0.50	0.33	1.64	2.67	7.34	0.89	5.69	6.71	7.67	7.68	8.63	9.64	14.51
		15	1	0.01	0.03	0.49	0.33	1.60	2.57	6.56	1.29	5.78	6.74	7.68	7.68	8.61	9.57	14.17

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.12: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology under the 2–local fault model with pulse-based algorithms and input skews 2ε and d^+.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^{3,2f}_{\text{adj}}$	2ε	0	0	0.01	0.06	1.50	1.10	4.17	5.63	8.20	−0.95	3.06	4.73	7.76	7.82	10.68	12.36	16.02
		15	0	0.01	0.05	1.76	0.89	6.66	7.65	28.43	−20.64	0.68	4.15	7.83	7.84	11.69	14.97	37.93
		15	1	0.01	0.05	1.69	0.87	6.40	7.48	20.55	−13.78	0.85	4.31	7.79	7.84	11.19	14.61	22.12
$\mathcal{A}^{3,2f}_{\text{any}}$	2ε	0	0	0.01	0.04	1.13	0.63	3.48	4.70	8.17	−0.73	3.80	5.23	7.68	7.68	10.12	11.57	15.87
		15	0	0.01	0.04	1.13	0.53	4.14	6.46	14.61	−7.05	2.65	5.23	7.67	7.67	10.08	12.67	19.93
		15	1	0.01	0.04	1.11	0.53	4.00	6.30	8.19	−1.02	2.87	5.30	7.68	7.68	10.04	12.45	16.07
$\mathcal{A}^{3,2f}_{\text{adj}}$	d^+	0	0	0.03	0.14	5.68	6.94	7.87	8.10	8.22	−8.64	−0.87	−0.17	7.58	7.70	15.12	15.50	16.39
		15	0	0.02	0.11	4.95	6.45	7.92	8.49	57.07	−49.91	−3.86	−0.20	7.54	7.73	15.12	15.63	60.56
		15	1	0.02	0.11	4.89	6.43	7.88	8.14	57.07	−49.91	−2.10	−0.16	7.53	7.73	15.07	15.50	45.62
$\mathcal{A}^{3,2f}_{\text{any}}$	d^+	0	0	0.02	0.10	5.03	6.58	7.83	8.08	8.20	−8.64	−0.82	−0.09	7.54	7.61	15.06	15.47	16.39
		15	0	0.01	0.07	4.05	4.98	7.83	8.13	40.31	−33.92	−3.17	−0.08	7.37	7.62	14.95	15.46	38.53
		15	1	0.02	0.08	4.05	5.02	7.80	8.09	32.00	−26.54	−1.50	−0.03	7.44	7.62	14.94	15.43	23.07

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.13: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology under the 2–local fault model with time-based algorithms with $d = 0$ and input skew 0 and ε. The inter-layer skews have been compensated by 25 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^{0,2f}_{\text{avg},25\,\text{ns}}$	0	0	0	0.00	0.01	0.18	0.15	0.43	0.57	1.06	6.97	7.40	7.54	7.88	7.88	8.22	8.37	8.96
		15	0	0.01	0.04	1.84	0.52	8.20	16.51	69.50	−24.83	0.97	6.45	9.55	8.16	16.79	25.82	80.89
		15	1	0.01	0.03	1.46	0.48	6.03	14.89	25.14	−24.31	3.80	6.84	9.25	8.14	15.68	23.43	76.02
$\mathcal{A}^{0,2f}_{\text{mid},25\,\text{ns}}$	0	0	0	0.00	0.01	0.18	0.15	0.44	0.57	1.00	6.95	7.39	7.53	7.87	7.87	8.22	8.36	9.00
		15	0	0.01	0.03	1.38	0.48	5.95	11.35	57.90	−15.94	3.95	7.23	9.05	8.06	14.92	20.83	68.35
		15	1	0.01	0.03	1.10	0.45	3.90	10.51	18.73	−13.19	6.19	7.33	8.83	8.05	12.92	18.66	65.86
$\mathcal{A}^{0,2f}_{\text{avg},25\,\text{ns}}$	ε	0	0	0.00	0.02	0.26	0.18	0.85	1.36	2.58	6.29	7.31	7.51	7.97	7.91	8.75	9.74	11.22
		15	0	0.01	0.04	1.89	0.66	8.22	16.54	65.19	−24.83	1.06	6.47	9.61	8.35	16.85	25.87	74.98
		15	1	0.01	0.04	1.51	0.62	6.00	14.85	25.14	−18.76	3.83	6.84	9.31	8.33	15.66	23.39	72.66
$\mathcal{A}^{0,2f}_{\text{mid},25\,\text{ns}}$	ε	0	0	0.00	0.02	0.26	0.19	0.87	1.26	2.06	6.53	7.34	7.51	7.96	7.90	8.74	9.74	10.83
		15	0	0.01	0.04	1.43	0.62	5.91	11.42	55.72	−15.82	3.94	7.21	9.12	8.16	14.91	20.87	65.58
		15	1	0.01	0.04	1.15	0.58	3.93	10.45	19.28	−10.18	6.19	7.31	8.90	8.15	12.93	18.68	63.37

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.14: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology under the 2–local fault model with time-based algorithms with $d = 0$ and input skew 2ε and d^+. The inter-layer skews have been compensated by 25 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^{0,2f}_{\text{avg},25\,\text{ns}}$	2ε	0	0	0.00	0.02	0.36	0.19	1.71	2.39	3.80	5.63	7.24	7.49	8.09	7.91	9.71	11.76	13.36
		15	0	0.01	0.05	1.96	0.74	8.22	16.55	63.08	−24.83	1.19	6.45	9.69	8.41	16.91	26.01	72.77
		15	1	0.01	0.04	1.57	0.69	5.97	14.78	25.14	−13.93	3.83	6.81	9.39	8.37	15.66	23.37	68.63
$\mathcal{A}^{0,2f}_{\text{mid},25\,\text{ns}}$	2ε	0	0	0.00	0.02	0.36	0.19	1.77	2.28	3.13	5.84	7.32	7.50	8.08	7.90	9.72	11.77	12.96
		15	0	0.01	0.04	1.50	0.67	5.94	11.46	56.54	−16.18	3.94	7.20	9.20	8.18	14.94	21.04	64.45
		15	1	0.01	0.04	1.22	0.63	3.95	10.41	20.15	−8.57	6.16	7.30	8.99	8.16	12.99	18.69	58.80
$\mathcal{A}^{0,2f}_{\text{avg},25\,\text{ns}}$	d^+	0	0	0.00	0.02	0.94	0.21	7.16	8.46	9.79	3.07	6.81	7.44	8.78	7.92	15.76	23.93	25.72
		15	0	0.01	0.05	2.48	0.88	8.81	16.61	77.67	−25.40	1.21	6.31	10.31	8.49	21.49	27.08	87.66
		15	1	0.01	0.05	2.11	0.81	8.20	14.74	30.46	−11.96	3.68	6.67	10.04	8.45	18.72	24.36	70.18
$\mathcal{A}^{0,2f}_{\text{mid},25\,\text{ns}}$	d^+	0	0	0.00	0.02	0.92	0.20	7.46	8.38	9.21	1.73	7.11	7.46	8.75	7.91	15.82	23.97	25.27
		15	0	0.01	0.05	2.04	0.76	8.35	11.79	70.84	−25.38	3.72	7.12	9.84	8.21	18.83	24.29	84.97
		15	1	0.01	0.04	1.76	0.71	8.00	10.60	28.05	−8.57	5.89	7.24	9.64	8.19	17.27	24.08	62.47

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.15: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology under the 2–local fault model with time-based algorithms with $d = 1$ and input skew 0 and ε. The inter-layer skews have been compensated by 25 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^{1,2f}_{\text{avg},\,25\,\text{ns}}$	0	0	0	0.00	0.02	0.20	0.17	0.49	0.62	1.05	6.68	7.24	7.39	7.77	7.78	8.15	8.31	8.98
		15	0	0.00	0.02	0.22	0.18	0.51	0.72	20.54	−10.07	7.21	7.37	7.78	7.78	8.17	8.38	32.80
		15	1	0.00	0.02	0.21	0.17	0.50	0.69	11.16	−9.18	7.23	7.38	7.78	7.78	8.16	8.36	24.58
$\mathcal{A}^{1,2f}_{\text{mid},\,25\,\text{ns}}$	0	0	0	0.00	0.02	0.20	0.17	0.49	0.62	1.05	6.68	7.24	7.39	7.77	7.78	8.15	8.31	8.98
		15	0	0.00	0.02	0.22	0.18	0.51	0.72	20.54	−10.07	7.21	7.37	7.78	7.78	8.17	8.38	32.80
		15	1	0.00	0.02	0.21	0.17	0.50	0.69	11.16	−9.18	7.23	7.38	7.78	7.78	8.16	8.36	24.58
$\mathcal{A}^{1,2f}_{\text{avg},\,25\,\text{ns}}$	ε	0	0	0.00	0.02	0.33	0.23	1.05	1.45	2.38	6.13	6.96	7.27	7.84	7.80	8.63	9.41	10.97
		15	0	0.00	0.02	0.35	0.24	1.10	1.62	20.30	−11.10	6.90	7.25	7.85	7.80	8.68	9.50	32.98
		15	1	0.00	0.02	0.34	0.23	1.08	1.54	11.18	−9.09	6.93	7.26	7.85	7.80	8.66	9.46	25.15
$\mathcal{A}^{1,2f}_{\text{mid},\,25\,\text{ns}}$	ε	0	0	0.00	0.02	0.33	0.23	1.05	1.45	2.38	6.13	6.96	7.27	7.84	7.80	8.63	9.41	10.97
		15	0	0.00	0.02	0.35	0.24	1.10	1.62	20.30	−11.10	6.90	7.25	7.85	7.80	8.68	9.50	32.98
		15	1	0.00	0.02	0.34	0.23	1.08	1.54	11.18	−9.09	6.93	7.26	7.85	7.80	8.66	9.46	25.15

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.16: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology under the 2–local fault model with time-based algorithms with $d = 1$ and input skew 2ε and d^+. The inter-layer skews have been compensated by 25 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^{1,2f}_{\text{avg},\,25\,\text{ns}}$	2ε	0	0	0.01	0.02	0.50	0.26	2.08	2.52	3.63	5.10	6.44	7.05	7.94	7.81	9.63	10.98	12.58
		15	0	0.01	0.02	0.53	0.26	2.12	2.77	21.70	−13.22	6.37	7.03	7.95	7.81	9.66	11.07	33.10
		15	1	0.01	0.02	0.51	0.26	2.09	2.63	13.39	−9.15	6.42	7.04	7.95	7.81	9.64	11.02	26.06
$\mathcal{A}^{1,2f}_{\text{mid},\,25\,\text{ns}}$	2ε	0	0	0.01	0.02	0.50	0.26	2.08	2.52	3.63	5.10	6.44	7.05	7.94	7.81	9.63	10.98	12.58
		15	0	0.01	0.02	0.53	0.26	2.12	2.77	21.70	−13.22	6.37	7.03	7.95	7.81	9.66	11.07	33.10
		15	1	0.01	0.02	0.51	0.26	2.09	2.63	13.39	−9.15	6.42	7.04	7.95	7.81	9.64	11.02	26.06
$\mathcal{A}^{1,2f}_{\text{avg},\,25\,\text{ns}}$	d^+	0	0	0.01	0.03	1.56	0.30	8.14	8.62	9.69	−1.02	3.41	4.98	8.57	7.83	15.86	20.12	21.66
		15	0	0.01	0.03	1.60	0.31	8.19	10.00	38.33	−29.97	3.28	5.07	8.58	7.83	15.79	20.25	53.59
		15	1	0.01	0.03	1.57	0.31	8.16	8.96	25.88	−10.25	3.35	5.14	8.58	7.83	15.75	20.19	40.25
$\mathcal{A}^{1,2f}_{\text{mid},\,25\,\text{ns}}$	d^+	0	0	0.01	0.03	1.56	0.30	8.14	8.62	9.69	−1.02	3.41	4.98	8.57	7.83	15.86	20.12	21.66
		15	0	0.01	0.03	1.60	0.31	8.19	10.00	38.33	−29.97	3.28	5.07	8.58	7.83	15.79	20.25	53.59
		15	1	0.01	0.03	1.57	0.31	8.16	8.96	25.88	−10.25	3.35	5.14	8.58	7.83	15.75	20.19	40.25

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.17: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology under the 2–local fault model with time-based algorithms with $d = 1$ and input skew 0 and ε. The inter-layer skews have been compensated by 80 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^{1,2f}_{\text{avg},\,80\,\text{ns}}$	0	0	0	0.00	0.02	0.20	0.17	0.49	0.62	1.05	6.68	7.24	7.39	7.77	7.78	8.15	8.31	8.98
		15	0	0.00	0.02	0.27	0.18	0.52	0.95	56.35	−42.08	7.20	7.37	7.79	7.78	8.18	8.44	74.05
		15	1	0.00	0.02	0.24	0.18	0.51	0.85	30.05	−36.68	7.21	7.38	7.79	7.78	8.17	8.41	52.08
$\mathcal{A}^{1,2f}_{\text{mid},\,80\,\text{ns}}$	0	0	0	0.00	0.02	0.20	0.17	0.49	0.62	1.05	6.68	7.24	7.39	7.77	7.78	8.15	8.31	8.98
		15	0	0.00	0.02	0.27	0.18	0.52	0.95	56.35	−42.08	7.20	7.37	7.79	7.78	8.18	8.44	74.05
		15	1	0.00	0.02	0.24	0.18	0.51	0.85	30.05	−36.68	7.21	7.38	7.79	7.78	8.17	8.41	52.08
$\mathcal{A}^{1,2f}_{\text{avg},\,80\,\text{ns}}$	ε	0	0	0.00	0.02	0.33	0.23	1.05	1.45	2.38	6.13	6.96	7.27	7.84	7.80	8.63	9.41	10.97
		15	0	0.00	0.02	0.39	0.24	1.14	1.88	55.82	−42.10	6.86	7.25	7.86	7.80	8.73	9.62	74.23
		15	1	0.00	0.02	0.36	0.24	1.12	1.72	29.55	−36.59	6.90	7.26	7.86	7.80	8.71	9.57	52.65
$\mathcal{A}^{1,2f}_{\text{mid},\,80\,\text{ns}}$	ε	0	0	0.00	0.02	0.33	0.23	1.05	1.45	2.38	6.13	6.96	7.27	7.84	7.80	8.63	9.41	10.97
		15	0	0.00	0.02	0.39	0.24	1.14	1.88	55.82	−42.10	6.86	7.25	7.86	7.80	8.73	9.62	74.23
		15	1	0.00	0.02	0.36	0.24	1.12	1.72	29.55	−36.59	6.90	7.26	7.86	7.80	8.71	9.57	52.65

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.18: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology under the 2–local fault model with time-based algorithms with $d = 1$ and input skew 2ε and d^+. The inter-layer skews have been compensated by 80 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^{1,2f}_{\text{avg},80\,\text{ns}}$	2ε	0	0	0.01	0.02	0.50	0.26	2.08	2.52	3.63	5.10	6.44	7.05	7.94	7.81	9.63	10.98	12.58
		15	0	0.01	0.02	0.57	0.26	2.16	3.08	56.34	−42.33	6.33	7.01	7.96	7.81	9.76	11.16	74.35
		15	1	0.01	0.02	0.54	0.26	2.13	2.79	29.41	−36.65	6.38	7.02	7.96	7.81	9.74	11.09	53.56
$\mathcal{A}^{1,2f}_{\text{mid},80\,\text{ns}}$	2ε	0	0	0.01	0.02	0.50	0.26	2.08	2.52	3.63	5.10	6.44	7.05	7.94	7.81	9.63	10.98	12.58
		15	0	0.01	0.02	0.57	0.26	2.16	3.08	56.34	−42.33	6.33	7.01	7.96	7.81	9.76	11.16	74.35
		15	1	0.01	0.02	0.54	0.26	2.13	2.79	29.41	−36.65	6.38	7.02	7.96	7.81	9.74	11.09	53.56
$\mathcal{A}^{1,2f}_{\text{avg},80\,\text{ns}}$	d^+	0	0	0.01	0.03	1.56	0.30	8.14	8.62	9.70	−1.03	3.41	4.99	8.57	7.83	15.86	20.12	21.74
		15	0	0.01	0.03	1.64	0.31	8.20	10.20	64.13	−54.31	3.27	5.02	8.59	7.83	15.84	20.27	86.48
		15	1	0.01	0.03	1.59	0.31	8.17	9.03	41.30	−37.75	3.34	5.09	8.59	7.83	15.80	20.20	60.92
$\mathcal{A}^{1,2f}_{\text{mid},80\,\text{ns}}$	d^+	0	0	0.01	0.03	1.56	0.30	8.14	8.62	9.70	−1.03	3.41	4.99	8.57	7.83	15.86	20.12	21.74
		15	0	0.01	0.03	1.64	0.31	8.20	10.20	64.13	−54.31	3.27	5.02	8.59	7.83	15.84	20.27	86.48
		15	1	0.01	0.03	1.59	0.31	8.17	9.03	41.30	−37.75	3.34	5.09	8.59	7.83	15.80	20.20	60.92

