Diplomarbeit

Bruchmechanische Simulation von Indentierungsversuchen in Gallium Nitrid

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Diplom-Ingenieurs (Dipl.-Ing.)
eingereicht an der TU Wien, Fakultät für Maschinenwesen und Betriebswissenschaften
von

Clemens REICHEL, BSc

Mat.Nr.: 0828258

unter der Leitung von

Univ.-Doz. Dipl.-Ing. Dr.techn. Heinz PETTERMANN

und

Dipl.-Ing. Dr.techn. Martin SPRINGER

Institut für Leichtbau und Struktur-Biomechanik, E317

Diese Arbeit wurde von KAI Kompetenzzentrum Automobil- und Industrieelektronik GmbH. Technologiepark Villach, Europastraße 8, 9524 Villach im Rahmen einer Anstellung als Dipomand unterstützt.

Ich nehme zur Kenntnis, dass ich zur Drucklegung meiner Arbeit unter der Bezeichnung

Diplomarbeit

nur mit Bewilligung der Prüfungskommission berechtigt bin.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass die vorliegende Arbeit nach den anerkannten Grundsätzen für wissenschaftliche Abhandlungen von mir selbstständig erstellt wurde. Alle verwendeten Hilfsmittel, insbesondere die zugrunde gelegte Literatur, sind in dieser Arbeit genannt und aufgelistet. Die aus den Quellen wörtlich entnommenen Stellen, sind als solche kenntlich gemacht.
Das Thema dieser Arbeit wurde von mir bisher weder im In- noch Ausland einer Beurteilerin/einem Beurteiler zur Begutachtung in irgendeiner Form als Prüfungsarbeit vorgelegt. Diese Arbeit stimmt mit der von den Begutachterinnen/Begutachtern beurteilten Arbeit überein.

Wien, Mai, 2018

Unterschrift

TECHNISCHE UNIVERSITÄT WIEN
Vienna | Austria

Master Thesis

Fracture Mechanical Simulation of Indentation Testing in Gallium Nitride

carried out for the purpose of obtaining the degree of Master of Science (MSc), submitted at TU Wien, Faculty of Industrial Engineering, by

Clemens REICHEL, BSc

Mat.Nr.: 0828258

under the supervision of

Assoc. Prof. Dipl.-Ing. Dr.techn. Heinz PETTERMANN

and

Dipl.-Ing. Dr.techn. Martin SPRINGER

Institute of Leightweight Structures and Structural Biomechanics, E317

This work was supported by KAI Kompetenzzentrum Automobil- und Industrieelektronik GmbH. Technologiepark Villach, Europastraße 8, 9524 Villach, within the framework of an employment as a diploma thesis student.

I confirm, that going to press of this thesis needs the confirmation of the examination committee.

Affidavit

I declare in lieu of oath, that I wrote this thesis and performed the associated research myself, using only literature cited in this volume. If text passages from sources are used literally, they are marked as such.
I confirm that this work is original and has not been submitted elsewhere for any examination, nor is it currently under consideration for a thesis elsewhere.

Vienna, May, 2018

Signature

Contents

Acknowledgement

I would like to thank Heinz Pettermann and Martin Springer, my academic supervisors at the ILSB at TUW, for introducing me to this challenging and interesting topic. They have my gratitude for their constant, endless support and their selfless time and care for many stimulating discussions. They improved my accurateness, my scientific approach and my style of writing, steps that were necessary to finish this thesis.

I would like to thank Balamurugan Karunamurthy, my supervisor at KAI, for his constant support and time. He provided me with literature on the topic and gave me insight in the manufacturing and experimental testing of gallium nitride stacks. He has my gratitude for keeping an open ear for many questions.

I would like to thank Josef Fugger, the R&D director at KAI, for his trust to employ me as a diploma thesis student at KAI and therefore providing the necessary financial support for the work on this thesis.

Thanks also to Thomas Detzel at Infineon for his remarks regarding gallium nitride technology.

Gratitude also to ILSB for providing me with the electronic infrastructure on which the simulations were carried out. I would also like to thank my colleagues at ILSB and KAI not only for providing a pleasant working environment, but also for the many stimulating discussions on the topic.

Special thanks to my parents, Romeo and Gertrude Reichel, for their endless personal support and also financial aid, making it possible to continue my studies on the thesis topic even after my employment at KAI enabling me to gain a deeper understanding on the topic and treating the single aspects of the thesis in more depth.

I would also like to thank Katrin Kernbichler for her contribution on proofreading of this thesis.

Zusammenfassung

Konzepte der Finite-Elemente Methode (FEM) und der linear-elastischen Bruchmechanik (LEBM) werden angewendet um das mechanische Versagen von spröden, geschichteten Strukturen, wie sie in Halbleiter-Bauelementen der Leistungselektronik verwendet werden, zu untersuchen. Spannungen werden durch thermomechanische Lasten und den Unterschied in den thermo-mechanischen Materialeigenschaften der Schichten hervorgerufen. Besonderes Interesse liegt auf Galliumnitrid basierten Strukturen. Diese bieten aufgrund ihrer besonderen elektronischen Eigenschaften großes Potential für den Einsatz in der Leistungselektronik, erfordern aber aufgrund der spröden Materialeigenschaften und der Beschränkung der Herstellungsverfahren auf die Dünnfilmabscheidung auf fremden Substraten eine Auslegung gegen sprödes Versagens während der Herstellung und im Betrieb. Unter den verwendeten Substraten liegt der Fokus auf Silizium, da es besonderes günstig ist und in großen Durchmessern erhältlich ist, aber der Unterschied in den Materialeigenschaften starke Zugspannungen im Abkühlungsvorgang des Abscheideprozesses in der Galliumnitridschicht erzeugen kann, wenn nicht geeignete Gegenmaßnahmen ergriffen werden.

Ein Modellierungsansatz bestehend aus einer geeigneten Abstraktion des Problems, einer Auswahl geeigneter Methoden der Numerischen Verfahren der Bruchmechanik und einer Auswahl geeigneter Materialeigenschaften und Parameter für

die verwendeten Methoden wird erarbeitet. Der Modellierungsansatz wird angewendet um den Einfluss von Restspannungen beim Risswachstum bei Nanoindentierung mit einem pyramidenförmigen, dreiseitigen Berkovich-Indenter zu untersuchen.

Numerische Methoden der Bruchmechanik werden auf ihre Anwendbarkeit zur Vorhersage der Rissentstehung und Rissausbreitung untersucht. Als mögliche Methoden werden J-Integral, Erweiterte Finite Elemente Methode (XFEM), Kohäsivzonenmethode (CZM) und Virtuelle Rissausbreitungstechnik (VCCT) in Betracht gezogen. Eine Literaturrecherche führt zeigt, dass komplexe Rissgeometrien durch den Testvorgang in der Struktur zu erwarten sind. Nur die Rissinitiierung zu modellieren scheint nicht ausreichend um das spröde Versagen zu beschreiben. Für ein umfangreicheres Verständnis sind Methoden, welche auch die Rissausbreitung simulieren können, erforderlich. Aufgrund der Notwendigkeit Risse im dreidimensionalen Raum zu beschreiben und der zu erwartenden komplexen Rissgeometrie werden die anwendbaren Methoden auf die Kohäsivzonenmethode in dieser Arbeit eingegrenzt.

Verfügbare Materialkennwerte aus der Literatur werden diskutiert und geeignete Werte, die am besten die Verhältnisse im vorliegenden Problem beschreiben, zusammengefasst. Die spröden Materialeigenschaften können zu einer sehr kleinen Prozesszone führen in der die Bindungslösung an der Rissspitze stattfindet und andererseits zu sehr großen Rissen bei kleinen Indentierungstiefen. Die Simulation des Risswachstums erfordert sowohl bei der Initiierung als auch bei fortschreitendem Risswachstum eine ausreichende Anzahl von Kohäsivelementen in der Prozesszone und erfordert außerdem einen ausreichend kleinen Zeitschritt. Daraus resultiert

ein sehr großer Speicher- und Rechenbedarf, welcher die Grenzen der zur Verfügung stehenden Computersysteme überschreiten kann. Außerdem können die spröden Materialeigenschaften der involvierten Materialien zu einer Phase der Instabilität während der Rissentstehung führen oder zu vollständiger Instabilität des Risswachstums, welche wiederum die Konvergenz des numerischen Lösungsalgorithmus verhindern kann und entsprechende Gegenmaßnahmen erfordert.

Um numerische Konvergenz zu erreichen und das Problem mit zur Verfügung stehenden Computersystemen lösen zu können, wird das zugrundeliegende Spannung–Separations-gesetzt in der Prozesszone modifiziert, was zu einer Verbesserung des Konvergenzverhaltens führt. Der Einfluss der Modifikation auf die Vorhersage der Rissinitiierung und des Risswachstums werden anhand eines zweidimensionalen Biegebalkens und einer zweidimensionalen Abstraktion des Indentierungsproblems untersucht. Das Ziel der Vorstudie ist insbesondere die Anzahl der Freiheitsgrade zu reduzieren und dabei die Konvergenz aufrecht zu erhalten. Die im zweidimensionalen Modell untersuchten Aspekte werden kombiniert und die Modellkomplexität sukzessive erhöht. Die Ergebnisse helfen, das dreidimensionale Problem zu modellieren. Das Finite-Elemente Modell wird durch Vergleich mit numerischen und experimentellen Ergebnissen verifiziert.