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.19: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology under the 2–local fault model with time-based algorithms with $d = 2$ and input skew 0 and ε. The inter-layer skews have been compensated by 25 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^{2,2f}_{\text{avg},\,25\,\text{ns}}$	0	0	0	0.00	0.02	0.22	0.19	0.54	0.68	1.26	6.40	7.10	7.26	7.68	7.68	8.10	8.26	9.10
		15	0	0.00	0.02	0.23	0.19	0.55	0.70	1.77	6.27	7.08	7.25	7.68	7.68	8.11	8.27	9.20
		15	1	0.00	0.02	0.22	0.19	0.54	0.69	1.30	6.36	7.09	7.26	7.68	7.68	8.10	8.26	9.07
$\mathcal{A}^{2,2f}_{\text{mid},\,25\,\text{ns}}$	0	0	0	0.00	0.02	0.22	0.19	0.54	0.68	1.26	6.40	7.10	7.26	7.68	7.68	8.10	8.26	9.10
		15	0	0.00	0.02	0.23	0.19	0.55	0.70	1.77	6.27	7.08	7.25	7.68	7.68	8.11	8.27	9.20
		15	1	0.00	0.02	0.22	0.19	0.54	0.69	1.30	6.36	7.09	7.26	7.68	7.68	8.10	8.26	9.07
$\mathcal{A}^{2,2f}_{\text{avg},\,25\,\text{ns}}$	ε	0	0	0.01	0.03	0.49	0.33	1.57	2.31	5.09	2.56	5.86	6.72	7.68	7.68	8.64	9.50	12.91
		15	0	0.01	0.03	0.50	0.33	1.64	2.67	7.22	0.79	5.70	6.71	7.67	7.68	8.63	9.64	14.51
		15	1	0.01	0.03	0.49	0.33	1.60	2.57	6.33	1.22	5.78	6.74	7.68	7.68	8.61	9.57	13.86
$\mathcal{A}^{2,2f}_{\text{mid},\,25\,\text{ns}}$	ε	0	0	0.01	0.03	0.49	0.33	1.57	2.31	5.09	2.56	5.86	6.72	7.68	7.68	8.64	9.50	12.91
		15	0	0.01	0.03	0.50	0.33	1.64	2.67	7.22	0.79	5.70	6.71	7.67	7.68	8.63	9.64	14.51
		15	1	0.01	0.03	0.49	0.33	1.60	2.57	6.33	1.22	5.78	6.74	7.68	7.68	8.61	9.57	13.86

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

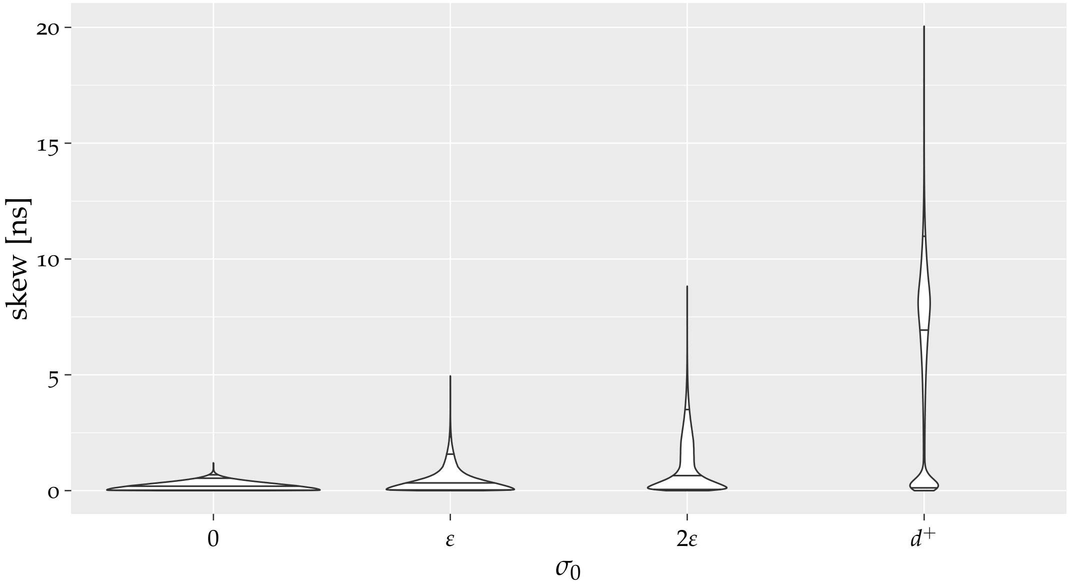
Table A.20: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_C$ topology under the 2–local fault model with time-based algorithms with $d = 2$ and input skew 2ε and d^+. The inter-layer skews have been compensated by 25 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^{2,2f}_{\text{avg},\,25\,\text{ns}}$	2ε	0	0	0.01	0.04	1.13	0.62	3.49	4.73	8.60	−0.97	3.77	5.24	7.68	7.68	10.12	11.59	16.08
		15	0	0.01	0.04	1.13	0.52	4.13	6.95	15.36	−7.09	2.57	5.25	7.67	7.67	10.06	12.75	22.28
		15	1	0.01	0.04	1.11	0.52	4.00	6.79	14.21	−6.44	2.81	5.32	7.68	7.68	10.02	12.51	21.39
$\mathcal{A}^{2,2f}_{\text{mid},\,25\,\text{ns}}$	2ε	0	0	0.01	0.04	1.13	0.62	3.49	4.73	8.60	−0.97	3.77	5.24	7.68	7.68	10.12	11.59	16.08
		15	0	0.01	0.04	1.13	0.52	4.13	6.95	15.36	−7.09	2.57	5.25	7.67	7.67	10.06	12.75	22.28
		15	1	0.01	0.04	1.11	0.52	4.00	6.79	14.21	−6.44	2.81	5.32	7.68	7.68	10.02	12.51	21.39
$\mathcal{A}^{2,2f}_{\text{avg},\,25\,\text{ns}}$	d^+	0	0	0.02	0.09	5.66	6.79	10.95	12.89	19.78	−11.83	−3.90	−1.69	7.68	7.68	17.05	19.26	27.73
		15	0	0.01	0.06	5.48	1.87	21.83	31.54	61.43	−55.13	−18.04	−2.68	7.63	7.67	17.78	33.20	65.53
		15	1	0.01	0.06	5.41	1.95	21.22	31.33	33.20	−48.94	−17.31	−2.18	7.67	7.68	17.46	32.54	41.21
$\mathcal{A}^{2,2f}_{\text{mid},\,25\,\text{ns}}$	d^+	0	0	0.02	0.09	5.66	6.79	10.95	12.89	19.78	−11.83	−3.90	−1.69	7.68	7.68	17.05	19.26	27.73
		15	0	0.01	0.06	5.48	1.87	21.83	31.54	61.43	−55.13	−18.04	−2.68	7.63	7.67	17.78	33.20	65.53
		15	1	0.01	0.06	5.41	1.95	21.22	31.33	33.20	−48.94	−17.31	−2.18	7.67	7.68	17.46	32.54	41.21

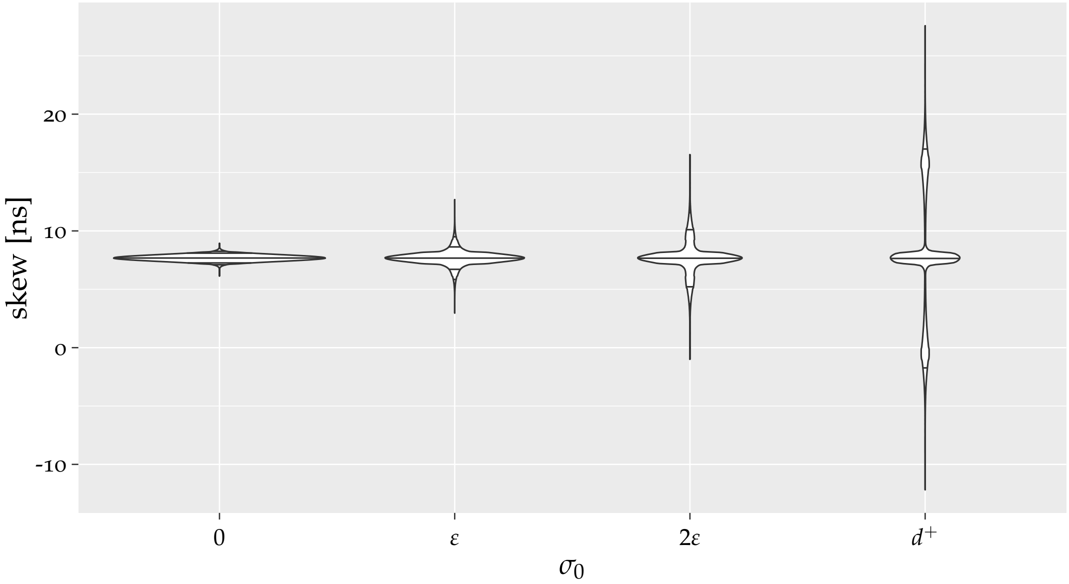
A.15 Plots with the 2–local fault Model in $\mathcal{G}_D$

In this section we also present the plots of the algorithms $\mathcal{A}^{3,2f}_{\text{any}}$, $\mathcal{A}^{0,2f}_{\text{avg},\,25\,\text{ns}}$ and $\mathcal{A}^{0,2f}_{\text{mid},\,25\,\text{ns}}$. We omit the common offset of 25 ns in the algorithm description. Further, recall that the provided inter-layer skews have been compensated by 25 ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



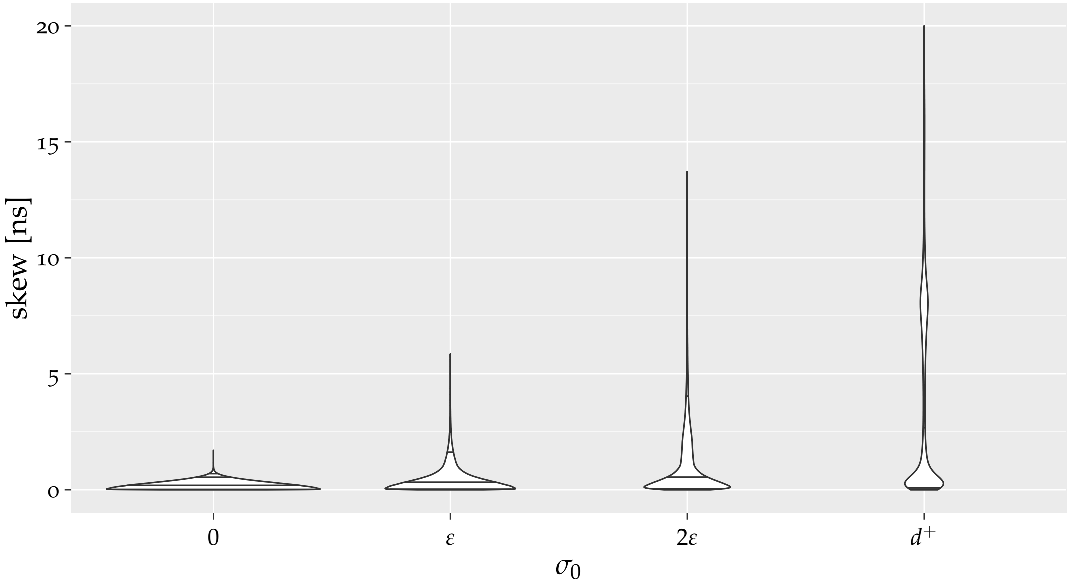
(a) intra-layer skews



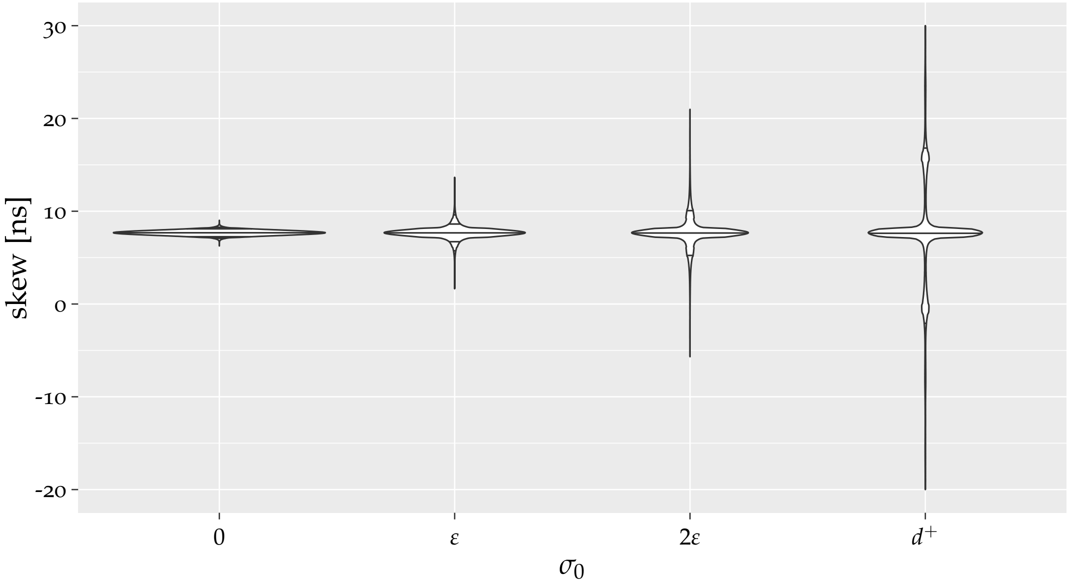
(b) inter-layer skews

Figure A.83: Skews in $\mathcal{G}_D$ with $\mathcal{A}_{\text{any}}^{3,2f}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



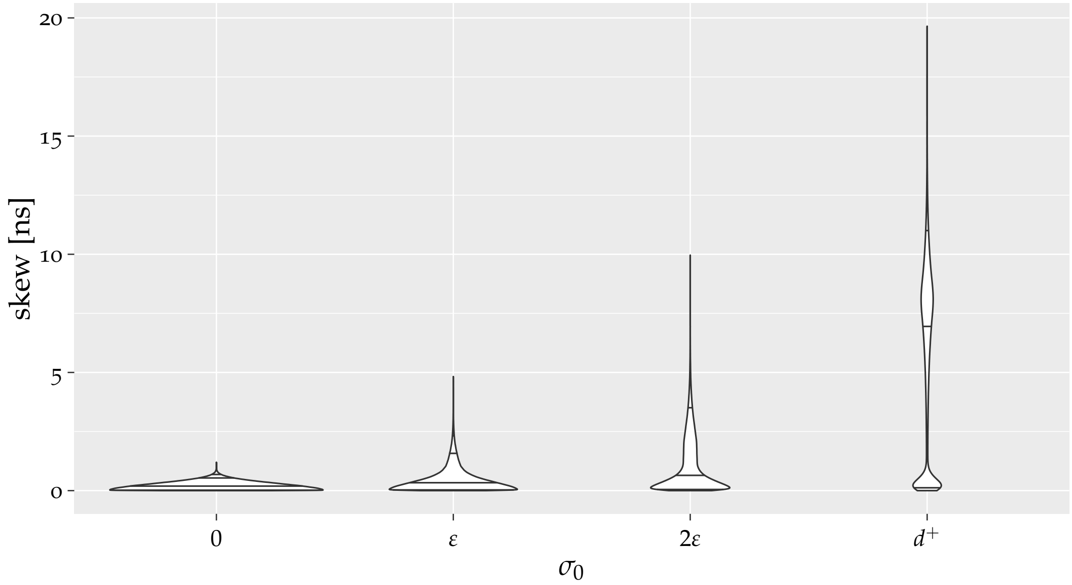
(a) intra-layer skews



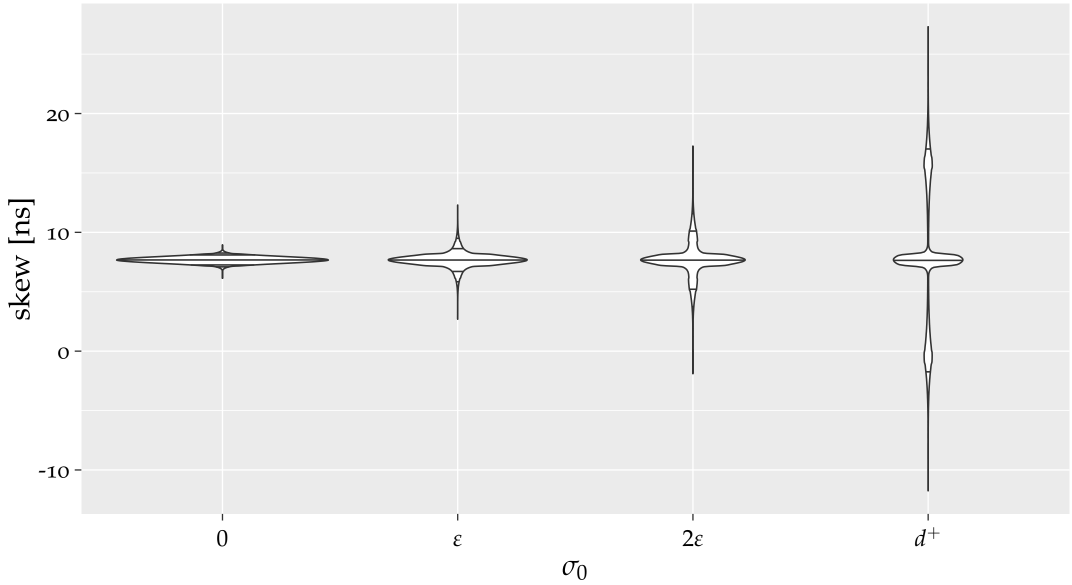
(b) inter-layer skews

Figure A.84: Skews in $\mathcal{G}_D$ with $\mathcal{A}_{\text{any}}^{3,2f}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-20, 30]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



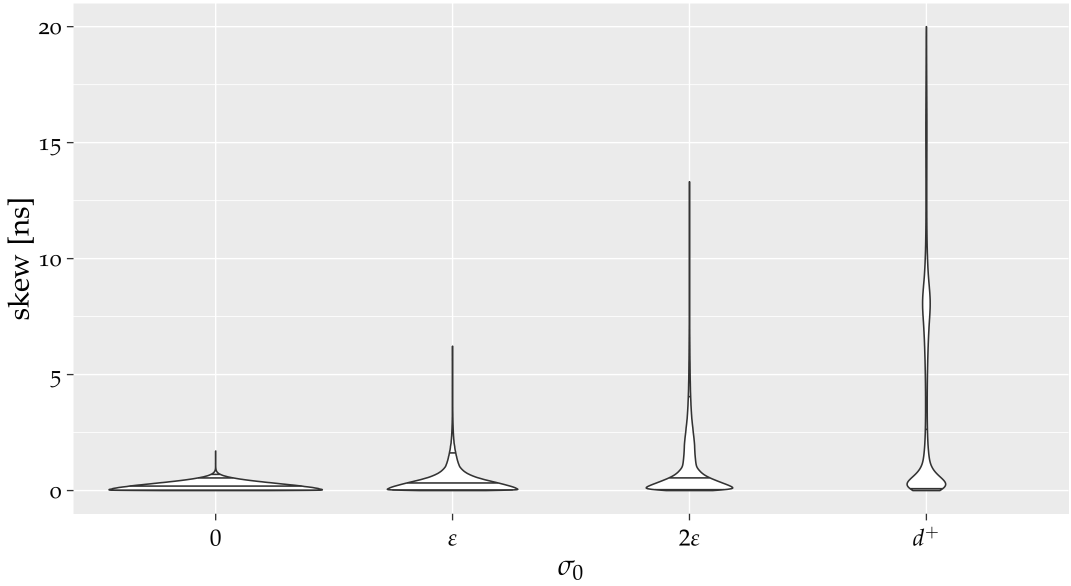
(a) intra-layer skews



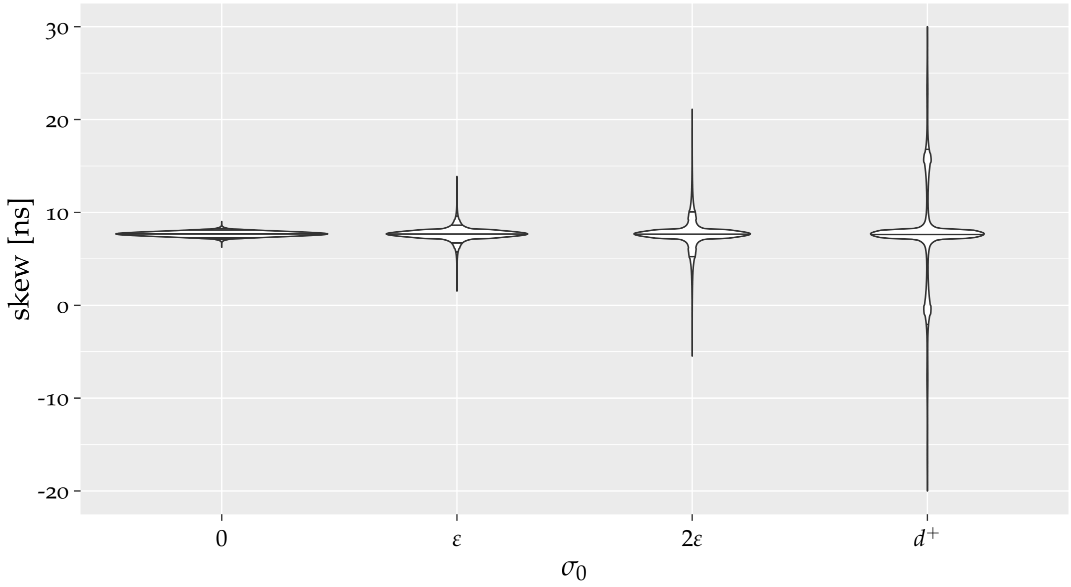
(b) inter-layer skews

Figure A.85: Skews in $\mathcal{G}_D$ with $\mathcal{A}_{\text{mid}}^{0,2f}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



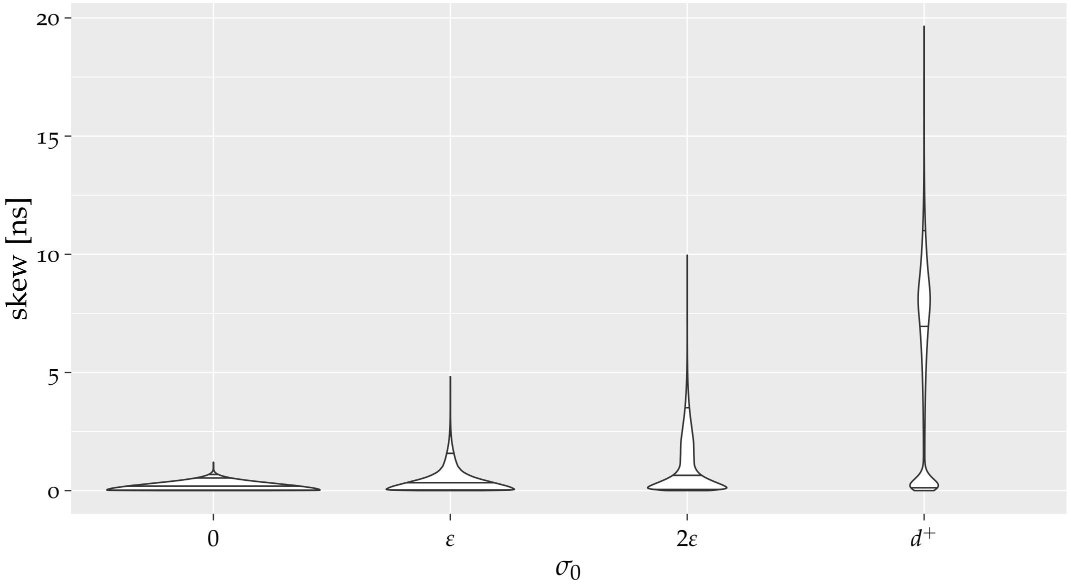
(a) intra-layer skews



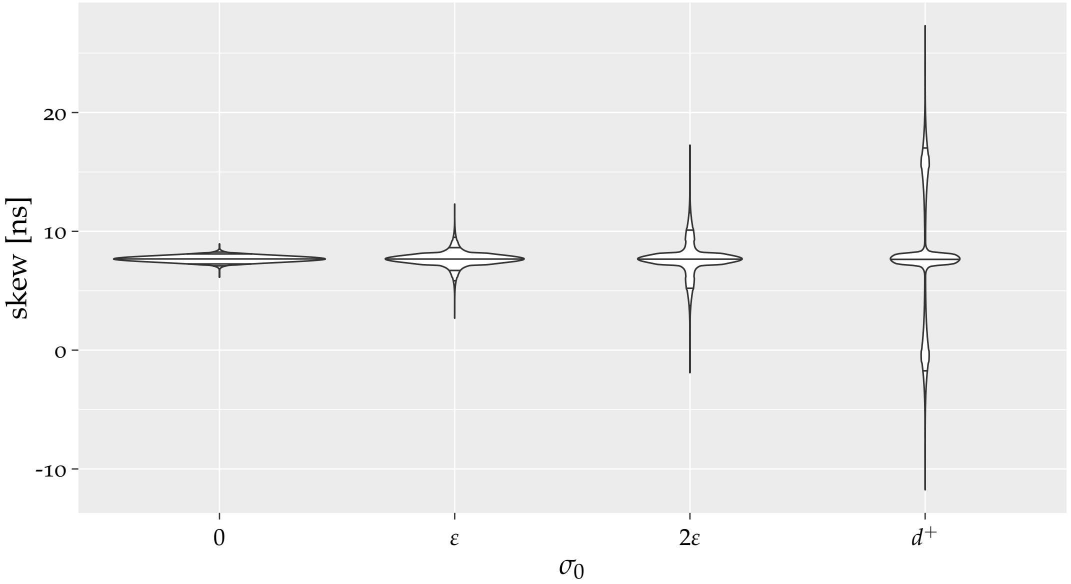
(b) inter-layer skews

Figure A.86: Skews in $\mathcal{G}_D$ with $\mathcal{A}_{\text{mid}}^{0,2f}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-20, 30]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



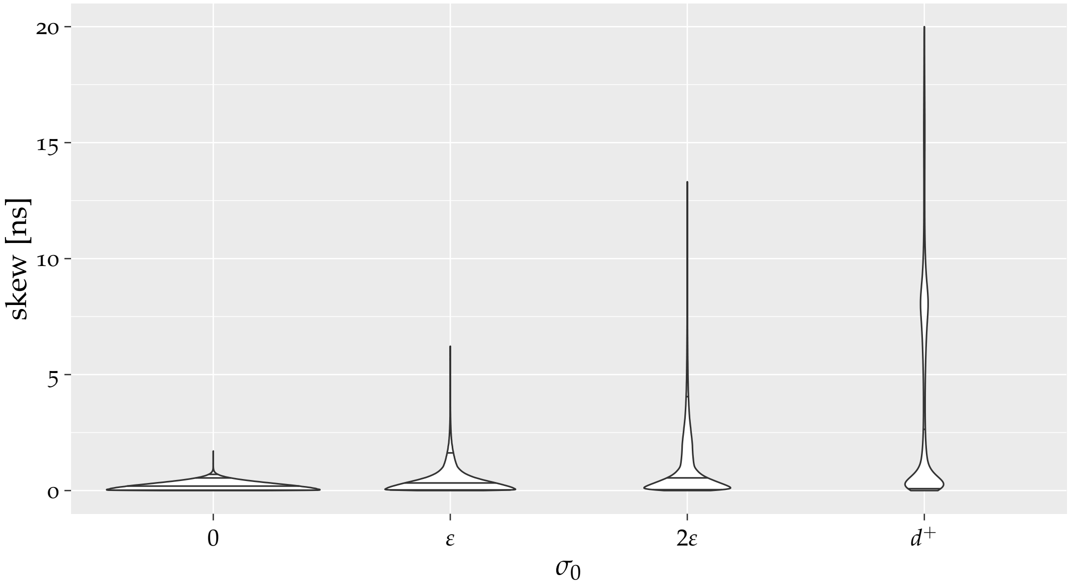
(a) intra-layer skews



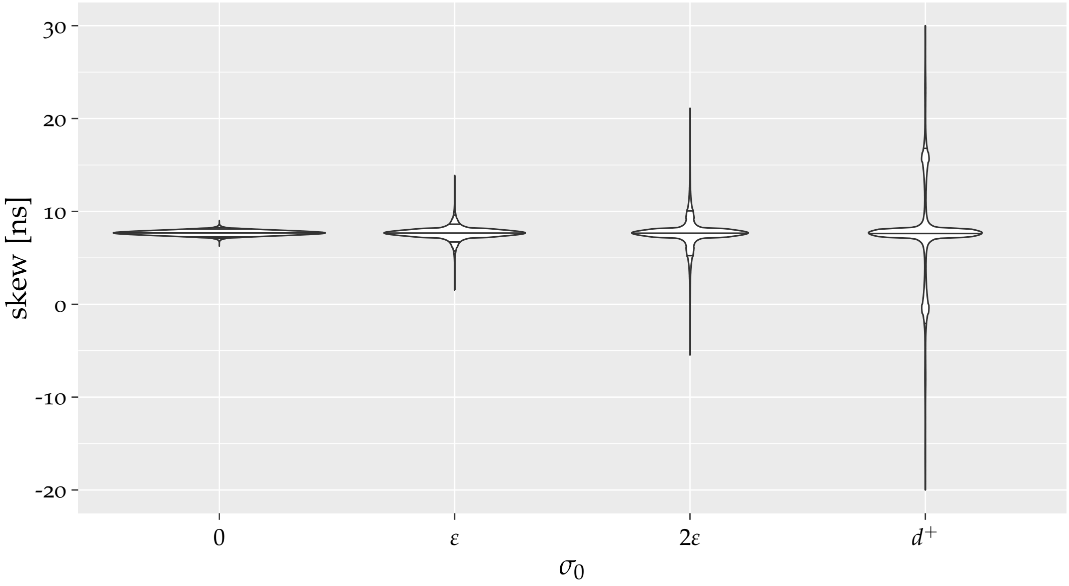
(b) inter-layer skews

Figure A.87: Skews in $\mathcal{G}_D$ with $\mathcal{A}_{\text{avg}}^{0,2f}$ without faults.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



(a) intra-layer skews



(b) inter-layer skews

Figure A.88: Skews in $\mathcal{G}_D$ with $\mathcal{A}_{\text{avg}}^{0,2f}$ with 15 Byzantine faults. For better visibility of the skew distribution, the intra-layer skews have been cropped at 20 ns, while the inter-layer skews have been limited to be within $[-20, 30]$ ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

A.16 Tables with the 2–local fault Model in $\mathcal{G}_D$

In this section we also present the tables of the algorithms $\mathcal{A}^{3,2f}_{\text{any}}$, $\mathcal{A}^{0,2f}_{\text{avg},\,25\,\text{ns}}$ and $\mathcal{A}^{0,2f}_{\text{mid},\,25\,\text{ns}}$. We omit the common offset of 25 ns in the algorithm description. Further, recall that the provided inter-layer skews have been compensated by 25 ns.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.21: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\mathrm{op}}$ (in [ns]) in the $\mathcal{G}_D$ topology under the 2–local fault model with pulse-based algorithms. The inter-layer skews of the time-based algorithms have been compensated by 25 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}_{\mathrm{any}}^{3,2f}$	0	0	0	0.00	0.02	0.22	0.19	0.54	0.68	1.19	6.14	7.10	7.26	7.68	7.68	8.10	8.26	8.94
		15	0	0.00	0.02	0.23	0.19	0.54	0.70	1.70	6.27	7.09	7.25	7.68	7.68	8.11	8.27	9.00
		15	1	0.00	0.02	0.22	0.19	0.54	0.69	1.30	6.41	7.10	7.26	7.68	7.68	8.10	8.26	9.00
$\mathcal{A}_{\mathrm{any}}^{3,2f}$	ε	0	0	0.01	0.03	0.49	0.33	1.58	2.32	4.94	2.96	5.85	6.72	7.68	7.68	8.64	9.51	12.68
		15	0	0.01	0.03	0.50	0.33	1.62	2.60	5.86	1.67	5.75	6.73	7.68	7.68	8.63	9.62	13.64
		15	1	0.01	0.03	0.49	0.33	1.59	2.52	5.86	1.67	5.80	6.74	7.68	7.68	8.62	9.56	13.64
$\mathcal{A}_{\mathrm{any}}^{3,2f}$	2ε	0	0	0.01	0.04	1.13	0.62	3.49	4.72	8.82	−1.00	3.78	5.23	7.68	7.68	10.13	11.58	16.53
		15	0	0.01	0.04	1.14	0.54	4.02	6.55	13.72	−5.67	2.84	5.26	7.68	7.68	10.10	12.51	20.96
		15	1	0.01	0.04	1.12	0.54	3.92	6.40	13.35	−5.56	3.01	5.30	7.68	7.68	10.05	12.35	20.57
$\mathcal{A}_{\mathrm{any}}^{3,2f}$	d^+	0	0	0.02	0.09	5.66	6.79	10.94	12.87	20.05	−12.17	−3.88	−1.70	7.68	7.68	17.05	19.24	27.57
		15	0	0.01	0.06	5.55	2.46	20.12	31.86	68.69	−61.19	−16.35	−2.53	7.68	7.68	17.87	31.70	76.65
		15	1	0.01	0.06	5.48	2.49	19.49	31.52	68.56	−60.48	−15.92	−2.22	7.68	7.68	17.57	31.28	76.65

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.22: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_D$ topology under the 2–local fault model with time-based algorithms with $d = 0$ and input skews 0 and ε. The inter-layer skews of the time-based algorithms have been compensated by 25 ns.