Der Modellierungsansatz wird angewendet um das Risswachstum bei Nanoindentierung eines Siliziumwafers mit einer 7µm dicken Schicht aus Galliumnitrid unter den folgenden Vereinfachungen und Annahmen zu simulieren. Das Materialverhalten jeder Schicht wird mit einem linear-elastischen isotropen Konstitutivgesetz be-

schrieben. Zwischen den Trennflächen der Schichten wird eine ideale Verbindung angenommen. Die Rissebenen stehen rechtwinklig auf die Oberfläche des Wafers und fallen jeweils mit einer Kante des Indenters zusammen. Diese Annahme ist motiviert durch die parallele Ausrichtung der Kanten des Indenters mit den schwächsten Kristallebenen.

Die Ergebnisse der Finite Elemente Analyse zeigen stabiles Risswachstum bei Verwendung dieses Modellierungsansatzes. Kreisförmige Risse entstehen unterhalb der Indenterspitze und reichen bei maximaler Indentierungstiefe von 1.7µm bis in die Si-Schicht. Mit Einbringen von biaxialen Zugspannungen von 400MPa in der Schicht aus Galliumnitrid wird das Risswachstum instabil. Ein Vergleich mit experimentellen Ergebnissen, die im Vergleich zur elastischen Energie eine hohe dissipierte Energie zeigen, motiviert die Notwendigkeit zusätzlicher Mechanismen wie Reibung oder Plastizität, da die Größenordnung des Produkts aus Bruchzähigkeit und Rissfläche dafür nicht ausreichen. Eine rein elastische Berechnung führt zu dem Ergebnis, dass die maximalen Schubspannungen, auch unter Berücksichtigung nur bestimmter Gleitebenen, einen Wert von 10GPa in einem Volumen von der Ausdehnung der Dicke der GaN-Schicht übersteigen. Dieses Resultat ist ein starkes Argument für das mögliche Auftreten von Plastizität. Die Ergebnisse dieser Arbeit bestätigen die starke Auswirkung von Restspannungen im spröden Schichtverbund und zeigen mögliche Richtungen für weiterführende Untersuchungen, um ein tiefergehendes Verständnis der Zusammenhänge zwischen Material, Struktur, Last und Versagen zu erlangen und somit die mechanische Stabilität von Galliumnitrid basierten Strukturen zu

optimieren.

Abstract

The finite element method and linear elastic fracture mechanics concepts are applied to study the mechanical failure of brittle layered structures, as used in power-electronic semiconductor devices. Stresses are caused by thermo-mechanical loads and the thermal mismatch between the layers, and can be superimposed by residual stresses from production. The structure of particular interest is a silicon wafer with a gallium nitride layer on one side. These structures show a great potential in the application in power electronic devices due to their favorable electronic properties, but brittle fracture during manufacturing and operation has to be critically assessed due to the brittle material behavior and the restriction to thin film deposition on foreign substrates. Upon available substrates, silicon is preferred because its cheap and available in big diameters, but it can also generate strong tensile stresses in the gallium nitride layer due to the unfavorable mismatch in material properties.

A modelling approach is developed including a suitable abstraction of the problem, the choice of suitable numerical fracture mechanical methods, and choice of suitable material parameters for the chosen methods. The modelling approach is applied to study the influence of residual stresses on fracture during nanoindentation by a pyramidal Berkovich indenter.

Fracture mechanics approaches available within the framework of the finite element method are reviewed. They are evaluated with respect to their capability

to predict crack initiation and propagation for the posed problems. Potential approaches taken into consideration are initiation analysis methods such as the evaluation of a stress intensity factor or J-integral, as well as methods to simulate crack propagation based on node and element splitting, the incorporation of a cohesive zone model in an extended finite element ansatz or an interface element ansatz, as well as the virtual crack closing technique. A literature review and initial simulations showed, that crack patterns from nanoindentation testing can be of complex shape. The modelling of crack-initiation only does not seem sufficient to describe the brittle failure. For a more complete view, crack propagation methods are required. Because of the three-dimensional nature of the problem and the expected complex crack-pattern, crack propagation methods are restricted to the cohesive zone method within the present study.

Available material properties from the literature are discussed and the values most suited to resemble the conditions in the structure of interest are summarized. The brittle material behavior can lead to a very small extent of the process zone, in which material separates at the crack tip, and very large crack increments at low indentation increments. The simulation requires a sufficiently small time step and a sufficient number of elements to resolve the process zone during crack initiation and propagation. The required disk space to save the results and computational effort to solve the system of equations can exceed that of available computer systems. Furthermore, the brittle material behavior can lead to instability during crack propagation and therefore prevent convergence of the numerical solution algorithm.

To achieve numerical convergence and solve the problem with available computer systems, a modification to the underlying traction–separation law is made, which improves convergence. The influence of the modifications on the prediction of crack initiation and propagation is studied by crack propagation in a two dimensional bending test and a two dimensional abstraction of the indentation problem. These preliminary studies aim on reduction of degrees of freedom while preserving consistency of the results. The investigated aspects are combined, increasing the model complexity. The findings help to approach the three-dimensional prototype problem. The finite element model is verified, comparing numerical with experimental results.

The modelling approach is applied to simulate crack initiation and propagation in a silicon wafer with a 7µm thick gallium nitride layer on one side under the following assumptions. A linear elastic, isotropic constitutive law is used for each layer. At the interfaces, a perfect bond is assumed. The crack planes are assumed to be perpendicular to the wafer surface and aligned with the indenter edges. This assumption is in agreement with orienting the indenter edges with the weakest planes of the uniformly aligned crystals and the absence of interface cracks in experimental test results from the literature.

The results from the finite element analysis show, that the modelling approach predicts stable crack growth. Penny cracks initiate under the indenter and reach in the silicon base layer at the highest indentation depth of 1.7µm. When biaxial tensile residual stresses of 400MPa are incorporated in the gallium nitride layer,

crack growth exhibits unstable behavior. The comparison with experimental results from the literature, in which high energy dissipation is recorded, motivates the requirement of incorporation of additional mechanisms, like friction or plasticity, because the product of critical energy release rate and crack area is too small to be responsible for the high amount of dissipation. A strong argument for plasticity is the occurrence of shear stresses, exceeding 10 GPa in regions of the size of the thickness of the gallium nitride layer, as revealed by a linear elastic analysis, including a refined analysis, in which only certain slip planes are considered. The results confirm the strong impact of residual stresses in brittle layered stacks and give directions for future studies to gain a deeper understanding of the relationship between material properties, stack and load, which may be used to optimize the mechanical stability of gallium nitride based stacks.

1 Introduction

The usage of gallium nitride (GaN) in power-electronic semiconductor devices offers great potential for the improvement of electronic properties (see Ratschinski et al. 2011). Due to the brittle nature of epitaxial GaN based layer structures, a detailed understanding of the stress states in combination with its fracture behavior is required, which is of particular importance for the manufacturing of GaN-on-Silicon wafers. Thermo-mechanical loads and mismatch in thermo-mechanical properties of the layers lead to high stresses and strains. Cracks may be formed and can propagate through the structure in case of non-optimized growth conditions. The resistance of the structure to brittle failure during manufacturing and during operation has to be considered in the stack design process. Therefore, a knowledge about the dependency of design parameters on crack initiation and propagation is necessary. The aim of this thesis is, to establish a finite element modelling strategy to study the influence of the stack and the residual stresses in the structure.

To provide a more complete view on the importance of extensive research with the aim to increase liability of GaN, the advantageous electronic properties of GaN as

well as its many applications in electronic devices and the economic importance of GaN are discussed in chapter 2. The device architecture, the crystal structure of the materials incorporated in GaN devices and the manufacturing process of GaN stacks as well as the choice of substrates on which stacks are grown are very important to understand the origin of the mechanical behavior including formation of dislocations, elastic and plastic deformation and fracture. They are therefore discussed in detail in chapter 3. To gain a better understanding of fracture in these materials and devices, experimental and numerical methods are widely used for the fracture mechanical characterization of GaN stacks. To provide a solid basis for the discussion of the application of these methods to GaN stacks, an introduction to fracture mechanics is provided in chapter 4. The fracture mechanical characterization of GaN stacks by nanoindentation is discussed in chapter 5 to give directions for the modelling approach set up in this thesis. The modelling approach is discussed in chapter 6. Single aspects of the modelling approach are studied by use of simple load cases and simple two-dimensional (2D) geometries. These preliminary studies are discussed in a subsection of the modelling approach. With a deeper understanding of the modelling aspects gained from the preliminary studies, the modelling approach is applied to the full three-dimensional (3D) stack loaded by nano-indentation to study the influence of residual stresses in GaN stacks in chapter 7. The thesis closes with the conclusions in chapter 8.

2 Gallium nitride technology and applications

GaN and its related alloys aluminium nitride (AlN) and indium nitride (InN), as well as ternary alloys as e.g. aluminium gallium nitride (AlGaN) constitute the group III-N material system, one of the most important material systems in electronic device manufacturing. It proved to have very favorable properties for its utilization in optoelectronic device structures such as light emitting diode (LED) and laser structures, as well as power electronic device structures such as high electron mobility transistors (HEMTs). To motivate the need for extensive research in GaN to which this thesis adds, the advantageous properties of GaN and their utilization in many applications, as well as future expectations in GaN and its economic importance are discussed in the following sections.

2.1 Electronic properties of gallium nitride and its related alloys

The material system has gained such an importance in optoelectronic structures, because it allows LEDs and lasers to emit wavelengths in a broad spectrum including the whole visible range and even ultraviolet wavelengths at a very high efficiency. This broad spectrum is based on the materials wide bandgaps, leading to red for InN to deep UV for AlN (see Weber et al. 2000). In addition, due to its bond strength, GaN is resistent to defect formation and motion, which would lead to material degradation during operation (see Weber et al. 2000).