				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^{0,2f}_{\text{avg},25\,\text{ns}}$	0	0	0	0.00	0.02	0.22	0.19	0.54	0.68	1.19	6.14	7.10	7.26	7.68	7.68	8.10	8.26	8.94
		15	0	0.00	0.02	0.23	0.19	0.54	0.70	1.70	6.27	7.09	7.25	7.68	7.68	8.11	8.27	9.00
		15	1	0.00	0.02	0.22	0.19	0.54	0.69	1.30	6.41	7.10	7.26	7.68	7.68	8.10	8.26	9.00
$\mathcal{A}^{0,2f}_{\text{mid},25\,\text{ns}}$	0	0	0	0.00	0.02	0.22	0.19	0.54	0.68	1.19	6.14	7.10	7.26	7.68	7.68	8.10	8.26	8.94
		15	0	0.00	0.02	0.23	0.19	0.54	0.70	1.70	6.27	7.09	7.25	7.68	7.68	8.11	8.27	9.00
		15	1	0.00	0.02	0.22	0.19	0.54	0.69	1.30	6.41	7.10	7.26	7.68	7.68	8.10	8.26	9.00
$\mathcal{A}^{0,2f}_{\text{avg},25\,\text{ns}}$	ε	0	0	0.01	0.03	0.49	0.33	1.58	2.32	4.82	2.70	5.85	6.72	7.68	7.68	8.64	9.51	12.30
		15	0	0.01	0.03	0.50	0.33	1.62	2.60	6.22	1.55	5.75	6.72	7.68	7.68	8.63	9.61	13.88
		15	1	0.01	0.03	0.49	0.33	1.60	2.52	5.96	1.66	5.80	6.74	7.68	7.68	8.62	9.56	13.64
$\mathcal{A}^{0,2f}_{\text{mid},25\,\text{ns}}$	ε	0	0	0.01	0.03	0.49	0.33	1.58	2.32	4.82	2.70	5.85	6.72	7.68	7.68	8.64	9.51	12.30
		15	0	0.01	0.03	0.50	0.33	1.62	2.60	6.22	1.55	5.75	6.72	7.68	7.68	8.63	9.61	13.88
		15	1	0.01	0.03	0.49	0.33	1.60	2.52	5.96	1.66	5.80	6.74	7.68	7.68	8.62	9.56	13.64

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Table A.23: Intra- and inter-layer skews σ^{op} and $\hat{\sigma}^{\text{op}}$ (in [ns]) in the $\mathcal{G}_D$ topology under the 2–local fault model with time-based algorithms with $d = 0$ and input skews 2ε and d^+. The inter-layer skews of the time-based algorithms have been compensated by 25 ns.

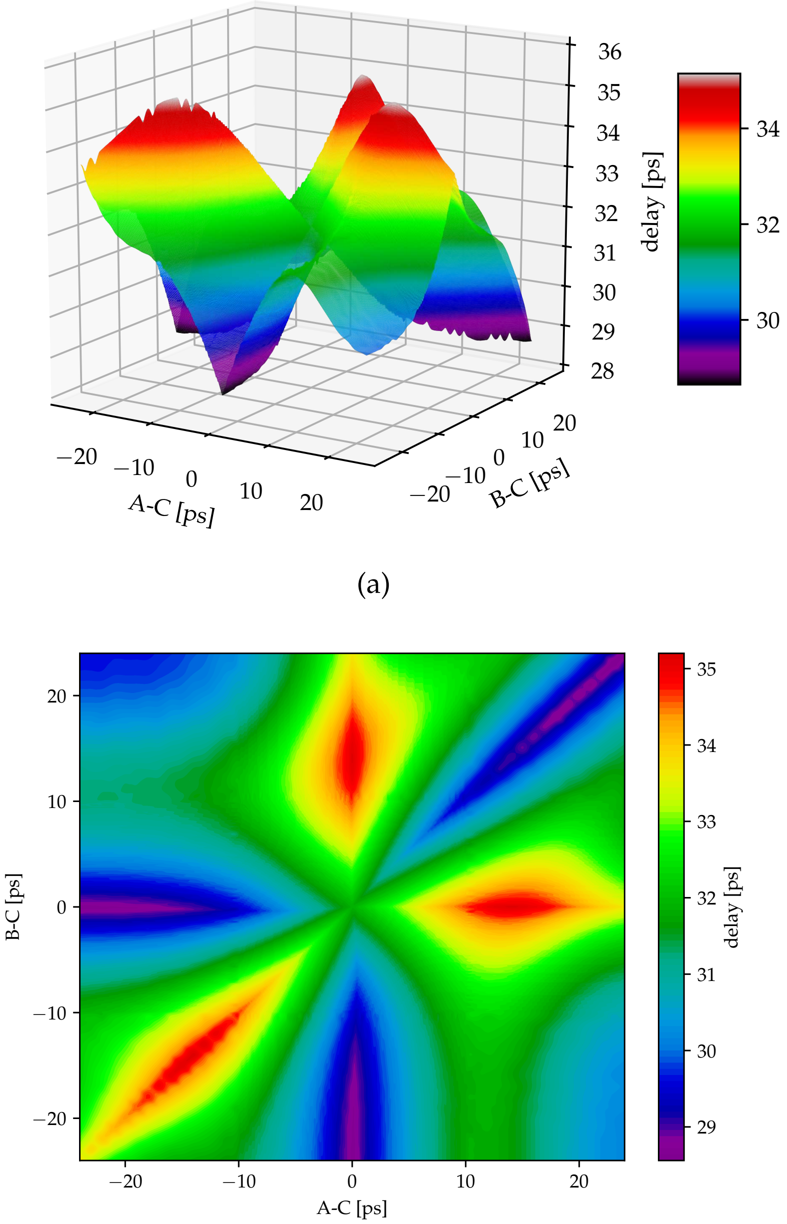
				intra-layer							inter-layer							
	σ_0	f	h	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max	min	q_1	q_5	avg	q_{50}	q_{95}	q_{99}	max
$\mathcal{A}^{0,2f}_{\text{avg},\,25\,\text{ns}}$	2ε	0	0	0.01	0.04	1.13	0.62	3.50	4.72	9.96	−1.88	3.77	5.23	7.68	7.68	10.13	11.59	17.25
		15	0	0.01	0.04	1.13	0.54	4.03	6.57	13.31	−5.45	2.82	5.27	7.68	7.68	10.09	12.54	21.10
		15	1	0.01	0.04	1.12	0.54	3.93	6.42	13.31	−5.30	2.98	5.31	7.68	7.68	10.05	12.37	20.94
$\mathcal{A}^{0,2f}_{\text{mid},\,25\,\text{ns}}$	2ε	0	0	0.01	0.04	1.13	0.62	3.50	4.72	9.96	−1.88	3.77	5.23	7.68	7.68	10.13	11.59	17.25
		15	0	0.01	0.04	1.13	0.54	4.03	6.57	13.31	−5.45	2.82	5.27	7.68	7.68	10.09	12.54	21.10
		15	1	0.01	0.04	1.12	0.54	3.93	6.42	13.31	−5.30	2.98	5.31	7.68	7.68	10.05	12.37	20.94
$\mathcal{A}^{0,2f}_{\text{avg},\,25\,\text{ns}}$	d^+	0	0	0.02	0.09	5.66	6.81	10.96	12.92	19.64	−11.75	−3.91	−1.70	7.68	7.68	17.05	19.27	27.30
		15	0	0.01	0.06	5.55	2.43	20.19	32.13	65.88	−58.01	−16.46	−2.51	7.68	7.68	17.86	31.81	73.65
		15	1	0.01	0.06	5.48	2.46	19.54	31.75	65.88	−57.95	−16.04	−2.20	7.68	7.68	17.55	31.40	73.65
$\mathcal{A}^{0,2f}_{\text{mid},\,25\,\text{ns}}$	d^+	0	0	0.02	0.09	5.66	6.81	10.96	12.92	19.64	−11.75	−3.91	−1.70	7.68	7.68	17.05	19.27	27.30
		15	0	0.01	0.06	5.55	2.43	20.19	32.13	65.88	−58.01	−16.46	−2.51	7.68	7.68	17.86	31.81	73.65
		15	1	0.01	0.06	5.48	2.46	19.54	31.75	65.88	−57.95	−16.04	−2.20	7.68	7.68	17.55	31.40	73.65

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

APPENDIX B

ADDITIONAL PLOTS OF CHAPTER 5

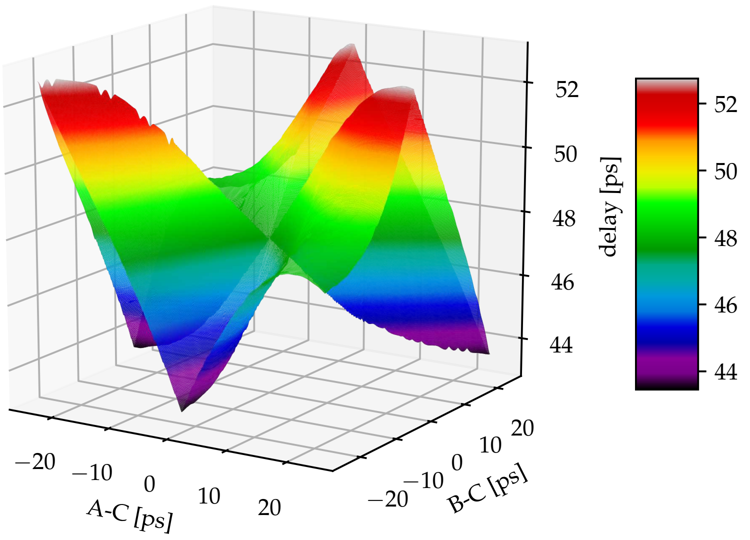


(a)

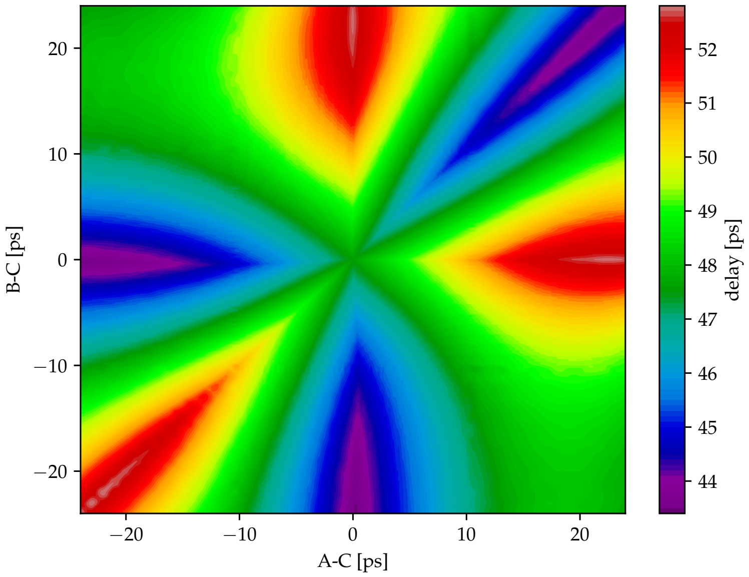
(b)

Figure B.1: Delay of the circuit from Figure 5.2 (Page 176), with the components in their fast parameter corner, as function of the relative times of the input transitions. It increases if two pulses together arrive early and the third one is late. It decreases if one pulse is early and the other two together arrive late.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub



(a)



(b)

Figure B.2: Delay of the circuit from Figure 5.2 (Page 176), with the components in their slow parameter corner, as function of the relative times of the input transitions. It increases if two pulses together arrive early and the third one is late. It decreases if one pulse is early and the other two together arrive late.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

TU Bibliothek Your knowledge hub WIEN
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

GLOSSARY

$[L]$ denotes the set $\{0, \ldots, L-1\}$. *Defined on pp.* 48, 114, 174

C_i is the slow clock of a clock domain i. A slow clock maps continuous real-time to discrete clock time. *Defined on p.* 194

d^- is the minimal time a trigger message is delayed between being sent on one and received at another node. *Defined on pp.* 48, 114

d^+ is the maximal time a trigger message is delayed between being sent on one and received at another node. *Defined on pp.* 48, 114

Δ_ℓ is the skew potential of layer ℓ, and is defined as $\Delta_\ell := \max_{i,j \in [W]}\{t_{\ell,i} - t_{\ell,j} - |i - j|_W d^-\}$. *Defined on p.* 53

ε is the maximal difference of the end-to-end delays, i.e., $d^+ - d^-$. *Defined on pp.* 48, 114

F_i^k is a fast clock of a clock domain i, started by tick k of the corresponding slow clock C_i. *Defined on p.* 195

Γ_i is the lower bound on the separation between two initial ticks of two fast clocks that belong to the same slow clock i. *Defined on p.* 196

h defines the outgoing h-hop neighborhood of a fault, which is discarded in the skew analysis. *Defined on pp.* 91, 131

IQD^p is for $0 < p < 1/2$ and a strictly-monotone continuous distribution function $F(x)$ defined as $q_{1-p} - q_p$, where $F(q_x) = x$. *Defined on p.* 84

I is the period interval. Every I^{th} slow clock tick, the address pointers are reset to their initial position. *Defined on p.* 198

L is the number of layers the grid has, which determine the height of the grid. *Defined on pp.* 48, 114, 174

$\boldsymbol{\lambda_0}$ defines the last layer were a path originating in some layer 0 node may deliver the trigger message not later than the fastest path originating in the same node but ending at layer ℓ. *Defined on p.* 56

$\boldsymbol{(\ell, i)}$ is the node located in layer ℓ at column i. *Defined on pp.* 48, 114, 174

$\boldsymbol{M}$ is minimum required size of the ring buffer. *Defined on p.* 202

$\boldsymbol{lzp^{i' \to (\ell,i)}}$ is a left zig-zag path. The path is composed of rightward and up-left links with the node (ℓ, i) as destination. The origin is either a node $(0, j)$ or (ℓ', i'). In the latter case, a left zig-zag path is called a triangular path. *Defined on p.* 51

$\boldsymbol{\Gamma_{\mathrm{min}}}$ is the minimal pulse separation time, i.e., the minimal time between the ticks of the same node, over the whole grid. *Defined on p.* 105

$\boldsymbol{\mathcal{S}}$ is the pulse separation time, i.e., the minimal time between the generation of consecutive pulses k and $k+1$ at layer 0. *Defined on p.* 68

$\boldsymbol{\sigma_\ell}$ is the intra-layer skew of layer ℓ, i.e., the skew between neighboring nodes in the same layer, assuming that the intra-layer links are symmetric. $\sigma_\ell := \max_{i \in [W]} \left\{ |t_{\ell,i} - t_{\ell,j}| \mid ((\ell, j), (\ell, i)) \in E \right\}$. *Defined on pp.* 53, 114, 175