GaN has a high critical electric field, as it is characteristic for wide bandgap materials which is very favorable for power electronic devices. In addition, GaN shows a high saturation velocity compared to conventional semiconductor materials. With these properties, a high breakdown voltage in conjunction with a high cutoff frequency can be achieved, as demanded by high frequency high power transistors. Breakdown voltage and cutoff frequency constitute a primary tradeoff, therefore increasing the breakdown voltage, decreases the cutoff-frequency (see Ueda 2017). The product of breakdown voltage and cutoff frequency is used as a figure of merit, called the Johnson's limit. The Johnson's limit characterizes the suitability of a material for high frequency high power operations. It is reported to be 90 000 GHzV for GaN, a multiple of the value of 300 GHzV for silicon (Si), providing a strong quantitative argument for the potential of GaN compared to Si in this field of application

(see Ueda 2017). These favorable properties make GaN a key material for communication systems, which demand an increase in power delivery in the microwave range. In addition, due to the high breakdown voltage, the distance of anode and cathode can be reduced for traditional power systems to decrease resistance loss and allow downsizing, leading to reduced system cost. The high bond strength and high temperature stability as well as high thermal conductivity adds to the advantages of GaN for high power systems, in which cooling can play an important role (see Ueda 2017). These advantageous properties compared to traditional Si and gallium arsenide (GaAs) based devices are the reason for the high demand for GaN and its incorporation in a wide field of applications.

2.2 Applications for gallium nitride based devices

Due to the properties of GaN discussed in the previous section, it is an indispensable element of LED and laser structures in a broad field of applications in various markets, as reported in the literature (see e.g. Ueda 2017, Weber et al. 2000, MarketsandMarkets 2017). GaN based LEDs are used as light sources for illumination of displays in devices such as laptops, notebooks and mobile displays, projectors, televisions, monitors and brilliant large TV screens. Great demand for GaN based LEDs exists in the automotive industry. Interior lights such as headlights and signal lights as well as exterior lights such as fog lights, stop lights and dome lights of cars utilize the high efficiency of GaN based LEDs. They are also incorporated in street

lights and traffic lights for industrial illumination purposes. Home illumination is a great example of a whole line of industry, which the introduction of III-N based LEDs was expected to revolutionize, because white light can be generated very efficiently by utilization of phosphor converters, although they are still in development due to challenges with excess heat. Another revolution was expected in home-video applications by switching to III-N based blue lasers, which enable an increase in data storage by a factor of 3 and hence allow 6 hours of continuous video storage (see Weber et al. 2000). This increase is achieved, because the reduced wavelength leads to a smaller focus diameter and hence results in an increased memory storage capacity. A well known product, which is based on this technology is the Blu-Ray™ disk with its reading device based on an ultraviolett laser. (see Weber et al. 2000, MarketsandMarkets 2017)

GaN was also incorporated with great success in the field of power electronic devices. Three power electronic application layers can be distinguished (see Ueda 2017). GaN is applied in integrated circuits (ICs) of the personal layer, which is characterized by voltages not exceeding 12 V. Another layer is characterzied by a voltage between 100 V and 240 V and addresses the home applications. In this layer, GaN is incorporated in field-effect transistors (FETs). HEMT structures are used for high power high frequency microwave amplifiers. GaN demonstrated high output power with high gain in these applications (see Ueda 2017). These amplifiers are incorporated in devices of wireless communication systems such as cell phones. With the expected increase in efficiency and signal power, number of satellites flying

in low earth orbits and ground stationary transmission stations can be reduced. In addition, a better reception in mobile communication can be achieved (see Weber et al. 2000, MarketsandMarkets 2017). In the field of electric energy conversion from renewables, conversion efficiency plays a major role. GaN technology is expected to play a key role in this context (see Ueda 2017).

Further fields of applications for the rather unique properties of this material system are visible-transparent detectors. The expected great radiation hardness of GaN makes it a possible candidate for devices operating in environments of strong radiation such as satellites, which operate in space (see Ueda 2017), but also for solar cells (see MarketsandMarkets 2017). GaN could also be a candidate for wireless power systems (see Doronzo & Virk 2016). Another possible application is the optical scanning light radar in automotive industry, in which the high voltage and high frequency of GaN based devices allow an increased detecting range at a higher definition (see Doronzo & Virk 2016).

An insight on the economic importance of GaN technology can be gained from market studies (see e.g. MarketsandMarkets 2017). A market volume of USD 16.5 billion in 2016 and an expected forecast of USD 22.47 billion in 2023 is stated in the literature (see MarketsandMarkets 2017). GaN challenges traditional Si devices in many market segments, because the devices are smaller, lighter and more efficient (see MarketsandMarkets 2017). An example for such a market segment is the wireless base station market, in which GaN based semiconductors are expected to cost about half of Si based laterally diffused metal oxide semiconductor (LDMOS) per

watt (see Doronzo & Virk 2016).

The many applications and the high market share of GaN in optoelectronic devices show the great success of the technology. But whereas the manufacturing process for optoelectronic GaN devices is considered to be matured since the first growth of a GaN LED (see Keller 2017), research on GaN for application in power electronic devices is considered to be in a very early stage, if not just at the beginning. For a similar success of GaN based devices in already addressed and future power electronics market segments, reliability during operation is required while keeping wafer cost low to compete with the other material systems. In the manufacturing process a lot of effort is necessary to achieve this reliability, such as optimization and growth of many thin interlayers for strain engineering. The manufacturing process is therefore discussed in the next chapter to assess the mechanical challenges it provides with the main focus on its relation to fracture.

3 Manufacturing of power electronic devices and its influence on fracture

Fracture in GaN stacks can be a problem during the manufacturing process or during further handling and processing steps and has to be critically assessed and prevented by a suitable stack design. Fracture is strongly related to the basic material properties, the temperatures necessary in the manufacturing process and the combination of different materials. Therefore, the manufacturing process and its relation to fracture is discussed in this section including the basic crystal structure of the involved materials, the processing steps, the choice of the substrate and methods to avoid fracture. In addition, examples of the structure of a GaN on Si wafer and a power electronic transistor are discussed with the focus on stacking sequence of materials including typical layer thicknesses to gain insight on the structural complexity, that a modelling approach, as developed in this thesis, has to deal with.

3.1 Typical structure of AlGaN/GaN transistors

AlGaN/GaN transistors are the predominant devices for electronic applications (see Mishra & Guidry 2017). Typical lateral transistor structures are reported by Keller (2017). Lateral transistors consist of a thick GaN layer and a thin AlGaN layer on top of it. At the interface, a 2D electron gas forms in the GaN layer, on which the working principle of the transistor is based. On top of the AlGaN layer, the source, drain and gate are deposited. This structure is the most basic of a transistor with only gallium (Ga) atoms on top of the GaN layer and is referred to as Ga-polar HEMT. For a GaN based transistor, with only nitrogen (N) atoms on top of the GaN layer, referred to as n-polar HEMT, an additional GaN layer can be place on top of the AlGaN layer to form an electron gas at the bottom of this additional GaN layer. The electronic structure can be even more complex, when instead of a sharp interface between GaN and AlGaN, grading is used. Layers of GaN either intentionally doped by foreign atoms or unintentionally doped by dislocations due to the growth on a different material increase complexity. An example of a very complex structure can be found in the literature consisting of more than 10 layers of thicknesses from 0.7 nm to 740 nm (see Mishra & Guidry 2017).

In addition to the electronically motivated structure, which is also referred to as the active region of the device, most lateral transistors are grown on a foreign substrate, adding to the complexity of the overall stack. Most common are foreign substrates such as thick layers of Si due to the lower cost and larger available size compared to GaN substrates. To grow GaN on this foreign substrate an AlN nucle-

ation layer is commonly added on top of the substrate. In addition, a buffer layer consisting of interlayers is added to gradually combine the material properties of the AlN nucleation layer and GaN active layer to avoid cracking.

An example of a Ga wafer on a Si substrate is provided in figure 3.1 (see Marcon & Stoffels 2017). It consists of a thick Si substrate, on which a 250nm AlN nucleation layer is grown. On the nucleation layer, a buffer layer consisting of a 350nm $Al_{0.75}GaN$, a 350nm $Al_{0.5}GaN$ and a 2 µm $Al_{0.25}GaN$ layer are deposited. The active region of the device consists of a 150 nm GaN layer and a 15 nm AlGaN top layer deposited on the buffer layer.