$\boldsymbol{\hat{\sigma}_\ell}$ is the inter-layer skew of layer $\ell > 0$, i.e., the skew between neighboring nodes in different layers. $\hat{\sigma}_\ell := \max_{i \in [W]} \left\{ t_{\ell,i} - t_{m,j} \mid ((m, j), (\ell, i)) \in E \wedge m \neq \ell \right\}$. *Defined on pp.* 53, 175

$\boldsymbol{\sigma_{i,j}}$ is the skew between the first ticks of two fast clocks. *Defined on p.* 196

$\boldsymbol{\sigma(f)}$ is meant to be a bound on the skew between any two correct neighboring nodes, assuming that the system has already "stabilized" and f faults remaining. *Defined on p.* 68

$\boldsymbol{\tau_r}$ is minimum initial distance that the reader's pointer must be ahead of the writer's pointer. *Defined on p.* 202

$\boldsymbol{\tau_w}$ is minimum initial position of the writer's pointer in the ring buffer. *Defined on p.* 202

$\boldsymbol{t_i^k}$ is the time when the slow clock of clock domain i generates its k^{th} tick. *Defined on p.* 195

$\boldsymbol{t_i^{k,b}}$ is the time when the k^{th} fast clock of clock domain i generates its b^{th} tick. *Defined on p.* 195

$\boldsymbol{t_{\ell,i}^k}$ is the triggering time of node (ℓ, i). It denotes the time when the node forwards the k^{th} pulse of the grid. *Defined on pp.* 50, 175

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

T_{link}^{-} is the shortest time a received trigger message must be memorized. *Defined on p.* 68

T_{link}^{+} is the longest time a received trigger message can be memorized. *Defined on p.* 68

T_{sleep}^{-} is the shortest time a node must be sleep after being triggered. *Defined on p.* 68

T_{sleep}^{+} is the longest time a node can be asleep after being triggered. *Defined on p.* 68

W is the number of columns the grid has, which determine the width of the grid. *Defined on pp.* 48, 114, 174

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

ACRONYMS

ADPLL *all digital PLL*

CDN *clock distribution network*

CMOS *complementary metal-oxide-semiconductor*

CPU *central processing unit*

ECC *error-correcting code*

EDA *electronic design automation*

EMI *electromagnetic interference*

FCR *fault-containment region*

FDIR *failure detection, isolation and recovery*

FSM *finite state machine*

GALS *globally asynchronous, locally synchronous*

GPU *graphical processing unit*

HDL *hardware description language*

HSM *hybrid state machine*

IC *integrated circuit*

IQD^p *inter-quantile distance at level* p

MEMS *micro-electro-mechanical system*

MOS *metal-oxide-semiconductor*

MTBF *mean time between failure*

NCL *Null Convention Logic*

NGU *never give up strategy*

NoC *Network-on-Chip*

PLL *phase locked loop*

PRNG *pseudo random number generator*

PVT *process-voltage-temperature*

SEE *single-event effect*

SET *single-event transient*

SEU *single-event upset*

SoC *System-on-Chip*

STG *signal transition graph*

TCL *Tool Command Language*

TDC *time-to-digital converter*

TMR *triple modular redundancy*

TSM *transition state machine*

TSV *through silicon via*

VHDL *very high speed integrated circuit hardware description language*

VLSI *very-large-scale integration*

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

BIBLIOGRAPHY

[ACB+15] F. Abouzeid, S. Clerc, C. Bottoni, B. Coeffic, J. M. Daveau, D. Croain, G. Gasiot, D. Soussan, and P. Roche, „28nm FD-SOI technology and design platform for sub-10pJ/cycle and SER-immune 32bits processors“, in *41st European Solid-State Circuits Conference*, 2015, pp. 108–111. DOI: `10.1109/ESSCIRC.2015.7313840`.

[ALRL04] A. Avižienis, J.-C. Laprie, B. Randell, and C. Landwehr, „Basic concepts and taxonomy of dependable and secure computing“, *IEEE Transactions on Dependable and Secure Computing*, vol. 1, no. 1, pp. 11–33, Jan. 2004. DOI: `10.1109/TDSC.2004.2`.

[AM70] M. Abramson and W. O. J. Moser, „More Birthday Surprises“, *The American Mathematical Monthly*, vol. 77, no. 8, pp. 856–858, Oct. 1970.

[AMB07] M. L. Aranda, M. S. Maza, and D. P. Bautista, „An Experimental Comparison of Clock Distribution Networks for Systems on Chip“, in *4th International Conference on Electrical and Electronics Engineering*, Sep. 2007, pp. 377–380. DOI: `10.1109/ICEEE.2007.4345044`.

[AW04] H. Attiya and J. Welch, *Distributed Computing: Fundamentals, Simulations and Advanced Topics*. USA: John Wiley & Sons, Inc., 2004, ISBN: 0471453242.

[Bau05] R. C. Baumann, „Radiation-induced soft errors in advanced semiconductor technologies“, *IEEE Transactions on Device and Materials Reliability*, vol. 5, no. 3, pp. 305–316, Sep. 2005. DOI: `10.1109/TDMR.2005.853449`.

[BBB+03] R. Belschner, J. Berwanger, F. Bogenberger, C. Ebner, H. Eisele, B. Elend, T. M. Forest, T. Führer, P. Fuhrmann, F. Hartwich, B. Hedenetz, P. Heuts, R. Hugel, M. Kühlewein, P. Lohrmann, A. Millsap, B. Müller, M. Peller, M. Rausch, A. Schedl, C. Temple, J. Ungermann, H. Weiler, T. Wuerz, and M. Zinke, „FlexRay communication protocol“, EP1355456A1, Oct. 22, 2003.

[BBD+14] T. Bandi, J. Baborowski, A. Dommann, H. Shea, F. Cardot, and A. Neels, „Evaluation of silicon tuning-fork resonators under space-relevant radiation conditions“, in *Reliability, Packaging, Testing, and Characterization of MOEMS/MEMS, Nanodevices, and Nanomaterials XIII*, International Society for Optics and Photonics, vol. 8975, 2014, p. 89750I. DOI: `10.1117/12.2044209`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU Bibliothek WIEN Your knowledge hub

[BBR+08] J. D. Black, D. R. Ball II, W. H. Robinson, D. M. Fleetwood, R. D. Schrimpf, R. A. Reed, D. A. Black, K. M. Warren, A. D. Tipton, P. E. Dodd, N. F. Haddad, M. A. Xapsos, H. S. Kim, and M. Friendlich, „Characterizing SRAM Single Event Upset in Terms of Single and Multiple Node Charge Collection“, *IEEE Transactions on Nuclear Science*, vol. 55, no. 6, pp. 2943–2947, Dec. 2008. DOI: `10.1109/TNS.2008.2007231`.

[BC09] J. Bhasker and R. Chadha, *Static Timing Analysis for Nanometer Designs: A Practical Approach*, 1st. Springer Publishing Company, Incorporated, 2009, ISBN: 0387938192.

[BLSC10] M. Benigni, V. Liberali, A. Stabile, and C. Calligaro, „Design of rad-hard SRAM cells: A comparative study“, in *27th International Conference on Microelectronics*, May 2010, pp. 279–282. DOI: `10.1109/MIEL.2010.5490481`.

[BM06] T. Bjerregaard and S. Mahadevan, „A Survey of Research and Practices of Network-on-chip“, *ACM Computing Surveys*, vol. 38, no. 1, Jun. 2006. DOI: `10.1145/1132952.1132953`.

[BRd+06] S. A. Bota, J. L. Rossello, C. de Benito, A. Keshavarzi, and J. Segura, „Impact of Thermal Gradients on Clock Skew and Testing“, *IEEE Design & Test of Computers*, vol. 23, no. 5, pp. 414–424, May 2006. DOI: `10.1109/MDT.2006.126`.

[BSW11] M. Biely, U. Schmid, and B. Weiss, „Synchronous consensus under hybrid process and link failures“, *Theoretical Computer Science*, vol. 412, no. 40, pp. 5602–5630, 2011, Stabilization, Safety and Security. DOI: `10.1016/j.tcs.2010.09.032`.

[BV05] V. Bhandari and N. H. Vaidya, „On Reliable Broadcast in a Radio Network“, in *24th annual ACM Symposium on Principles of Distributed Computing*, Las Vegas, NV, USA, 2005, pp. 138–147. DOI: `10.1145/1073814.1073841`.

[BV06] E. Beigne and P. Vivet, „Design of on-chip and off-chip interfaces for a GALS NoC architecture“, in *12th IEEE International Symposium on Asynchronous Circuits and Systems*, Mar. 2006, pp. 172–183. DOI: `10.1109/ASYNC.2006.16`.

[BW01] S. Biaz and J. Welch, „Closed Form Bounds for Clock Synchronization Under Simple Uncertainty Assumptions“, *Information Processing Letters*, vol. 80, no. 3, pp. 151–157, 2001. DOI: `10.1016/S0020-0190(01)00151-X`.

[Cat66] I. Catt, „Time Loss Through Gating of Asynchronous Logic Signal Pulses“, *IEEE Transactions on Electronic Computers*, vol. EC-15, no. 1, pp. 108–111, Feb. 1966. DOI: `10.1109/PGEC.1966.264407`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[CCH12] S. Chellappa, L. T. Clark, and K. E. Holbert, „A 90-nm Radiation Hardened Clock Spine“, *IEEE Transactions on Nuclear Science*, vol. 59, no. 4, pp. 1020–1026, Aug. 2012. DOI: `10.1109/TNS.2012.2183647`.

[CCK+12] R. Chipana, E. Chielle, F. L. Kastensmidt, J. Tonfat, and R. Reis, „Soft-Error Probability Due to SET in Clock Tree Networks“, in *IEEE Computer Society Annual Symposium on VLSI*, Aug. 2012, pp. 338–343. DOI: `10.1109/ISVLSI.2012.39`.

[CCL+08] B. H. Calhoun, Y. Cao, X. Li, K. Mai, L. T. Pileggi, R. A. Rutenbar, and K. L. Shepard, „Digital Circuit Design Challenges and Opportunities in the Era of Nanoscale CMOS“, *Proceedings of the IEEE*, vol. 96, no. 2, pp. 343–365, Feb. 2008. DOI: `10.1109/JPROC.2007.911072`.

[CD03] W. S. Coates and R. J. Drost, „Congestion and starvation detection in ripple FIFOs“, in *9th International Symposium on Asynchronous Circuits and Systems*, May 2003, pp. 36–45. DOI: `10.1109/ASYNC.2003.1199164`.

[CG03] A. Chakraborty and M. R. Greenstreet, „Efficient self-timed interfaces for crossing clock domains“, in *9th International Symposium on Asynchronous Circuits and Systems*, May 2003, pp. 78–88. DOI: `10.1109/ASYNC.2003.1199168`.

[CH09] J. Chabloz and A. Hemani, „A flexible communication scheme for rationally-related clock frequencies“, in *2009 IEEE International Conference on Computer Design*, Oct. 2009, pp. 109–116. DOI: `10.1109/ICCD.2009.5413166`.

[Cha84] D. M. Chapiro, „Globally-Asynchronous Locally-Synchronous Systems“, PhD thesis, Stanford University, Oct. 1984.

[CK14] R. Chipana and F. L. Kastensmidt, „SET Susceptibility Analysis of Clock Tree and Clock Mesh Topologies“, in *2014 IEEE Computer Society Annual Symposium on VLSI*, 2014, pp. 559–564. DOI: `10.1109/ISVLSI.2014.33`.

[CKK+02] J. Cortadella, M. Kishinevsky, A. Kondratyev, L. Lavagno, and A. Yakovlev, *Logic Synthesis for Asynchronous Controllers and Interfaces*. Springer Science & Business Media, 2002, ISBN: 978-3-642-55989-1.

[CKK+97] J. Cortadella, M. Kishinevsky, A. Kondratyev, L. Lavagno, and A. Yakovlev, „Petrify: A Tool for Manipulating Concurrent Specifications and Synthesis of Asynchronous Controllers“, *IEICE Transactions on Information and Systems*, vol. E80-D, pp. 315–325, Mar. 1997.

[CKT+11] R. Chipana, F. L. Kastensmidt, J. Tonfat, R. Reis, and M. Guthaus, „SET susceptibility analysis in buffered tree clock distribution networks“, in *12th European Conference on Radiation and Its Effects on Components and Systems*, Sep. 2011, pp. 256–261. DOI: `10.1109/RADECS.2011.6131404`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[CKTR12] R. Chipana, F. L. Kastensmidt, J. Tonfat, and R. Reis, „SET susceptibility estimation of clock tree networks from layout extraction“, in *13th Latin American Test Workshop*, 2012, pp. 1–6. DOI: `10.1109/LATW.2012.6261256`.

[CKW18] L. Cardelli, M. Kwiatkowska, and M. Whitby, „Chemical reaction network designs for asynchronous logic circuits“, *Natural Computing*, vol. 17, no. 1, pp. 109–130, Mar. 2018, ISSN: 1572-9796. DOI: `10.1007/s11047-017-9665-7`.

[Cla67] W. A. Clark, „Macromodular Computer Systems“, in *Spring Joint Computer Conference*, Atlantic City, New Jersey, 1967, pp. 335–336. DOI: `10.1145/1465482.1465536`.

[CLM+14] Y. P. Chen, T. D. Loveless, P. Maillard, N. J. Gaspard, S. Jagannathan, A. L. Sternberg, E. X. Zhang, A. F. Witulski, B. L. Bhuva, T. W. Holman, and L. W. Massengill, „Single-Event Transient Induced Harmonic Errors in Digitally Controlled Ring Oscillators“, *IEEE Transactions on Nuclear Science*, vol. 61, no. 6, pp. 3163–3170, Dec. 2014. DOI: `10.1109/TNS.2014.2364813`.

[CN04] T. Chelcea and S. M. Nowick, „Robust interfaces for mixed-timing systems“, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 12, no. 8, pp. 857–873, Aug. 2004. DOI: `10.1109/TVLSI.2004.831476`.

[CNV+00] P. Cheynet, B. Nicolescu, R. Velazco, M. Rebaudengo, M. Sonza Reorda, and M. Violante, „Experimentally evaluating an automatic approach for generating safety-critical software with respect to transient errors“, *IEEE Transactions on Nuclear Science*, vol. 47, no. 6, pp. 2231–2236, Dec. 2000. DOI: `10.1109/23.903758`.

[Con02] C. Constantinescu, „Impact of deep submicron technology on dependability of VLSI circuits“, in *International Conference on Dependable Systems and Networks*, Jun. 2002, pp. 205–209. DOI: `10.1109/DSN.2002.1028901`.

[Cpp11] ISO, *ISO/IEC 14882:2011 Information technology — Programming languages — C++*, Third. Geneva, Switzerland: International Organization for Standardization, Sep. 2011. [Online]. Available: `https://www.iso.org/standard/50372.html`.

[CW75] G. R. Couranz and D. F. Wann, „Theoretical and Experimental Behavior of Synchronizers Operating in the Metastable Region“, *IEEE Transactions on Computers*, vol. C-24, no. 6, pp. 604–616, Jun. 1975. DOI: `10.1109/T-C.1975.224273`.

[DB99] C. Dike and E. Burton, „Miller and Noise Effects in a Synchronizing Flip-Flop“, *IEEE Journal of Solid-State Circuits*, vol. SC-34, no. 6, pp. 849–855, 1999. DOI: `10.1109/4.766819`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[DC93] P. H. Dietz and L. R. Carley, „An analog circuit technique for finding the median“, in *IEEE Custom Integrated Circuits Conference*, May 1993, pp. 6.1.1–6.1.4. DOI: `10.1109/CICC.1993.590571`.

[DDS87] D. Dolev, C. Dwork, and L. Stockmeyer, „On the Minimal Synchronism Needed for Distributed Consensus“, *Journal of the ACM*, vol. 34, no. 1, pp. 77–97, Jan. 1987. DOI: `10.1145/7531.7533`.

[DFL+13] D. Dolev, M. Függer, C. Lenzen, M. Perner, and U. Schmid, „HEX: Scaling Honeycombs is Easier than Scaling Clock Trees“, in *25th ACM Symposium on Parallelism in Algorithms and Architectures*, 2013, pp. 164–175.

[DFL+14] D. Dolev, M. Függer, C. Lenzen, M. Posch, U. Schmid, and A. Steininger, „Rigorously Modeling Self-Stabilizing Fault-Tolerant Circuits: An Ultra-Robust Clocking Scheme for Systems-on-Chip“, *Journal of Computer and System Sciences*, vol. 80, no. 4, pp. 860–900, 2014. DOI: `10.1016/j.jcss.2014.01.001`.