3.2 Crystal structure of gallium nitride, its related alloys and the substrate material silicon

GaN, its related compounds AlN and InN as well as the substrates on which these materials are grown, are incorporated as thin layers of single-crystal structure in the devices. Electrical properties, mechanical properties and the manufacturing process of the brittle layered structures are strongly related to the crystal structure of these materials. In addition, the stresses that constitute during the manufacturing process and can lead to cracking, are strongly related to the differences in crystal structure of the involved materials. Therefore, the basic properties of the involved materials are discussed in this section to build a basis on which an introduction to the manufacturing process and challenges due to the crystal structure can be

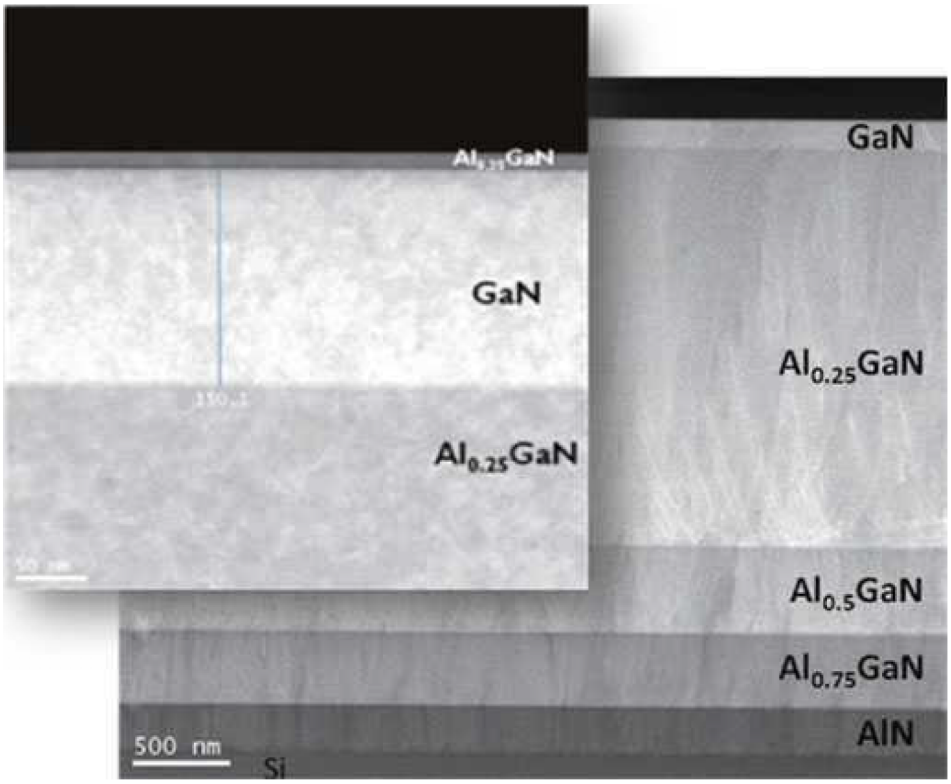

Figure 3.1: TEM image of the full GaN-on-Si epitaxial stack. The upper part, i.e., active region of the stack is magnified in the inset (Reprinted from Marcon & Stoffels 2017, p.54. Copyright ©2017 by Springer International Publishing Switzerland. Reprinted with permission.).

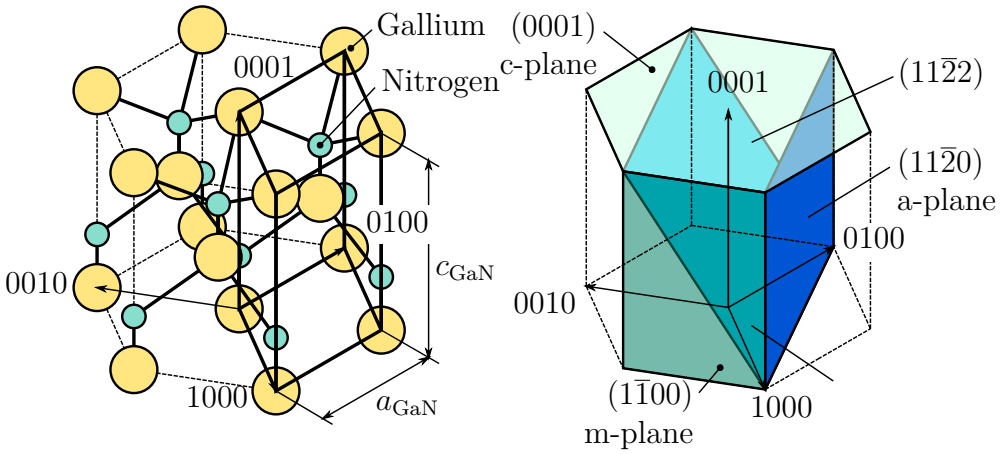

Figure 3.2: Crystal structure of GaN (left) and selected crystal planes (right).

discussed.

GaN, AlN and InN consist of an equal number of elements of each component. The first component in each compound, hence gallium, aluminium and indium, belong to the third group in the periodic system, whereas the second component nitrogen belongs to the fifth group. These compounds therefore constitute the group III nitrides, a subgroup of the group III-V compounds, from which many have properties that are very useful for semiconductors. Gallium Nitride is a compound, first synthetisized in 1931 by Johnson et al. (1931) (see Morkoç 2009). It grows, as all other group III-nitrides in Wurtzite structure, which is the only thermodynamically stable structure. Other possible structures are the unstable zinc-blende and the rock-salt structure. (Referring to Morkoç (2009), GaN and InN were also successfully grown in zinc-blende structure by epitaxial growth on Si, silicon carbide (SiC) and other cubic substrates.) The Wurtzite structure of GaN is shown in figure 3.2 by an isometric axonometric projection and is defined in its size by two lattice constants a_{GaN} and c_{GaN}. The lattice constant a_{GaN} denotes the distance between two Ga

atoms in planes where arragements of Ga atoms form hexagonal structures and c_{GaN} the distance between two Ga atoms perpendicular to such planes. In figure 3.2 they are shown as the horizontal plane and the vertical direction. The lattice constants can be considered as vectors, that span a lattice unit cell, as demonstrated by the thick black line in figure 3.2 (left). Taking into account the hexagonal structure, the hexagonal plane is spanned by three axes, pairwise including an angle of 120°. Any position in the crystal lattice can then be described by a linear combination of the three hexagonal unit vectors and the one vertical unit vector $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ and $\boldsymbol{c}$ with the coefficients h, k, i and l, written as a series of four digits, hence $hkil$ (e.g. 0001 denotes the vertical unit vector). To denote a set of parallel directions, the quadruple is surrounded by square brackets (e.g. [1000]) and to denote a set of equivalent crystal directions, the quadruple is surrounded by angle brackets (e.g. [1000], [0100] and [0010] are in the set of equivalent directions $\langle 1000 \rangle$). A plane is described by the coordinates, at which the plane intersects with the axis. Negative values are denoted by a line over the digit. The quadruple is surrounded by brackets to denote a set of parallel crystal planes (e.g. $(000\overline{1})$ for the hexagonal planes intersecting with the vertical axis at -1) which is referred to as the Miller-Bravais indices. To denote a set of equivalent crystal planes, the quadruple is surrounded by curly brackets (e.g. (1000), (0100) and (0010) are in the set of equivalent planes $\{1000\}$). The notation therefore follows Borchardt-Ott (2009), in which further information on crystal symmetries is provided. The crystal planes, used for manufacturing or exhibiting special properties are shown in figure 3.2 (right). The (0001) planes

are polar planes and referred to as c-planes. These planes also represent the basal planes of the crystal structure. In the vertical direction, c-planes with only one of the two components alternate. The 0001 direction is so important, because it is the most common growth direction in epitaxial growth of GaN (see Morkoç 2009). In addition the non-polar $(1\bar{1}00)$ m-plane and $(11\bar{2}0)$ a-plane are shown, which are also used for manufacturing of certain devices. These planes contain Ga and N atoms. The $\{11\bar{2}2\}$ planes are important slip-planes, as addressed in further sections. The previous considerations are also valid for AlN and InN, despite the different lattice constants.

The most promising and most challenging substrate to grow structures of the group III-nitrides to manufacture power electronic structures is silicon and therefore its crystal structure is addressed next. Silicon, grows in diamond cubic structure, as shown in figure 3.3. Positions and planes of this crystal structure are described by the coordinates h, k and l with respect to the three unit vectors. They are written as a series of digits hkl, which constitute the Miller indices. The covalent bonds are very strong and therefore lead to the high hardness of silicon. The diamond cubic structure has the same geometric arrangement as the Zinc-Blend structure (see Borchardt-Ott 2009). GaN wafers are most commonly grown in (0001) direction and therefore use a Si substrate with the (111) axis as growth axis, as it shows hexagonal symmetry, as shown in figure 3.3 (right).

An overview of the crystal structure, the lattice constants as well as the melting points and linear thermal expansion coefficient are listed for selected III-N materials

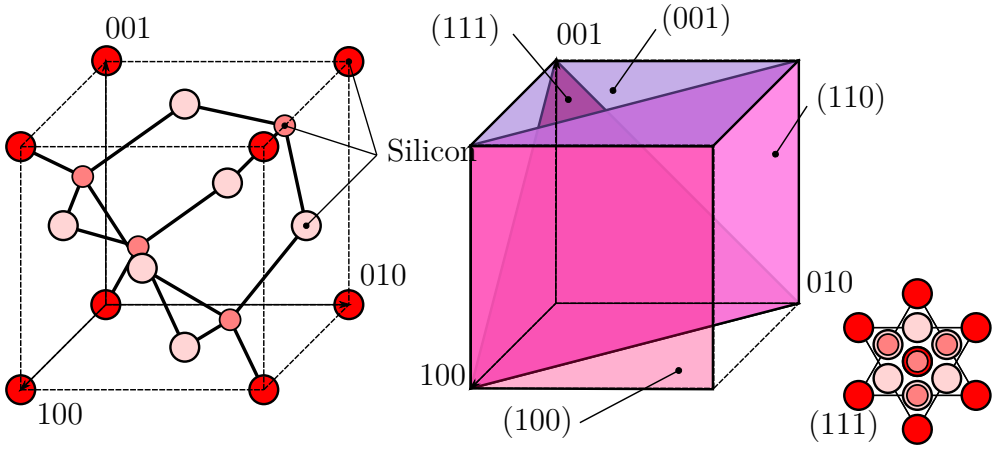

Figure 3.3: Crystal structure of Si (left), selected crystal planes (center) and projection of the lattice cell in 111 direction (right).

in table 3.1 and for the common substrates in table 3.2.

3.3 Epitaxial growth

The basic material properties of GaN and its related alloys have strong influence on how GaN can be primary shaped. GaN has a high melting point and very high vapor pressure of 45kbar at 2800°C (see Jackson & Schröter 2000), which make single crystal growth from a melt, as widely used for Si growth in the Czochralsky process, technically impossible, because at technically feasible pressures, the nitrogen would gasify from the melt.

GaN and stacks of GaN and its related alloys for electronic device manufacturing are therefore grown by epitaxial growth. This process belongs to the thin film deposition technology. With this technology very thin layers of a material are applied onto a substrate. The thin film deposition techniques share the basic steps in the deposition process. Chemical compounds, which contain the matter of which the

Table 3.1: Properties of the crystal structure of group III-nitrides (see Keller 2017)

		Material		
Property	Unit	GaN	AlN	InN
Crystal structure	-	Hexagonal, cubic	Wurtzite	Wurtzite
Lattice constant	(Å)	Hex 3.189/a, 5.185/c, cubic 4.52	3.1114/a, 4.9792/c	3.544/a, 5.718/c
Melting point	(°C)	2573 at 60kbar	2200	1100
Linear thermal expansion coefficient	($10^{-6}K^{-1}$)	5.6/a, 3.2/c	5.27/a, 4.15/c	3.8/a, 2.9/c

Table 3.2: Properties of the crystal structure of common substrates (see Keller 2017)

		Material		
Property	Unit	Si	SiC	Sapphire
Crystal structure	-	Diamond	Hexagonal, cubic	Trigonal
Lattice constant	(Å)	5.431	(6H) 3.086/a, 15.117/c, (4H) 3.073/a, 10.053/c, (3C) 4.3596/a, 3.073/c	4.75/a, 12.982/c
Melting point	(°C)	1415	2830	2044
Linear thermal expansion coefficient	($10^{-6}K^{-1}$)	2.6	4.4	7.5

crystal is grown, are transported to the surface of the substrate. At the surface, the compounds are decomposed and the species migrate on the surface. Finally, the crystal components bond at the growth front (see Kuech & Tischler 2000). Two categories of thin film deposition techniques can be distinguished considering the processes involved in transport of material to the surface and incorporation in the structure. These are the chemical deposition techniques and physical deposition techniques.

The most widely used techniques to grow GaN structures are metal-organic vapor phase epitaxy (MOVPE), hydride vapor phase epitaxy (HVPE) and molecular beam epitaxy (MBE). The choice of the technique has influence on the liability to fracture. Temperatures of the compounds and of the substrate are different depending on the process used. Therefore a different amount of residual stresses are incorporated in the structure. MOVPE and HVPE are chemical thin film deposition techniques and are based on crystal growth from the gas phase. These techniques utilize compounds of the elements of the crystal that have a high vapor pressure. The vapor is transported to the surface and to the growth front, where it reacts. The elements are then incorporated in the structure (see Kuech & Tischler 2000). HVPE uses the halide compound GaCl as Ga source and NH_3 as N source. As an advantage of this technique, thick layers of GaN can be grown with deposition rates of 100 μmh^{-1} (see Weber et al. 2000). It is still not the preferred technique, because the process cannot be controlled as well as with e.g. MOVPE leading to worse quality of the film (see Weber et al. 2000). In addition, a high amount of

stresses is incorporated in the structures, leading to wafer bending. An example of a wafer of 50mm in diameter and with a vertical height of up to 160μm, hence 3% of the diameter is provided in the literature demonstrating the strong wafer bow, that can occur and providing an imagination of the high amount of residual stresses (see Weber et al. 2000). Despite these disadvantages, the technique has great potential in combination with laser lift-off technique, by which energy is transported through the light-transparent substrate sapphire to the interface which leads to lift off of the GaN layer. By this lift-off the stresses are released as well, leading to a stress-free freestanding GaN layer, which can be used as a substrate for further growth (see Weber et al. 2000).

In contrast to HVPE, MOVPE uses metal organic precursors, such as $Ga(CH_3)_3$ instead of GaCl, transported by the carrier gas H_2 or N_2. The metal organic precursor leads to better process control and makes this technique the preferred technique in manufacturing advanced III/V nitrides also because of the possibility to deposit buffer layers and ternary compounds such as AlGaN. (see Weber et al. 2000). The material is deposited at temperatures around 1000°C and at atmospheric pressure (see Weber et al. 2000).

MBE is a physical thin film deposition technique, based on ballistic flight of the nutrients from a source to the substrate in a vacuum environment. A practical limit for the substrate temperature is 800°C, due to the necessity to control desorption at the surface (see Weber et al. 2000). The lower surface mobility leads to a worse quality compared to MOVPE.

A more detailed view on epitaxial growth and the chemical and pyhsical processes applying the techniques can be found in the literature (see e.g. Kuech & Tischler 2000, Weber et al. 2000, Keller 2017).

The overview on deposition techniques shows, that all processes are performed under high temperatures. After cool down to room temperature, stresses and strains remain in the stacks, that can lead to bending and fracture. The choice of substrate is very important as each substrate poses different challenges in preventing bending and fracture of the stack. Common substrates are therefore discussed in the following section.

3.4 Substrate choice and influence on fracture

GaN and GaN based heterostructures can be grown on different substrates by the deposition techniques discussed in the previous section. Common substrates are sapphire, SiC, Si and GaN. Sapphire was the first substrate GaN was grown on, whereas later growth on SiC and Si was made possible. With achievements in growing thick free-standing layers of GaN by growth on sapphire and separation from the substrate by epitaxial lateral overgrowth (ELOG), the growth on thick GaN substrates was made possible as well.

When growing GaN on a foreign substrate, such as SiC, Si and sapphire, stresses are incorporated due to the lattice mismatch between substrate and GaN and due to the thermal expansion mismatch. These stresses can lead to deformations of the

structure, such as wafer bow, and also to fracture. The substrates differ significantly in their lattice mismatch to GaN and its related alloys, as well as in their thermomechanical properties. Therefore the deposition of a certain architecture on a substrate poses different challenges. The substrates differ also significantly in cost per wafer area. In addition, the available diameter of the substrate depends on the material. Both properties have to be considered when aiming for efficient device manufacturing. Another important aspect is that each substrate causes different mechanical challenges during manufacturing and processing. The thermomechanical properties and lattice mismatch to GaN differ between the substrates and therefore lead to either tensile stresses or compressive stresses in the deposited layers. Tensile stresses can lead to nucleation and propagation of cracks, if not critically assessed and prevented by a suitable stack design. The lattice mismatch is highest for sapphire and Si, but with different signs. Thermal expansion coefficients of SiC and Si are lower than the coefficient of GaN, leading to tensile stresses, whereas that of sapphire is higher leading to compressive stresses. Therefore, Si poses the most challenges due to the combination of large lattice mismatch and large thermal expansion mismatch. It should be noted, that the lattice mismatch and thermal expansion mismatch depend not only on their corresponding lattice properties, but also on the temperature and the angle between the lattices, as e.g. GaN is rotated by an angle of 30° when grown on sapphire to reduce the lattice mismatch from 33% to 16% (see Keller 2017, Kukushkin et al. 2008). The resulting stresses depend also on the elastic tensors of the materials.

In addition to fracture mechanical problems, elastic deformation is important, because further processing steps demand a low deformation of the wafer generated during cool down by the differences in thermomechanical properties. The wafer bow can be controlled by optimization of deposition within certain bounds, but can even increase the problem of fracture if tensile stresses are necessary in the deposited layers to decrease the wafer bow. Above considerations are not only valid for growth of GaN, but also for growth of stacks that consist of several layers of different group III nitrides. Therefore, the choice of a substrate is most importantly influenced by the type of device grown. Further considerations in the choice of substrate include electronic properties of the substrate, as e.g. resistance of the substrate and also transparency for special technologies such as laser lift-off, by which deposited layers such as thick GaN layers can be separated from the substrate.

The main focus in this thesis lies on Si substrates, which are most rewarding for large scale wafer and device manufacturing but pose challenges to the stack design to prevent fracture. The mechanical challenges, methods to circumvent fracture, as well as the potential of GaN growth on this substrate is therefore discussed in more detail. Si substrates are very favorable for efficient production of lateral structures in large quantities, because they are available in large diameters up to 12" and are also the least expensive substrates compared to Sapphire, SiC and GaN substrates. As a downgrade, it is also the material with the largest lattice mismatch to GaN and a mismatch in thermal expansion coefficient. Strong optimization of the process is necessary to control the strong wafer bow when growing thick GaN layers. As the

thermal expansion coefficient of Si is lower than that of GaN, tension is incorporated in the GaN layer during cooldown, which can lead to fracture. Additional methods are therefore necessary to prevent fracture. These methods are based on exchange of the sudden change of material properties at the interface to a stepwise or even gradual change in material properties. The different methods are discussed in the next section. Despite these challenges GaN layers were grown up to a thickness of 14μm (see Keller 2017).