[DFL+16] D. Dolev, M. Függer, C. Lenzen, M. Perner, and U. Schmid, „HEX: Scaling Honeycombs is Easier than Scaling Clock Trees“, *Journal of Computer and System Sciences*, vol. 82, no. 5, pp. 929–956, 2016. DOI: `10.1016/j.jcss.2016.03.001`.

[DFLS14] D. Dolev, M. Függer, C. Lenzen, and U. Schmid, „**F**ault-tolerant **A**lgorithms for **T**ick-generation in **A**synchronous **L**ogic: Robust Pulse Generation“, *Journal of the ACM*, vol. 61, no. 5, 30:1–30:74, Sep. 2014. DOI: `10.1145/2560561`.

[DGKC09] R. Dash, R. Garg, S. P. Khatri, and G. Choi, „SEU hardened clock regeneration circuits“, in *10th International Symposium on Quality Electronic Design*, Mar. 2009, pp. 806–813. DOI: `10.1109/ISQED.2009.4810396`.

[DGS04] R. Dobkin, R. Ginosar, and C. Sotiriou, „Data synchronization issues in GALS SoCs“, in *10th International Symposium on Asynchronous Circuits and Systems*, Apr. 2004, pp. 170–179. DOI: `10.1109/ASYNC.2004.1299298`.

[DHP+04] K. Driscoll, B. Hall, M. Paulitsch, P. Zumsteg, and H. Sivencrona, „The real Byzantine Generals“, in *23rd Digital Avionics Systems Conference*, vol. 2, 2004, pp. 6.D.4–61. DOI: `10.1109/DASC.2004.1390734`.

[Dij74] E. W. Dijkstra, „Self-Stabilizing Systems in Spite of Distributed Control“, *Communications of the ACM*, vol. 17, no. 11, pp. 643–644, Nov. 1974. DOI: `10.1145/361179.361202`.

[DKM05] A. J. Drake, A. KleinOsowski, and A. K. Martin, „A self-correcting soft error tolerant flop-flop“, in *12th NASA Symposium on VLSI Design, Coeur d'Alene*, 2005, pp. 4–5.

[Dol00] S. Dolev, *Self-Stabilization*. MIT Press, 2000, p. 208, ISBN: 0262041782.

[Dol82] D. Dolev, „The Byzantine Generals Strike Again“, *Journal of Algorithms*, vol. 3, no. 1, pp. 14–30, 1982. DOI: `10.1016/0196-6774(82)90004-9`.

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[DW11] A. Dixit and A. Wood, „The impact of new technology on soft error rates“, in *2011 International Reliability Physics Symposium*, Apr. 2011, 5B.4.1–5B.4.7. doi: 10.1109/IRPS.2011.5784522.

[Ebe91] J. C. Ebergen, „A formal approach to designing delay-insensitive circuits“, *Distributed Computing*, vol. 5, no. 3, pp. 107–119, Dec. 1991. doi: 10.1007/bf02252954.

[Ein05] A. Einstein, „Zur Elektrodynamik bewegter Körper“, *Annalen der Physik*, vol. 322, no. 10, pp. 891–921, 1905. doi: 10.1002/andp.19053221004.

[Fai06] S. Fairbanks, *Method and Apparatus for a Distributed Clock Generator*, US patent no. 7,126,405 B2, Filed December 2nd, 2003, Issued October 24th, 2006, 2006.

[FB96] K. M. Fant and S. A. Brandt, „NULL Convention Logic: a complete and consistent logic for asynchronous digital circuit synthesis“, in *International Conference on Application Specific Systems, Architectures and Processors*, Aug. 1996, pp. 261–273. doi: 10.1109/ASAP.1996.542821.

[FDS10] M. Függer, A. Dielacher, and U. Schmid, „How to Speed-Up Fault-Tolerant Clock Generation in VLSI Systems-on-Chip via Pipelining“, in *8th European Dependable Computing Conference*, 2010, pp. 230–239. doi: 10.1109/EDCC.2010.35.

[FFL18] S. Friedrichs, M. Függer, and C. Lenzen, „Metastability-Containing Circuits“, *IEEE Transactions on Computers*, vol. 67, no. 8, pp. 1167–1183, Aug. 2018. doi: 10.1109/TC.2018.2808185.

[FFS09] G. Fuchs, M. Függer, and A. Steininger, „On the Threat of Metastability in an Asynchronous Fault-Tolerant Clock Generation Scheme“, in *15th IEEE Symposium on Asynchronous Circuits and Systems*, May 2009, pp. 127–136. doi: 10.1109/ASYNC.2009.15.

[FFSK06] M. Ferringer, G. Fuchs, A. Steininger, and G. Kempf, „VLSI Implementation of a Fault-Tolerant Distributed Clock Generation“, in *21st IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems*, Oct. 2006, pp. 563–571. doi: 10.1109/DFT.2006.67.

[FH90] P. Forshaw and R. Hahn, „Synchronous design: the right technique for digital ASICs“, in *3rd Annual IEEE Proceedings on ASIC Seminar and Exhibit*, Sep. 1990, P6/1.1–P6/1.5. doi: 10.1109/ASIC.1990.186126.

[Fid91] C. Fidge, „Logical time in distributed computing systems“, *Computer*, vol. 24, no. 8, pp. 28–33, Aug. 1991. doi: 10.1109/2.84874.

[FKG+12] X. Fan, M. Krstić, E. Grass, B. Sanders, and C. Heer, „Exploring Pausible Clocking Based GALS Design for 40-nm System Integration“, in *Design, Automation and Test in Europe*, Mar. 2012, pp. 1118–1121. doi: 10.1109/DATE.2012.6176663.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[FKLW18] M. Függer, A. Kinali, C. Lenzen, and B. Wiederhake, „Fast All-Digital Clock Frequency Adaptation Circuit for Voltage Droop Tolerance“, in *24th IEEE International Symposium on Asynchronous Circuits and Systems*, May 2018, pp. 68–77. DOI: `10.1109/ASYNC.2018.00025`.

[FL06] R. Fan and N. Lynch, „Gradient clock synchronization“, *Distributed Computing*, vol. 18, no. 4, pp. 255–266, Mar. 2006. DOI: `10.1007/s00446-005-0135-6`.

[FLP85] M. J. Fischer, N. A. Lynch, and M. S. Paterson, „Impossibility of Distributed Consensus with One Faulty Process“, *Journal of the ACM*, vol. 32, no. 2, pp. 374–382, Apr. 1985. DOI: `10.1145/3149.214121`.

[FM05] S. Fairbanks and S. Moore, „Self-Timed Circuitry for Global Clocking“, in *11th Symposium on Asynchronous Circuits and Systems*, 2005, pp. 86–96. DOI: `10.1109/ASYNC.2005.29`.

[FP86] E. G. Friedman and S. Powell, „Design and analysis of a hierarchical clock distribution system for synchronous standard cell/macrocell VLSI“, *IEEE Journal of Solid-State Circuits*, vol. 21, no. 2, pp. 240–246, Apr. 1986. DOI: `10.1109/JSSC.1986.1052510`.

[Fri01] E. G. Friedman, „Clock Distribution Networks in Synchronous Digital Integrated Circuits“, *Proceedings of the IEEE*, vol. 89, no. 5, pp. 665–692, May 2001. DOI: `10.1109/5.929649`.

[FS11] G. Fuchs and A. Steininger, „VLSI Implementation of a Distributed Algorithm for Fault-Tolerant Clock Generation“, *Journal of Electrical and Computer Engineering*, vol. 2011, no. 936712, 2011. DOI: `10.1155/2011/936712`.

[FS12] M. Függer and U. Schmid, „Reconciling Fault-Tolerant Distributed Computing and Systems-on-Chip“, *Distributed Computing*, vol. 24, no. 6, pp. 323–355, 2012. DOI: `10.1007/s00446-011-0151-7`.

[FSFK06] M. Függer, U. Schmid, G. Fuchs, and G. Kempf, „Fault-Tolerant Distributed Clock Generation in VLSI Systems-on-Chip“, in *6th European Dependable Computing Conference*, Oct. 2006, pp. 87–96. DOI: `10.1109/EDCC.2006.11`.

[GC00] V. Gutnik and A. Chandrakasan, „Active GHz Clock Network Using Distributed PLLs“, *IEEE Journal of Solid-State Circuits*, vol. 35, no. 11, pp. 1553–1560, Nov. 2000. DOI: `10.1109/4.881199`.

[GCRC15] A. Gujja, S. Chellappa, C. Ramamurthy, and L. T. Clark, „Redundant Skewed Clocking of Pulse-Clocked Latches for Low Power Soft Error Mitigation“, in *15th European Conference on Radiation and Its Effects on Components and Systems*, 2015, pp. 1–7. DOI: `10.1109/RADECS.2015.7365658`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[GD98] G. Geannopoulos and X. Dai, „An adaptive digital deskewing circuit for clock distribution networks“, in *IEEE International Solid-State Circuits Conference*, Feb. 1998, pp. 400–401. DOI: `10.1109/ISSCC.1998.672552`.

[GDM+08] G. Gielen, P. De Wit, E. Maricau, J. Loeckx, J. Martin-Martinez, B. Kaczer, G. Groeseneken, R. Rodriguez, and M. Nafria, „Emerging Yield and Reliability Challenges in Nanometer CMOS Technologies“, in *Design, Automation and Test in Europe*, Mar. 2008, pp. 1322–1327. DOI: `10.1109/DATE.2008.4484862`.

[GEB+06] M. J. Gadlage, P. H. Eaton, J. M. Benedetto, M. Carts, V. Zhu, and T. L. Turflinger, „Digital Device Error Rate Trends in Advanced CMOS Technologies“, *IEEE Transactions on Nuclear Science*, vol. 53, no. 6, pp. 3466–3471, Dec. 2006. DOI: `10.1109/TNS.2006.886212`.

[Gho14] S. Ghosh, *Distributed Systems: An Algorithmic Approach, Second Edition*, 2nd. Chapman & Hall/CRC, 2014, ISBN: 1466552972.

[Gin03] R. Ginosar, „Fourteen ways to fool your synchronizer“, in *9th International Symposium on Asynchronous Circuits and Systems*, May 2003, pp. 89–96. DOI: `10.1109/ASYNC.2003.1199169`.

[GK08] R. Garg and S. P. Khatri, „A novel, highly SEU tolerant digital circuit design approach“, in *IEEE International Conference on Computer Design*, Oct. 2008, pp. 14–20. DOI: `10.1109/ICCD.2008.4751834`.

[GM00] A. F. González and P. Mazumder, „Redundant arithmetic, algorithms and implementations“, *Integration*, vol. 30, no. 1, pp. 13–53, 2000. DOI: `10.1016/S0167-9260(00)00015-8`.

[Gre95] M. R. Greenstreet, „Implementing a STARI chip“, in *International Conference on Computer Design: VLSI in Computers and Processors*, Oct. 1995, pp. 38–43. DOI: `10.1109/ICCD.1995.528788`.

[Ham50] R. W. Hamming, „Error detecting and error correcting codes“, *The Bell System Technical Journal*, vol. 29, no. 2, pp. 147–160, Apr. 1950. DOI: `10.1002/j.1538-7305.1950.tb00463.x`.

[Hau95] S. Hauck, „Asynchronous design methodologies: an overview“, *Proceedings of the IEEE*, vol. 83, no. 1, pp. 69–93, Jan. 1995. DOI: `10.1109/5.362752`.

[HLL99] A. Hajimiri, S. Limotyrakis, and T. H. Lee, „Jitter and phase noise in ring oscillators“, *IEEE Journal of Solid-State Circuits*, vol. 34, no. 6, pp. 790–804, Jun. 1999. DOI: `10.1109/4.766813`.

[HOW06] B. Heidergott, G. J. Olsder, and J. van der Woude, *Max Plus at Work: Modeling and Analysis of Synchronized Systems: A Course on Max-Plus Algebra and Its Applications*. Princeton University Press, 2006, ISBN: 9780691117638.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[HSG09] A. Hansson, M. Subburaman, and K. G. W. Goossens, „Aelite: A Flit-Synchronous Network on Chip with Composable and Predictable Services“, in *Design, Automation and Test in Europe*, Nice, France, Apr. 2009, pp. 250–255. DOI: `10.1109/DATE.2009.5090666`.

[Huf57] D. A. Huffman, „The Design and Use of Hazard-Free Switching Networks“, *Journal of the ACM*, vol. 4, no. 1, pp. 47–62, Jan. 1957. DOI: `10.1145/320856.320866`.

[IEE93] IEEE-SA Standards Board, „IEEE Standard Multivalue Logic System for VHDL Model Interoperability (Std_logic_1164)“, *IEEE Std 1164-1993*, pp. 1–24, May 1993. DOI: `10.1109/IEEESTD.1993.115571`.

[Iwa09] H. Iwai, „Roadmap for 22nm and beyond“, *Microelectronic Engineering*, vol. 86, no. 7-9, pp. 1520–1528, Jul. 2009. DOI: `10.1016/j.mee.2009.03.129`.

[JH01] X. Jiang and S. Horiguchi, „Statistical skew modeling for general clock distribution networks in presence of process variations“, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 9, no. 5, pp. 704–717, Oct. 2001. DOI: `10.1109/92.953503`.

[JYG09] I. W. Jones, S. Yang, and M. Greenstreet, „Synchronizer Behavior and Analysis“, in *15th IEEE Symposium on Asynchronous Circuits and Systems*, May 2009, pp. 117–126. DOI: `10.1109/ASYNC.2009.8`.

[KB03] H. Kopetz and G. Bauer, „The time-triggered architecture“, *Proceedings of the IEEE*, vol. 91, no. 1, pp. 112–126, Jan. 2003. DOI: `10.1109/JPROC.2002.805821`.

[KBD+01] N. A. Kurd, J. S. Barkarullah, R. O. Dizon, T. D. Fletcher, and P. D. Madland, „A multigigahertz clocking scheme for the Pentium(R) 4 microprocessor“, *IEEE Journal of Solid-State Circuits*, vol. 36, no. 11, pp. 1647–1653, Nov. 2001. DOI: `10.1109/4.962284`.

[KBY02] D. J. Kinniment, A. Bystrov, and A. V. Yakovlev, „Synchronization Circuit Performance“, *IEEE Journal of Solid-State Circuits*, vol. SC-37, no. 2, pp. 202–209, 2002. DOI: `10.1109/4.982426`.

[KC87] L. Kleeman and A. Cantoni, „On the Unavoidability of Metastable Behavior in Digital Systems“, *IEEE Transactions on Computers*, vol. C-36, no. 1, pp. 109–112, Jan. 1987. DOI: `10.1109/TC.1987.5009455`.

[KCOW08] A. KleinOsowski, E. H. Cannon, P. Oldiges, and L. Wissel, „Circuit design and modeling for soft errors“, *IBM Journal of Research and Development*, vol. 52, no. 3, pp. 255–263, May 2008. DOI: `10.1147/rd.523.0255`.

[KCS+10] A. Korniienko, E. Colinet, G. Scorletti, E. Blanco, D. Galayko, and J. Juillard, „A Clock Network of Distributed ADPLLs Using an Asymmetric Comparison Strategy“, in *IEEE International Symposium on Circuits and Systems*, May 2010, pp. 3212–3215. DOI: `10.1109/ISCAS.2010.5537932`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[KGGV07] M. Krstić, E. Grass, F. K. Gürkaynak, and P. Vivet, „Globally Asynchronous, Locally Synchronous Circuits: Overview and Outlook“, *IEEE Design Test of Computers*, vol. 24, no. 5, pp. 430–441, Sep. 2007. DOI: `10.1109/MDT.2007.164`.

[KHL16] A. Kinali, F. Huemer, and C. Lenzen, „Fault-Tolerant Clock Synchronization with High Precision“, in *IEEE Computer Society Annual Symposium on VLSI*, Jul. 2016, pp. 490–495. DOI: `10.1109/ISVLSI.2016.88`.

[Kin08] D. J. Kinniment, *Synchronization and Arbitration in Digital Systems*. Wiley Publishing, 2008, ISBN: 9780470510827. DOI: `10.1002/9780470517147`.

[KK07] I. Koren and C. M. Krishna, *Fault-Tolerant Systems*, 1st. Morgan Kaufmann Publishers Inc., 2007, ISBN: 9780120885251.