3.5 Crack mitigation in gallium nitride wafers grown on a silicon substrate

Several methods to mitigate cracking due to the lattice and thermal expansion mismatch between GaN and Si are reported in the literature (see Keller 2017). These methods utilize that the thermal expansion coefficient and lattice constant, as well as the Young's modulus of AlN lies between that of Si and GaN. A buffer layer between the thin AlN nucleation layer deposited on the Si substrate and the GaN film is constructed of layers of AlN, GaN and graded AlGaN to change the structure of the stress state. This procedure is also referred to as strain management. Four characteristic types of buffer layers are distinguished. The buffer layer can consist of a graded AlGaN layer, a AlN/GaN supperlattice with a great number of very thin alternating AlN and GaN layers, a combination of GaN layers with thinner AlN interlayers and step graded AlGaN layers (see Keller 2017). The graded layers always

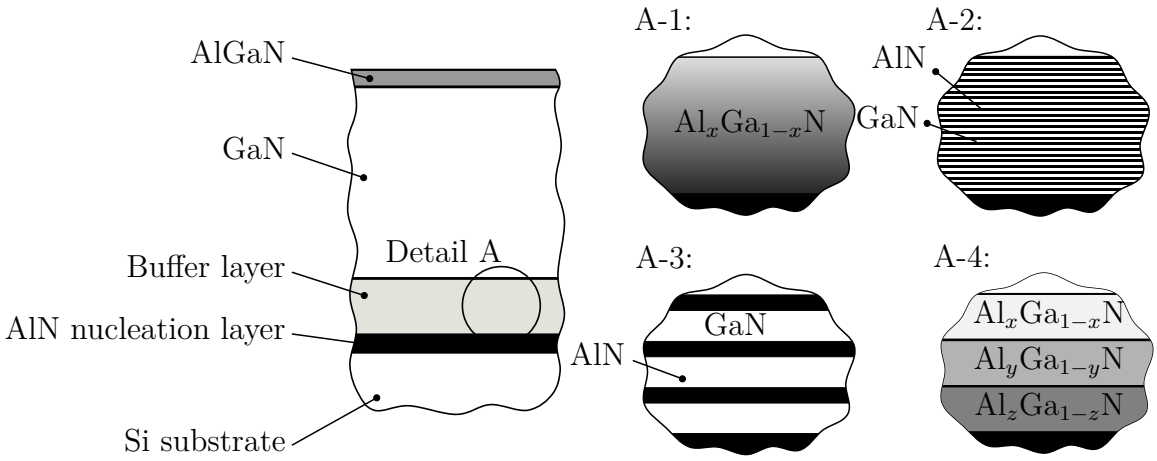

Figure 3.4: Example of a transistor stack (left) and buffer layer types for strain management (right); (A-1) Graded AlGaN layer, (A-2) GaN layers with AlN interlayers, (A-3) Superlattice of AlN and GaN, (A-4) Step-graded AlGaN layers; (based on Keller 2017, fig. 2.5)

increase the amount of Al_x from the nucleation layer to the GaN layer (x is the number of Al atoms compared to the number of Ga atoms). The incorporation of such a structure leads to compressive stresses at the growth temperature (see Keller 2017) and therefore reduce the liability to fracture during cooldown. A schematic of a transistor stack grown on a silicon substrate followed by a nucleation layer, a buffer layer, the GaN layer and an AlGaN top-layer is shown in figure 3.4 (left) and the four types of buffer layers are shown in figure 3.4 (right)

3.6 Summary

In this chapter, the attributes of GaN devices that influence or trigger brittle fracture were identified from a literature study with focus on the manufacturing process of power GaN devices. Important attributes are the layered composition of the device

with different materials, the differences in thermomechanical properties of these materials and the residual stresses due to lattice mismatch and thermomechanical mismatch that arise during cooldown. When studying fracture in these devices, stresses and strains that arise during further processing steps and quick handling steps are to be considered as well, because they can superimpose the residual stress state. A suitable approach to model fracture in these devices therefore requires to incorporate these attributes. The fracture mechanical characterization of such brittle structures by either experimental testing methods or numerical simulation is based on linear elastic fracture mechanics, which is introduced in the following chapter.

4 Fracture mechanics review

To characterize brittle materials such as GaN in terms of their fracture behavior, experimental testing methods such as nano indentation are widely applied. With this method a body of defined shape and material is pressed into the surface of the tested object. Cracks emanate from the object and measures for the resistance to cracking can be evaluated from the length of the cracks and the force and depth of indentation. The model describing the relation of the experimental results and the fracture resistance of this experimental testing method is based on linear elastic fracture mechanics (LEFM), which is widely employed to study fracture in brittle materials. It introduces terms and concepts such as crack initiation, crack propagation, crack modes, crack lengths, fracture toughness, energy release rate, the K-concept and the cohesive zone concept. Employing these concepts in a finite element method (FEM) environment, numerical fracture mechanical methods have been developed, which can be used to predict crack growth in complex structures. This chapter aims on introducing the concepts of LEFM in section 4.1 to provide a basis for the discussion of fracture behavior of GaN during nanoindentation, as

targeted in chapter 5, and the numerical methods to model fracture in GaN in a FEM environment, which are reviewed and assessed for their applicability to model fracture in GaN stacks loaded by nanoindentation in chapter 6.

4.1 Linear elastic fracture mechanics

LEFM is a theory to describe fracture on the basis of continuum mechanics in materials that exhibit only elastic material deformation, despite in a small area where material separates. The confinement of inelastic deformation to a small area during fracture is called localization, which is characteristic for many brittle materials and is therefore also widely used in the macroscopic description of fracture in fairly single crystal materials such as epitaxially grown GaN.

The continuum mechanics description of a fracture process within the theory of LEFM always starts with the assumption of a pre-existing or hypothetic crack in the structure (see Gross & Seelig 2011). A crack consists of a region without material bonding, the crack faces and in the case of a perfectly sharp crack, as it is the case for the assumption of the LEFM, a line, which marks the transition from separated to bonded material, the crack front. In a two dimensional projection, or where a crack front meets the surface, the crack front degenerates to a crack tip. A geometric sketch is shown in figure 4.1. When a crack gets instationary and starts to grow is called crack initiation. When a crack grows further it is called crack propagation. Crack nucleation, which is the formation of a region which enables a crack to initiate

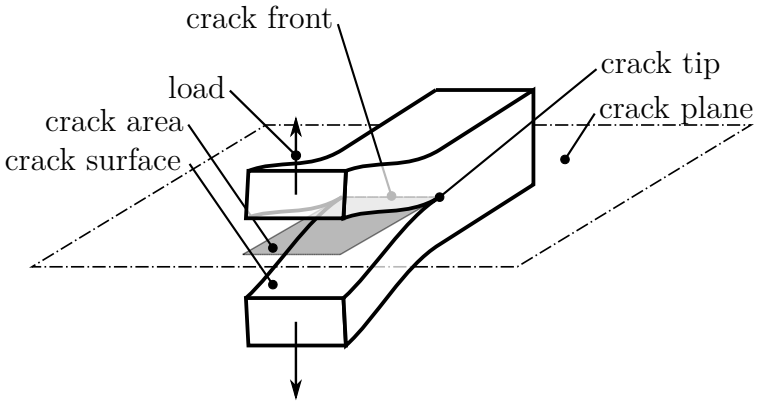

Figure 4.1: The geometric description of a crack. The crack tip or crack front marks the transition from bonded to separated material.

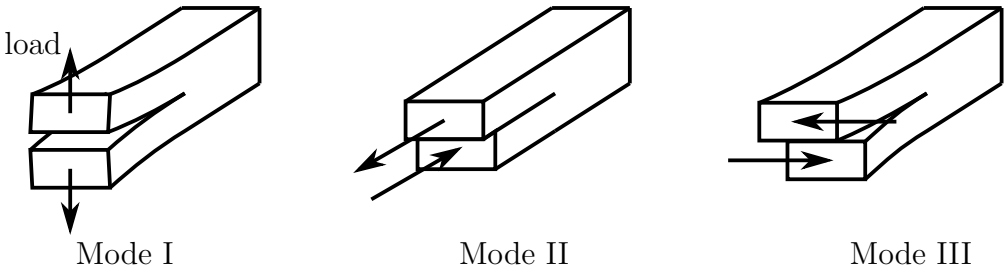

Figure 4.2: Three basic load cases for a crack. Arrows denote the direction of the load. Mode I describes tension perpendicular to the crack area, Mode II in-plane shear and Mode III out-of-plane shear.

by e.g. dislocation movement and annihilation, is not task of fracture mechanics but of continuum damage mechanics (see Gross & Seelig 2011).

Three basic crack modes exist, that describe the local deformation around the crack tip, as shown in figure 4.2. Mode I denotes opening due to tension perpendicular to the crack faces and mode II and mode III, crack opening due to shear stress states. Often, a superposition of two modes or all three modes occurs and is referred to as a mixed-mode.

For the case of a simple 2D problem, an analytical solution can be found for the stress and strain components around the crack tip. The analytical stress field result of the boundary value problem of a crack in an infinite plate loaded by mode I can be expressed as an expansion series of eigenfunctions. For each stress component, the first order term is dominant close to the crack tip and is given by

$$\sigma^{\mathrm{I}}_{rr}(r,\varphi) = \frac{1}{\sqrt{2\pi r}} K_{\mathrm{I}} \tilde{\sigma}^{\mathrm{I}}_{rr}(\varphi) \tag{4.1}$$

$$\sigma^{\mathrm{I}}_{\varphi\varphi}(r,\varphi) = \frac{1}{\sqrt{2\pi r}} K_{\mathrm{I}} \tilde{\sigma}^{\mathrm{I}}_{\varphi\varphi}(\varphi) \tag{4.2}$$

$$\tau^{\mathrm{I}}_{r\varphi}(r,\varphi) = \frac{1}{\sqrt{2\pi r}} K_{\mathrm{I}} \tilde{\sigma}^{\mathrm{I}}_{r\varphi}(\varphi) \tag{4.3}$$

with the stress components $\sigma^{\mathrm{I}}_{rr}(r,\varphi)$, $\sigma^{\mathrm{I}}_{\varphi\varphi}(r,\varphi)$ and $\tau^{\mathrm{I}}_{r\varphi}(r,\varphi)$ described in a polar coordinate system with its origin at the crack tip. The stress components are a function of the polar distance to the crack tip r and approach infinity by inverse of $\sqrt{r}$. Angular dependencies depend on the stress component and the crack mode and are subsumed in the functions $\tilde{\sigma}^{\mathrm{I}}_{rr}(\varphi)$, $\tilde{\sigma}^{\mathrm{I}}_{\varphi\varphi}(\varphi)$ and $\tilde{\tau}^{\mathrm{I}}_{r\varphi}(\varphi)$. The magnitude of the stresses depends on the geometry of the model and applied loads. These dependencies are summarized in the stress intensity factor K_{I} with I denoting that the factor is related to mode I loading. In the same way, solutions can be obtained for the other modes. The solutions scale by the stress intensity factors K_{I}, K_{II} and K_{III}. In the case of a three-dimensional crack-front, load cases can be superimposed,

leading to the expression

$$\sigma_{ij} = \frac{1}{\sqrt{2\pi r}} \left(K_{\mathrm{I}} \tilde{\sigma}_{ij}^{\mathrm{I}}(\varphi) + K_{\mathrm{II}} \tilde{\sigma}_{ij}^{\mathrm{II}}(\varphi) + K_{\mathrm{III}} \tilde{\sigma}_{ij}^{\mathrm{III}}(\varphi) \right) \tag{4.4}$$

for the stress components σ_{ij} with angular dependencies σ_{ij}^{I}, $\sigma_{ij}^{\mathrm{II}}$ and $\sigma_{ij}^{\mathrm{III}}$ (see Gross & Seelig 2011).