[KL16] P. Khanchandani and C. Lenzen, „Self-stabilizing Byzantine Clock Synchronization with Optimal Precision“, in *Stabilization, Safety, and Security of Distributed Systems*, Cham: Springer International Publishing, 2016, pp. 213–230. DOI: `10.1007/978-3-319-49259-9_18`.

[KLP19] A. Kinali, C. Lenzen, and M. Perner, „Fault-tolerant High-Performance Clock Distribution“, E191 - Institut für Computer Engineering; Technische Universität Wien, Tech. Rep. TUW-278925, 2019. [Online]. Available: `https://publik.tuwien.ac.at/files/publik_278925.pdf`.

[Koo04] C.-Y. Koo, „Broadcast in Radio Networks Tolerating Byzantine Adversarial Behavior“, in *23rd annual ACM Symposium on Principles of Distributed Computing*, St. John's, Newfoundland, Canada, 2004, pp. 275–282. DOI: `10.1145/1011767.1011807`.

[Kop97] H. Kopetz, *Real-Time Systems, Design Pinciples for Distributed Embedded Applications*. Kluwer Academic Publishers, 1997.

[KW76] D. J. Kinniment and J. V. Woods, „Synchronisation and arbitration circuits in digital systems“, *Proceedings of the Institution of Electrical Engineers*, vol. 123, no. 10, pp. 961–966, Oct. 1976. DOI: `10.1049/piee.1976.0212`.

[Lac08] R. C. Lacoe, „Improving Integrated Circuit Performance Through the Application of Hardness-by-Design Methodology“, *IEEE Transactions on Nuclear Science*, vol. 55, no. 4, pp. 1903–1925, Aug. 2008. DOI: `10.1109/TNS.2008.2000480`.

[Lam12] L. Lamport, „Buridan's Principle“, *Foundations of Physics*, vol. 42, no. 8, pp. 1056–1066, Aug. 2012. DOI: `10.1007/s10701-012-9647-7`.

[Lam78] ——, „Time, Clocks, and the Ordering of Events in a Distributed System“, *Communications of the ACM*, vol. 21, no. 7, pp. 558–565, Jul. 1978. DOI: `10.1145/359545.359563`.

[LC66] W. M. Littlefield and T. J. Chaney, „The glitch phenomenon“, Computer Systems Laboratory, Washington University, Tech. Rep., 1966.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[LE09] R. Leidenfrost and W. Elmenreich, „Firefly Clock Synchronization in an 802.15.4 Wireless Network“, *EURASIP Journal on Embedded Systems*, vol. 2009, no. 1, p. 186 406, Jul. 2009. DOI: `10.1155/2009/186406`.

[LKM10] D.-J. Lee, M.-C. Kim, and I. Markov, „Low-power Clock Trees for CPUs“, in *IEEE/ACM International Conference on Computer-Aided Design*, Nov. 2010, pp. 444–451. DOI: `10.1109/ICCAD.2010.5653738`.

[LL88] J. Lundelius-Welch and N. A. Lynch, „A New Fault-Tolerant Algorithm for Clock Synchronization“, *Information and Computation*, vol. 77, no. 1, pp. 1–36, 1988. DOI: `10.1016/0890-5401(88)90043-0`.

[LLW10] C. Lenzen, T. Locher, and R. Wattenhofer, „Tight Bounds for Clock Synchronization“, in *Journal of the ACM*, vol. 57, Feb. 2010, 8:1–8:42. DOI: `10.1145/1667053.1667057`.

[LM11] D. Lee and I. L. Markov, „Multilevel tree fusion for robust clock networks“, in *IEEE/ACM International Conference on Computer-Aided Design*, Nov. 2011, pp. 632–639. DOI: `10.1109/ICCAD.2011.6105396`.

[LML+08] I. Loi, S. Mitra, T. H. Lee, S. Fujita, and L. Benini, „A Low-overhead Fault Tolerance Scheme for TSV-based 3D Network on Chip Links“, in *IEEE/ACM International Conference on Computer-Aided Design*, 2008, pp. 598–602. DOI: `10.1109/ICCAD.2008.4681638`.

[LR93] P. Lincoln and J. Rushby, „A formally verified algorithm for interactive consistency under a hybrid fault model“, in *23rd International Symposium on Fault-Tolerant Computing*, Jun. 1993, pp. 402–411. DOI: `10.1109/FTCS.1993.627343`.

[LSH+11] C. Lung, Y. Su, S. Huang, Y. Shi, and S. Chang, „Fault-tolerant 3D clock network“, in *48th ACM/EDAC/IEEE Design Automation Conference*, 2011, pp. 645–651. DOI: `10.1145/2024724.2024872`.

[LSP82] L. Lamport, R. Shostak, and M. Pease, „The Byzantine generals problem“, *ACM Transactions on Programming Languages and Systems*, vol. 4, no. 3, pp. 382–401, Jul. 1982. DOI: `10.1145/357172.357176`.

[LV62] R. E. Lyons and W. Vanderkulk, „The Use of Triple-modular Redundancy to Improve Computer Reliability“, *IBM Journal of Research and Development*, vol. 6, no. 2, pp. 200–209, Apr. 1962. DOI: `10.1147/rd.62.0200`.

[MA03] M. S. Maza and M. L. Aranda, „Interconnected Rings and Oscillators as Gigahertz Clock Distribution Nets“, in *13th ACM Great Lakes Symposium on VLSI*, New York, NY, USA, 2003, pp. 41–44. DOI: `10.1145/764808.764819`.

[MA04] M. S. Maza and M. L. Aranda, „Analysis and Verification of Interconnected Rings As Clock Distribution Networks“, in *14th ACM Great Lakes Symposium on VLSI*, Boston, MA, USA, 2004, pp. 312–315. DOI: `10.1145/988952.989027`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[Mar10] S. Marlow, Ed., *Haskell 2010 Language Report*. 2010.

[Mar81] L. R. Marino, „General Theory of Metastable Operation“, *IEEE Transactions on Computers*, vol. C-30, no. 2, pp. 107–115, Feb. 1981. DOI: `10.1109/TC.1981.6312173`.

[Mar90] A. J. Martin, „The Limitations to Delay-insensitivity in Asynchronous Circuits“, in *6th MIT Conference on Advanced Research in VLSI*, Boston, Massachusetts, USA, 1990, pp. 263–278, ISBN: 0-262-04109-X.

[Mat89] F. Mattern, „Virtual time and global states of distributed systems“, *International Workshop on Parallel and Distributed Algorithms*, vol. 1, no. 23, pp. 215–226, 1989.

[Max65] J. C. Maxwell, „A dynamical theory of the electromagnetic field“, *Philosophical Transactions of the Royal Society of London*, vol. 155, pp. 459–513, 1865.

[MB56] D. E. Muller and W. S. Bartky, *A theory of asynchronous circuits I*. University of Illinois, Graduate College, Digital Computer Laboratory, 1956.

[McP07] J. W. McPherson, „Reliability Trends with Advanced CMOS Scaling and the Implications for Design“, in *IEEE Custom Integrated Circuits Conference*, Sep. 2007, pp. 405–412. DOI: `10.1109/CICC.2007.4405763`.

[MDM04] C. Metra, S. Di Francescantonio, and T. M. Mak, „Implications of Clock Distribution Faults and Issues with Screening them During Manufacturing Testing“, *IEEE Transactions on Computers*, vol. 53, no. 5, pp. 531–546, May 2004. DOI: `10.1109/TC.2004.1275295`.

[MDMR01] C. Metra, S. Di Francescantonio, T. M. Mak, and B. Ricco, „Evaluation of clock distribution networks' most likely faults and produced effects“, in *IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems*, Oct. 2001, pp. 357–365. DOI: `10.1109/DFTVS.2001.966789`.

[Mes90] D. G. Messerschmitt, „Synchronization in Digital System Design“, *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 8, pp. 1404–1419, Oct. 1990. DOI: `10.1109/49.62819`.

[MG07] I. Miro Panades and A. Greiner, „Bi-Synchronous FIFO for Synchronous Circuit Communication Well Suited for Network-on-Chip in GALS Architectures“, in *1st International Symposium on Networks-on-Chip*, May 2007, pp. 83–94. DOI: `10.1109/NOCS.2007.14`.

[MGC+16] V. Malherbe, G. Gasiot, S. Clerc, F. Abouzeid, J. L. Autran, and P. Roche, „Investigating the single-event-transient sensitivity of 65 nm clock trees with heavy ion irradiation and Monte-Carlo simulation“, in *IEEE International Reliability Physics Symposium*, 2016, SE-3-1–5. DOI: `10.1109/IRPS.2016.7574639`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[Mil61] R. E. Miller, „An introduction to speed independent circuit theory“, in *1st and 2nd Annual Symposium on Switching Circuit Theory and Logical Design*, Oct. 1961, pp. 87–93. DOI: `10.1109/FOCS.1961.6`.

[MK11] T. Mittal and C.-K. Koh, „Cross Link Insertion for Improving Tolerance to Variations in Clock Network Synthesis“, in *International Symposium on Physical Design*, 2011, pp. 29–36. DOI: `10.1145/1960397.1960407`.

[MM95] R. Manohar and A. J. Martin, „Quasi-delay-insensitive circuits are Turing-complete“, California Institute of Technology Pasadena, Department of Computer Science, Tech. Rep., 1995.

[Moo06] G. E. Moore, „Moore's law at 40“, in *Understanding Moore's Law: Four Decades of Innovation*, D. Brook, Ed., Chemical Heritage Foundation, 2006, ch. 7, pp. 67–84, ISBN: 0941901416.

[Moo65] ——, „Cramming more components onto integrated circuits“, *Electronics*, vol. 38, no. 8, pp. 114–117, Apr. 1965.

[MSF14] M. Mounir Mahmoud, N. Soin, and H. A. H. Fahmy, „Design Framework to Overcome Aging Degradation of the 16 nm VLSI Technology Circuits“, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 33, no. 5, pp. 691–703, May 2014. DOI: `10.1109/TCAD.2014.2299713`.

[MTMR02] S. Moore, G. Taylor, R. Mullins, and P. Robinson, „Point to point GALS interconnect“, in *8th International Symposium on Asynchronous Circuits and Systems*, Apr. 2002, pp. 69–75. DOI: `10.1109/ASYNC.2002.1000297`.

[Mul55] D. E. Muller, *Theory of asynchronous circuits*. Urbana, Illinois: University of Illinois, Graduate College, Digital Computer Laboratory, 1955.

[MZ08] A. Mallajosyula and P. Zarkesh-Ha, „A Robust Single Event Upset Hardened Clock Distribution Network“, in *IEEE International Integrated Reliability Workshop Final Report*, Oct. 2008, pp. 121–124. DOI: `10.1109/IRWS.2008.4796101`.

[Nan08] NanGate, Inc, *NanGate 45nm Open Cell Library*, 2008. [Online]. Available: `http://www.nangate.com/?page_id=2325`.

[Nic05] M. Nicolaidis, „Design for soft error mitigation“, *IEEE Transactions on Device and Materials Reliability*, vol. 5, no. 3, pp. 405–418, Sep. 2005. DOI: `10.1109/TDMR.2005.855790`.

[NS14] R. Najvirt and A. Steininger, „Equivalence of clock gating and synchronization with applicability to GALS communication“, in *24th International Workshop on Power and Timing Modeling, Optimization and Simulation*, Sep. 2014, pp. 1–8. DOI: `10.1109/PATMOS.2014.6951873`.

[NS15] ——, „How to Synchronize a Pausible Clock to a Reference“, in *21st IEEE International Symposium on Asynchronous Circuits and Systems*, May 2015, pp. 9–16. DOI: `10.1109/ASYNC.2015.10`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[NT96] P. Nilsson and M. Torkelson, „A monolithic digital clock-generator for on-chip clocking of custom DSP's", *IEEE Journal of Solid-State Circuits*, vol. 31, no. 5, pp. 700–706, May 1996. DOI: `10.1109/4.509852`.

[NYH+15] E. Ng, Y. Yang, V. A. Hong, C. H. Ahn, D. B. Heinz, I. Flader, Y. Chen, C. L. M. Everhart, B. Kim, R. Melamud, R. N. Candler, M. A. Hopcroft, J. C. Salvia, S. Yoneoka, A. B. Graham, M. Agarwal, M. W. Messana, K. L. Chen, H. K. Lee, S. Wang, G. Bahl, V. Qu, C. F. Chiang, T. W. Kenny, A. Partridge, M. Lutz, G. Yama, and G. J. O'Brien, „The long path from MEMS resonators to timing products", in *28th IEEE International Conference on Micro Electro Mechanical Systems*, Jan. 2015, pp. 1–2. DOI: `10.1109/MEMSYS.2015.7050869`.

[OJ09] J. K. Ousterhout and K. Jones, *Tcl and the Tk Toolkit*. Addison-Wesley Professional, 2009, ISBN: 0-321-33633-X.

[ON03] T. Olsson and P. Nilsson, „Portable Digital Clock Generator for Digital Signal Processing Applications", *Electronics Letters*, vol. 39, no. 19, pp. 1372–1374, Sep. 2003, ISSN: 0013-5194. DOI: `10.1049/el:20030910`.

[ONM+00] T. Olsson, P. Nilsson, T. Meincke, A. Hemam, and M. Torkelson, „A Digitally Controlled Low-power Clock Multiplier for Globally Asynchronous Locally Synchronous Designs", in *Symposium on Circuits and Systems*, vol. 3, 2000, pp. 13–16. DOI: `10.1109/ISCAS.2000.855983`.

[ORM04] M. Omana, D. Rossi, and C. Metra, „Fast and low-cost clock deskew buffer", in *19th IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems*, Oct. 2004, pp. 202–210. DOI: `10.1109/DFTVS.2004.1347841`.

[ORM05] ——, „Low cost scheme for on-line clock skew compensation", in *23rd IEEE VLSI Test Symposium*, May 2005, pp. 90–95. DOI: `10.1109/VTS.2005.52`.

[Per13] M. Perner, „Self-Stabilizing Byzantine Fault-Tolerant Clock Distribution in Grids", Master's thesis, Technische Universität Wien, Institut für Technische Informatik, Treitlstr. 3/2/182-2, 1040 Vienna, Austria, 2013.

[PHS09] T. Polzer, T. Handl, and A. Steininger, „A Metastability-Free Multi-Synchronous Communication Scheme for SoCs", in *11th Symposium on Stabilization, Safety, and Security of Distributed Systems*, 2009, pp. 578–592. DOI: `10.1007/978-3-642-05118-0_40`.

[PK13] H. Park and T. Kim, „Comprehensive technique for designing and synthesizing TSV Fault-tolerant 3D clock trees", in *IEEE/ACM International Conference on Computer-Aided Design*, 2013, pp. 691–696. DOI: `10.1109/ICCAD.2013.6691190`.

[PK15] ——, „Synthesis of TSV Fault-Tolerant 3-D Clock Trees", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 2, pp. 266–279, 2015. DOI: `10.1109/TCAD.2014.2379645`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[PM95] C. L. Portmann and T. H. Y. Meng, „Supply Noise and CMOS Synchronization Errors“, *IEEE Journal of Solid-State Circuits*, vol. SC-30, no. 9, pp. 1015–1017, 1995. DOI: `10.1109/4.406400`.

[PMAS04] A. K. Palit, V. Meyer, W. Anheier, and J. Schloeffel, „Modeling and analysis of crosstalk coupling effect on the victim interconnect using the ABCD network model“, in *19th IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems*, Oct. 2004, pp. 174–182. DOI: `10.1109/DFTVS.2004.1347838`.

[PN95] G. A. Pratt and J. Nguyen, „Distributed synchronous clocking“, *IEEE Transactions on Parallel and Distributed Systems*, vol. 6, no. 3, pp. 314–328, Mar. 1995. DOI: `10.1109/71.372779`.

[PP05] A. Pelc and D. Peleg, „Broadcasting with Locally Bounded Byzantine Faults“, *Information Processing Letters*, vol. 93, no. 3, pp. 109–115, Feb. 2005. DOI: `10.1016/j.ipl.2004.10.007`.

[PS18a] M. Perner and U. Schmid, „Self-Stabilizing High-Speed Communication in Multi-Synchronous GALS Architectures“, in *24th IEEE International Symposium on On-Line Testing and Robust System Design*, Platja d'Aro, Spain, 2018.

[PS18b] ——, „Self-Stabilizing High-Speed Communication in Multi-Synchronous GALS Architectures“, Technische Universität Wien, Tech. Rep. TUW-268547, 2018. [Online]. Available: `http://publik.tuwien.ac.at/files/publik_268547.pdf`.

[PS19] ——, *Self-Stabilizing and Metastability-free Communication in Multi-Synchronous Architectures*, (submitted to MICPRO), 2019.

[PS57] J. H. F. Priebe and J. A. E. Spencer, *Transistor multivibrator circuits*, US Patent 2,787,712, Apr. 1957.

[PSL80] M. Pease, R. Shostak, and L. Lamport, „Reaching Agreement in the Presence of Faults“, *Journal of the ACM*, vol. 27, no. 2, pp. 228–234, Apr. 1980. DOI: `10.1145/322186.322188`.

[PSSL13] M. Perner, M. Sigl, U. Schmid, and C. Lenzen, „Byzantine Self-Stabilizing Clock Distribution with HEX: Implementation, Simulation, Clock Multiplication“, in *6th Conference on Dependability*, Barcelona, Spain, Aug. 2013, pp. 6–15, ISBN: 978-1-61208-301-8.

[PW72] W. W. Peterson and E. J. Weldon Jr., *Error-correcting Codes*. The MIT Press, 1972, ISBN: 0262527316.

[RAGM13] P. Roche, J. Autran, G. Gasiot, and D. Munteanu, „Technology downscaling worsening radiation effects in bulk: SOI to the rescue“, in *IEEE International Electron Devices Meeting*, Dec. 2013, pp. 31.1.1–31.1.4. DOI: `10.1109/IEDM.2013.6724728`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[RCN08] J. M. Rabaey, A. Chandrakasan, and B. Nikolic, *Digital Integrated Circuits*, 3rd. Upper Saddle River, NJ, USA: Prentice Hall Press, 2008, ISBN: 0132219107.

[RC018] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2018. [Online]. Available: `https://www.R-project.org/`.

[RFZJ13] M. Radetzki, C. Feng, X. Zhao, and A. Jantsch, „Methods for Fault Tolerance in Networks-on-chip", *ACM Computing Surveys*, vol. 46, no. 1, 8:1–8:38, Jul. 2013. DOI: `10.1145/2522968.2522976`.

[RHM06] A. Rajaram, J. Hu, and R. Mahapatra, „Reducing clock skew variability via crosslinks", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 25, no. 6, pp. 1176–1182, Jun. 2006. DOI: `10.1109/TCAD.2005.855928`.

[RMW+01] P. J. Restle, T. G. McNamara, D. A. Webber, P. J. Camporese, K. F. Eng, K. A. Jenkins, D. H. Allen, M. J. Rohn, M. P. Quaranta, D. W. Boerstler, C. J. Alpert, C. A. Carter, R. N. Bailey, J. G. Petrovick, B. L. Krauter, and B. D. McCredie, „A clock distribution network for microprocessors", *IEEE Journal of Solid-State Circuits*, vol. 36, no. 5, pp. 792–799, May 2001. DOI: `10.1109/4.918917`.

[RRA15] A. Rahman, M. Rahman, and F. Arifin, „Implimentation and evaluation of an efficient clock distribution network for deep-submicron technology", in *2nd International Conference on Electrical Information and Communication Technologies*, Dec. 2015, pp. 239–242. DOI: `10.1109/EICT.2015.7391953`.

[RWM06] S. Reddy, G. Wilke, and R. Murgai, „Analyzing Timing Uncertainty in Mesh-based Clock Architectures", in *Design, Automation and Test in Europe*, 2006, pp. 1097–1102. DOI: `10.1109/DATE.2006.243962`.

[SAA06] K. Sundaresan, P. E. Allen, and F. Ayazi, „Process and temperature compensation in a 7-MHz CMOS clock oscillator", *IEEE Journal of Solid-State Circuits*, vol. 41, no. 2, pp. 433–442, Feb. 2006. DOI: `10.1109/JSSC.2005.863149`.

[SAGZ14] C. Shan, F. Anceau, D. Galayko, and E. Zianbetov, „"Swimming pool"-like distributed architecture for clock generation in large many-core SoC", in *IEEE International Symposium on Circuits and Systems*, Jun. 2014, pp. 2768–2771. DOI: `10.1109/ISCAS.2014.6865747`.

[Sch86] F. B. Schneider, „A Paradigm for Reliable Clock Synchronization", in *Advanced Seminar of Local Area Networks*, Bandol, France, Apr. 1986, pp. 85–104.

[Sch93] M. Schneider, „Self-stabilization", *ACM Computing Surveys*, vol. 25, no. 1, pp. 45–67, Mar. 1993. DOI: `10.1145/151254.151256`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[SCY98] H.-S. Siu, Y.-H. Chin, and W.-P. Yang, „Byzantine agreement in the presence of mixed faults on processors and links“, *IEEE Transactions on Parallel and Distributed Systems*, vol. 9, no. 4, pp. 335–345, Apr. 1998. DOI: `10.1109/71.667895`.

[SD13] F. SalarKaleji and A. Dayyani, „A survey on Fault Detection, Isolation and Recovery (FDIR) module in satellite onboard software“, in *6th International Conference on Recent Advances in Space Technologies*, Jun. 2013, pp. 545–548. DOI: `10.1109/RAST.2013.6581270`.

[SEE96] M. Shams, J. C. Ebergen, and M. I. Elmasry, „A comparison of CMOS implementations of an asynchronous circuits primitive: the C-element“, in *International Symposium on Low Power Electronics and Design*, Aug. 1996, pp. 93–96. DOI: `10.1109/LPE.1996.542737`.

[Sei80] C. L. Seitz, „System timing“, in *Introduction to VLSI Systems*, Addison-Wesley, 1980.

[SFGP09] Y. Shi, S. B. Furber, J. Garside, and L. A. Plana, „Fault Tolerant Delay Insensitive Inter-chip Communication“, in *15th IEEE Symposium on Asynchronous Circuits and Systems*, May 2009, pp. 77–84. DOI: `10.1109/ASYNC.2009.21`.

[SG03] Y. Semiat and R. Ginosar, „Timing Measurements of Synchronization Circuits“, in *9th Symposium on Asynchronous Circuits and Systems*, May 2003, pp. 68–77. DOI: `10.1109/ASYNC.2003.1199167`.

[SGAZ14] C. Shan, D. Galayko, F. Anceau, and E. Zianbetov, „A reconfigurable distributed architecture for clock generation in large many-core SoC“, in *9th International Symposium on Reconfigurable and Communication-Centric Systems-on-Chip*, Montpellier, France: IEEE, May 2014, pp. 1–8. DOI: `10.1109/ReCoSoC.2014.6861349`.

[Sha38] C. E. Shannon, „A symbolic analysis of relay and switching circuits“, *Electrical Engineering*, vol. 57, no. 12, pp. 713–723, Dec. 1938. DOI: `10.1109/EE.1938.6431064`.

[SiT14] SiTime, „Shock and Vibration Performance Comparison of MEMS and Quartz-based Oscillators“, SiTime Corporation, Sunnyvale, California, Tech. Rep. SiT-AN10032 Rev 1.1, Feb. 2014, p. 13.

[SiT15] ——, „Resilience and Reliability of Silicon MEMS Oscillators“, SiTime Corporation, Sunnyvale, California, Tech. Rep. Sit-AN10045 Rev 2.2, Mar. 2015, p. 13.

[SKK+02] P. Shivakumar, M. Kistler, S. W. Keckler, D. Burger, and L. Alvisi, „Modeling the effect of technology trends on the soft error rate of combinational logic“, in *International Conference on Dependable Systems and Networks*, Jun. 2002, pp. 389–398. DOI: `10.1109/DSN.2002.1028924`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.
TU WIEN Bibliothek Your knowledge hub

[SM00] A. E. Sjogren and C. J. Myers, „Interfacing synchronous and asynchronous modules within a high-speed pipeline“, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 8, no. 5, pp. 573–583, Oct. 2000. DOI: 10.1109/92.894162.

[SS01] M. Saint-Laurent and M. Swaminathan, „A Multi-PLL Clock Distribution Architecture for Gigascale Integration“, in *IEEE Computer Society Workshop on VLSI*, Apr. 2001, pp. 30–35. DOI: 10.1109/IWV.2001.923136.

[SS04] M. Saint-Laurent and M. Swaminathan, „Impact of power-supply noise on timing in high-frequency microprocessors“, *IEEE Transactions on Advanced Packaging*, vol. 27, no. 1, pp. 135–144, Feb. 2004. DOI: 10.1109/TADVP.2004.825480.

[SSP+05] N. Seifert, P. Shipley, M. D. Pant, V. Ambrose, and B. Gill, „Radiation-induced clock jitter and race“, in *34rd Annual IEEE International Reliability Physics Symposium*, Apr. 2005, pp. 215–222. DOI: 10.1109/RELPHY.2005.1493087.

[ST87] T. K. Srikanth and S. Toueg, „Optimal Clock Synchronization“, *Journal of the ACM*, vol. 34, no. 3, pp. 626–645, Jul. 1987. DOI: 10.1145/28869.28876.

[Sut89] I. E. Sutherland, „Micropipelines“, *Communications of the ACM*, vol. 32, pp. 720–738, 6 Jun. 1989, ISSN: 0001-0782. DOI: 10.1145/63526.63532.

[SWL90] B. Simons, J. L. Welch, and N. Lynch, „Fault-tolerant Distributed Computing“, in, B. Simons and A. Spector, Eds., Springer-Verlag, 1990, ch. An Overview of Clock Synchronization, pp. 84–96, ISBN: 0-387-97385-0.

[SWR02] U. Schmid, B. Weiss, and J. Rushby, „Formally verified Byzantine agreement in presence of link faults“, in *22nd International Conference on Distributed Computing Systems*, Jul. 2002, pp. 608–616. DOI: 10.1109/ICDCS.2002.1022311.

[Tam09] S. Tam, „Modern Clock Distribution Systems“, in *Clocking in Modern VLSI Systems*, T. Xanthopoulos, Ed. Springer US, 2009, pp. 9–65, ISBN: 978-1-4419-0261-0. DOI: 10.1007/978-1-4419-0261-0_2.

[TGL07] P. Teehan, M. Greenstreet, and G. Lemieux, „A Survey and Taxonomy of GALS Design Styles“, *IEEE Design & Test of Computers*, vol. 24, no. 5, pp. 418–428, Sep. 2007. DOI: 10.1109/MDT.2007.151.

[TKP+15] A. S. Tararaksin, L. N. Kessarinskiy, A. A. Pechenkin, A. V. Demidova, A. V. Yanenko, D. V. Boychenko, and A. Y. Nikiforov, „Experimental Investigation of SELs in SiT8003 MEMS-Oscillators“, in *IEEE Radiation Effects Data Workshop*, Jul. 2015, pp. 1–4. DOI: 10.1109/REDW.2015.7336709.

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[Tur86] A. M. Turing, „Lecture to the London Mathematical Society on 20 February 1947“, in *Charles Babbage Institute Reprint Series for the History of Computing*, vol. 10, MIT Press, 1986.

[UMC03] UMC 90 nm, *90 nanometer*, Retrieved: March 28, 2019 from `http://www.umc.com/English/pdf/90nm_DM.pdf`, 2003.

[UMC05] UMC 65 nm, *65 nanometer*, Retrieved: March 28, 2019 from `http://www.umc.com/English/pdf/UMC65nm.pdf`, 2005.

[Vee17] H. J. M. Veendrick, „Robustness of Nanometer CMOS Designs: Signal Integrity, Variability and Reliability“, in *Nanometer CMOS ICs*, Springer International Publishing, 2017, pp. 429–493. DOI: `10.1007/978-3-319-47597-4_9`.

[Von56] J. Von Neumann, „Probabilistic logics and the synthesis of reliable organisms from unreliable components“, *Automata studies*, vol. 34, pp. 43–98, 1956.

[VPS+13] V. S. Veeravalli, T. Polzer, A. Steininger, U. Schmid, M. Hofbauer, K. Schweiger, H. Dietrich, K. Schneider-Hornstein, H. Zimmermann, K.-O. Voss, B. Merk, and M. Hajek, „An Infrastructure for Accurate Characterization of Single-Event Transients in Digital Circuits“, *Microprocessors and Microsystems*, vol. 37, no. 8-A, pp. 772–791, 2013. DOI: `10.1016/j.micpro.2013.04.011`.

[VPSS12] V. S. Veeravalli, T. Polzer, A. Steininger, and U. Schmid, „Architecture and Design Analysis of a Digital Single-Event Transient/Upset Measurement Chip“, in *15th Euromicro Conference on Digital System Design*, Sep. 2012, pp. 8–17. DOI: `10.1109/DSD.2012.26`.

[Wat97] S. Watts, „Overview of radiation damage in silicon detectors – Models and defect engineering“, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 386, no. 1, pp. 149–155, 1997, 5th International Workshop on Vertex Detectors. DOI: `10.1016/S0168-9002(96)01110-2`.

[WHG+09] L. Wissel, D. F. Heidel, M. S. Gordon, K. P. Rodbell, K. Stawiasz, and E. H. Cannon, „Flip-Flop Upsets From Single-Event-Transients in 65 nm Clock Circuits“, *IEEE Transactions on Nuclear Science*, vol. 56, no. 6, pp. 3145–3151, Dec. 2009. DOI: `10.1109/TNS.2009.2033997`.

[Wic16] H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2016, ISBN: 978-3-319-24277-4. [Online]. Available: `https://ggplot2.tidyverse.org`.

[WMC+16] H. B. Wang, N. Mahatme, L. Chen, M. Newton, Y. Q. Li, R. Liu, M. Chen, B. L. Bhuva, K. Lilja, S. J. Wen, R. Wong, R. Fung, and S. Baeg, „Single-Event Transient Sensitivity Evaluation of Clock Networks at 28-nm CMOS Technology“, *IEEE Transactions on Nuclear Science*, vol. 63, no. 1, pp. 385–391, 2016. DOI: `10.1109/TNS.2015.2509443`.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

[WS07] J. Widder and U. Schmid, „Booting clock synchronization in partially synchronous systems with hybrid process and link failures“, *Distributed Computing*, vol. 20, no. 2, pp. 115–140, Aug. 2007. DOI: `10.1007/s00446-007-0026-0`.

[WS09] ——, „The Theta-Model: achieving synchrony without clocks“, *Distributed Computing*, vol. 22, no. 1, pp. 29–47, Apr. 2009. DOI: `10.1007/s00446-009-0080-x`.

[WY03] T. Watanabe and S. Yamauchi, „An all-digital PLL for frequency multiplication by 4 to 1022 with seven-cycle lock time“, *IEEE Journal of Solid-State Circuits*, vol. 38, no. 2, pp. 198–204, Feb. 2003. DOI: `10.1109/JSSC.2002.807405`.

[YWC+06] C. Yeh, G. Wilke, H. Chen, S. Reddy, H. Nguyen, T. Miyoshi, W. Walker, and R. Murgai, „Clock distribution architectures: a comparative study“, in *7th International Symposium on Quality Electronic Design*, Mar. 2006, pp. 85–91. DOI: `10.1109/ISQED.2006.33`.

[ZS05] M. Zhang and N. R. Shanbhag, „A CMOS design style for logic circuit hardening“, in *43rd Annual IEEE International Reliability Physics Symposium*, Apr. 2005, pp. 223–229. DOI: `10.1109/RELPHY.2005.1493088`.

TU WIEN Bibliothek Your knowledge hub
Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar.
The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

Curriculum Vitæ
Martin Perner

Personal Information

Date of birth	January 3rd, 1988
Citizenship	Austrian

Education

2013 – 2019	**PhD program in Computer Engineering**, *TU Wien*.
2011 – 2013	**Master program in Computer Engineering**, *TU Wien*. Graduation with distinction and nomination for the best thesis award (EPILOG)
2007 – 2011	**Bachelor program in Computer Engineering**, *TU Wien*. Graduation with distinction
2002 – 2007	**Secondary College of Electrical Engineering**, *HTL Waidhofen/Ybbs*. General qualification for university entrance

---

Working Experience

2017 – 2018	**Research assistant (Projektassistent)**, *Embedded Computing Systems Group, TU Wien*.
2013 – 2017	**Assistant professor (Universitätsassistent)**, *Embedded Computing Systems Group, TU Wien*.
since 2015	**CTO/Head of Infrastructure and Software**, *Gaminside GmbH*.
Winter Term 2011 Winter Term 2012	**Tutor**, *Embedded Computing Systems Group, TU Wien*. In the Hardware-Software Codesign laboratory.