The near field solution is valid only in a certain region, the K-determined region, confined outwards by the effect of higher-order terms and inwards by a zone of inelastic material behavior. If the assumption holds, that the zone of inelastic behavior is small compared to the K-dominated region, then the processes in this zone are also determined by the stress intensity factor, which can therefore be seen as a state variable (see Gross & Seelig 2011). For simple geometries and loads, stress intensity factors can be evaluated analytically. As an example, the solution is given for a crack of length a in an infinite plate by

$$K_{\mathrm{I}} = \sigma_{\infty} \sqrt{\pi a} \tag{4.5}$$

with σ_{∞} denoting the uniaxial far field stress. The solution for various load cases and geometries can be found in the literature (see e.g. Gross & Seelig 2011, or other special literature). For more complex geometries, the FEM is widely used for the evaluation of the stress intensity factor or other measures of the loading of the crack front. The considerations above also hold for the case of a three-dimensional crack front, despite that the stress intensity factors for all three modes depend on the

position along the crack front (see Gross & Seelig 2011).

A fracture criterion can be formulated by comparing the stress intensity factor K_{I} with a critical material property K_{Ic}, the fracture toughness, which is obtained from experiments. In the case of single-mode I loading, the criterion for fracture is given by

$$K_{\mathrm{I}} = K_{\mathrm{I,c}}. \tag{4.6}$$

Similar expressions can be obtained for the other modes. In the case of mixed-mode fracture, combined criterions can be formulated as

$$f(K_{\mathrm{I}}, K_{\mathrm{II}}, K_{\mathrm{III}}, K_{\mathrm{I,c}}, K_{\mathrm{II,c}}, K_{\mathrm{III,c}}) = 0. \tag{4.7}$$

The formulation of a mixed criterion depends on the material and assumed direction of crack propagation (see Gross & Seelig 2011).

Critical stress intensity factors are evaluated from theoretical considerations, analytical and numerical models or experimental fracture mechanics. The assumptions of LEFM must hold for the specimen. To get reliable properties, the material applied should be tested at the characteristic length scale and with specimen from the material applied to the structure of interest (see Gross & Seelig 2011). This is difficult to achieve for pure GaN due to the difficulty in growing thick GaN films suitable for traditional experimental testing methods, and even more for GaN in layered devices due to the number of different layers of different thickness, residual stress state and different defect concentration. This aspect will be addressed in detail in chapter 5.

4.1.1 Energy release rate

The energy release rate

$$\mathcal{G} = -\frac{\mathrm{d}\left(\Pi_{\mathrm{int}} + \Pi_{\mathrm{ext}}\right)}{\mathrm{d}A} \tag{4.8}$$

is another measure for the load at the crack tip. It is equivalent to the derivative of the global energy balance before and after an incremental crack growth with respect to an infinitesimal extension of the crack area $\mathrm{d}A$. This global energy balance consists of the sum of the potential of elastic energy in the material Π_{int} and the potential of external forces Π_{ext}. If the assumptions of LEFM are applied, energy release rates and stress intensity factors are uniquely related for each mode by

$$\mathcal{G}_{\mathrm{I}} = \frac{K_{\mathrm{I}}^2}{E_{\mathrm{eff}}} \tag{4.9}$$

$$\mathcal{G}_{\mathrm{II}} = \frac{K_{\mathrm{II}}^2}{E_{\mathrm{eff}}} \tag{4.10}$$

$$\mathcal{G}_{\mathrm{III}} = \frac{K_{\mathrm{III}}^2}{2G} \tag{4.11}$$

with the shear modulus G and the effective modulus E_{eff}. For a general crack load, modes can be superimposed and the total energy release rate is given by

$$\mathcal{G} = \mathcal{G}_{\mathrm{I}} + \mathcal{G}_{\mathrm{II}} + \mathcal{G}_{\mathrm{III}} \tag{4.12}$$

which is the sum of the energy release rates of each mode. The effective modulus, which describes the relation between stress and strain components in the $r - \varphi$ plane, depends on the stiffness in the direction perpendicular to this plane due to

the Poisson effect. Considering again the example of a crack in an infinite plate, the stress components perpendicular to the $r-\varphi$ plane σ_{zi} are zero if the model thickness is considered to be very thin and loads only act within this plane. This is called a plane stress state. On the other hand if the material is very thick and deformations perpendicular to the plane are locked by boundary conditions or due to the thickness of the model, the strain components $\varepsilon_{\mathrm{zi}}$ perpendicular to the $r-\varphi$ direction can be assumed to be zero, which is referred to as a plane strain state. The effective modulus is expressed by

$$E_{\mathrm{eff}} = \begin{cases} E/(1-\nu^2) & \text{plane strain state} \\ E & \text{plane stress state} \end{cases} \tag{4.13}$$

as a function of the Young's modulus E and the Poisson ratio ν. Considering a cube with its faces aligned with a cartesian coordinate system and loaded by a normal stress in x-direction without constraint for deformations in y and z direction, the Poisson ratio

$$\nu = -\frac{\varepsilon_{\mathrm{yy}}}{\varepsilon_{\mathrm{xx}}} \tag{4.14}$$

is the ratio of normal strain in a direction perpendicular to the load to the normal strain in direction of the load for an isotropic material with ν independent of the load and strain components if the material is linear elastic.

A fracture criterion based on energy considerations can be formulated by

$$\mathcal{G} \geqslant \mathcal{G}_{\mathrm{c}} \tag{4.15}$$

comparing the energy release rate with a critical value $\mathcal{G}_{\mathrm{c}}$. If the energy released for an infinitesimal crack extension is greater than the critical energy necessary for an infinitesimal crack extension, the crack grows. This criterion is called Griffith crack criterion. The critical value can be derived from the critical fracture toughness, from experimental results or theoretical considerations, similar to the fracture toughness.

4.1.2 Stability of crack growth

A crack gets instationary and grows, when the fracture criterion of equation (4.6) or equation (4.15) is fullfilled. This section addresses the question whether this crack growth is stable or unstable. For the case of stable crack growth, there is a relation between the increase of the applied load expressed either by an increment of an applied force ΔF or an applied displacement Δu and the crack growth by an increment in crack area ΔA. For the case of unstable crack growth, as soon as the crack starts to grow, the crack growth gets independent of the applied load. The crack then grows until it becomes stable again or until it reaches the material boundaries. The equilibrium condition for an elastic body subject to quasi stationary

crack growth is given by a stationary point of the total energy

$$\frac{\mathrm{d}\left(\Pi_{\mathrm{int}} + \Pi_{\mathrm{ext}} + \Gamma\right)}{\mathrm{d}A} = 0 \tag{4.16}$$

and therefore of the derivative of the sum of the elastic energy Π_{int}, potential of external forces Π_{ext} and the fracture energy Γ with respect to the created fracture area. The fracture energy arises due to irreversible mechanisms during crack growth. A stable growth is then found by a minimum of the total energy

$$\frac{\mathrm{d}^2\left(\Pi_{\mathrm{int}} + \Pi_{\mathrm{ext}} + \Gamma\right)}{\mathrm{d}A^2} > 0 \tag{4.17}$$

and hence by a positive second derivative, or a positive first non-zero derivative of the total energy with respect to an infinitesimal crack extension. Fracture energy and critical energy release rate are related by

$$\mathcal{G}_{\mathrm{c}} = \frac{\mathrm{d}\Gamma}{\mathrm{d}A} \tag{4.18}$$

and hence equations (4.16) and (4.15) are equivalent. Incorporating (4.18) in the stability criterion (4.17) leads to the simplified stability criterion

$$\frac{\mathrm{d}\mathcal{G}}{\mathrm{d}A} < 0 \tag{4.19}$$

when the critical energy release rate is independent of the crack length ($\mathrm{d}^2\Gamma/\mathrm{d}A^2 = 0$) and taking into account equation (4.8). As an example, the growth of a crack in a double cantilever beam (DCB) specimen is considered. The compliance of the plate $C(a)$ depends on the crack length a. The plate is loaded by a displacement u_s perpendicular to the crack applied to a spring of constant stiffness C_s. In the limit of vanishing compliance of the spring ($C_\mathrm{s} \to 0$), which represents the application of a pure displacement, crack growth is always stable. In the case of infinite compliance of the spring, representing the application of a dead load, crack growth is unstable. In between the two extrema, the stability depends on the compliance of spring and plate, first and second derivative of the compliance of the plate with respect to crack length and the load, which itself depends on which force is necessary for a crack to initiate. The instability mechanism for a dead load in this example provides also information on the influence of residual stresses on stability. The dead load can be interpreted as uniaxial residual stresses incorporated in the structure. If residual stresses are strong enough to initiate a crack, it grows unstable. Even if residual stresses are lower than the load required to initiate a crack, an additional stress peak at the location of stress initiation, as it occurs under indentation load can lead to unstable crack growth.

Another example, which resembles some aspects of crack growth in GaN and for which an analytical solution exists, is that of crack growth in a thin layer subject to residual tension. The layer must be thinner than a critical thickness (see Gross &

Seelig 2011, eq. (4.111))

$$h_{\mathrm{c}} = \frac{4\sqrt{2}\mathcal{G}_{\mathrm{c}}E}{\sigma_{\infty}^{2}(1-\nu^{2})} \tag{4.20}$$

to prevent the layer to crack. The critical thickness is small if the critical energy release rate is small and if high residual stresses are incorporated, as it is the case for GaN stacks.

5 Fracture mechanical characterization by nanoindentation

The fracture mechanical behavior of GaN stacks is widely studied by application of nanoindentation. This experimental testing method enables the evaluation of fracture mechanical properties such as the fracture toughness and critical energy release rate from experimental results and an analysis based on the concepts of LEFM which were introduced in chapter 4. With conventional nanoindentation, an indenter of defined shape, is pressed into the surface. Cracks initiate under the indenter and emanate from it as the load is increased. After releasing the indenter, an indent remains in the surface. The maximum load, the indentation depth, the size of the indent and the crack length are experimental results of the test and used to evaluate the fracture mechanical properties.

Nanoindentation is widely used on GaN stacks, because it requires only small

portions of material and no specialized geometry of the specimen is necessary. The importance of these advantages becomes especially evident, considering the efforts necessary to grow thick, free-standing GaN or bulk GaN, necessary to prepare specimen of defined geometry, as required by specifications of classical bending tests. Nanoindentation therefore circumvents the limitations of GaN stacks in fulfilling the requirements of classical testing methods. In addition, because only small portions of material are necessary, many measurements can be carried out on a single specimen, allowing to study the spatial distribution of fracture mechanical properties over the wafer surface. From such tests, information on the differences in material properties close to the boundaries of a wafer compared to that in the interior can be gained. Even more information, including the structural changes in the material, can be gained by recording the load displacement history during the indentation process. This improved nanoindentation technique is referred to as instrumented indentation technique (IIT). Supplementary techniques such as cutting, imaging and spectroscopy techniques allow to get an insight into the actual structural as well as morphological or crystallographic changes under the visible surface.

This chapter provides an overview on the testing procedure to provide a suitable basis for discussion of its application to GaN and for modelling the prototype problem. A typical testing procedure, as well as a discussion of the versatility of nanoindentation based testing procedure is therefore provided in section 5.1. Typical crack-patterns, load displacement-curves and also deformation mechanisms based on the application of nanoindentation as found in the literature, are discussed in sec-

tion 5.2. Section 5.3 provides a summary of aspects which are expanded into the modelling approach.

5.1 Nanoindentation testing procedure

A detailled overview on IIT to evaluate Young's modulus and hardness is provided in the literature from which important aspects are recapitulated (see e.g. Hay 2009). Details from experimental procedures used to apply nanoindentation to study GaN are added to the discussion of the testing procedure (see e.g. Drory et al. 1996, Yonenaga et al. 2000, Yonenaga 2001, Tsai et al. 2008, Ratschinski et al. 2011, 2012, 2013).

A typical nanoindentation testing procedure is shown in figure 5.1 by example of the evolution of the load drawn against time (left, top), the evolution of the load drawn against the displacement of the indenter (left, bottom) and cross-sections of the indentation at characteristic loads (center; right, top), as well as a top view on the surface after release of the indenter (right, bottom). In the first step, shown in figure 5.1 (1), the indenter is brought into contact with the surface of the specimen until a response by the stiffness of the specimen is sensed. This can be repeated several times to exclude parasitic phenomena from the measurement (see Tsai et al. 2008). After this procedure, the displacement of the indenter is increased, as shown in figure 5.1 (2). Increasing the displacement further, at some point a crack forms under the indenter, outlined by the grey region marking a penny-wise crack in figure

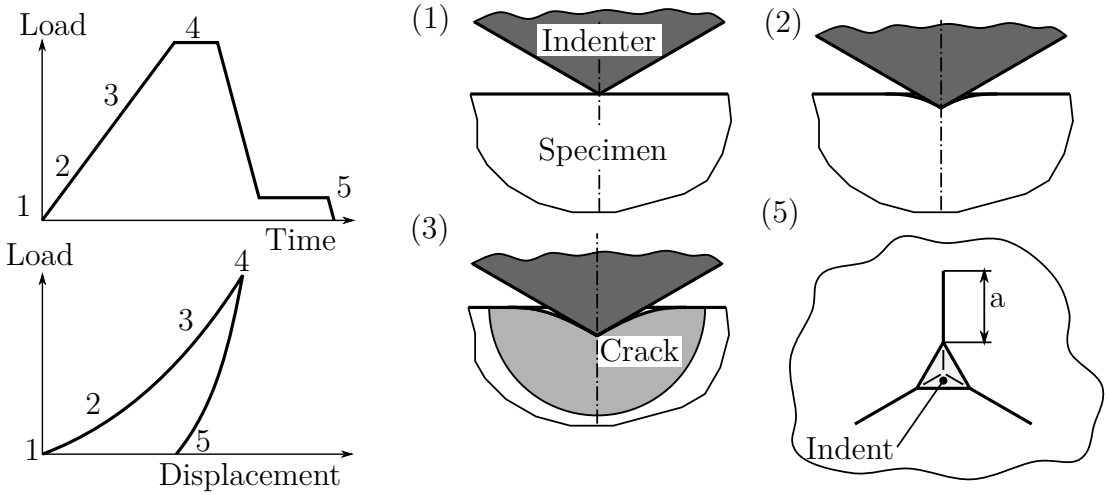

Figure 5.1: Steps of the indentation process outlined by means of the load–time and load–displacement curve (left) as well as cross-sections of the indentation (right)

5.1 (3). After reaching the maximum load (4), the indenter is released to a portion of the maximum load. It is often hold at this load for a certain time to determine the influence of thermal drift (see Hay 2009). After fully removing the indenter, an indent remains in the surface, as shown in figure 5.1 (5) by example of an indent from a pyramidal three–sided Berkovich indenter. The intersection of the cracks with the surface are shown by the solid lines oriented with the indenter edges. The length of the crack a as well as the area of the remaining indent A can be measured and are used in the analysis to evaluate the mechanical and fracture mechanical properties.

The indentation procedure can be performed with intenders of different shapes. The indenter geometries mainly discussed in this thesis include the four sided Vickers and the three sided Berkovich indenter, which is also the indenter used for the prototype problem. Other indenter shapes used for nanoindentation in GaN include

the spherical and Knoop indenter (see Tsai et al. 2008). Indenters are characterized by the area function

$$A_{\text{contact}} = f(h_{\text{contact}}) \tag{5.1}$$

providing the relation between the cross section of the indenter at the transition line between region of contact between the specimen and the indenter and a region without contact A_{contact}, and the depth over which the indenter is in contact with the surface h_{contact}, referred to as the contact depth.

The Berkovich diamond indenter is the most common indenter because it has the advantage that it is easy to manufacture, it induces plasticity at smaller loads than other indenter types and it has a large angle, which leads to less influence of friction compared to other indenters (see Hay 2009). In addition, Vickers Hardness can be evaluated from Berkovich indentation tests. It is therefore not surprising, that most literature found on nanoindentation in GaN stacks is based on this indenter type. In the analysis stages, the mechanical and fracture-mechanical properties of the material are evaluated from experimental testing results. The evaluation of Young's modulus E and hardness H can be found in the literature (see e.g. Hay 2009). The evaluation of E and H require information on the contact area or contact depth, which are related by the area function (5.1). To circumvent a direct measurement of the contact area, the Oliver Pharr model is widely used, which relates the contact area A_{contact} to the total indentation depth v, unloading stiffness S and indentation load F.

The choice of equation for the evaluation of the fracture toughness from exper-

imental results depends on the crack system, which is discussed in the following section. The equation typically incorporates the elastic modulus, the hardness and the measured crack-length (see Anstis et al. 1981). When plasticity is present, the growth of cracks in the unloading stage may have to be taken into account (see Anstis et al. 1981).

5.2 Deformation behavior of gallium nitride stacks loaded by nanoindentation

The deformation behavior of GaN was extensively studied by application of nanoindentation in the literature. Three deformation stages, characteristic for single crystal materials, which are also observed for nanoindentation in GaN, are distinguished (see Ratschinski et al. 2011). In the elastic stage of the nanoindentation process, the deformations are fully reversible. In the elasto-plasic stage, dislocations are generated and deformation movement occurs. In a third stage, crack seeds form, macroscopic cracks emerge and emmanate from the indenter outwards, which can lead to the destruction of the specimen (see Ratschinski et al. 2011). The arrangement of dislocations and crack pattern, are determined by the type of stack, the type of indenter, the orientation of the indenter on the specimen and the applied load, as results of Vickers and Berkovich indentations in the literature reveal. The different stages are typically observed at different indentation loads.

5.2.1 Elasto-plastic deformation

At all indentation loads, indentation induced dislocations emerge which overlay the irregular in-grown dislocation network (see Ratschinski et al. 2011). The dislocations typically consist of 3 groups of 2 branches arranged in the shape of a rosetta and can be made visible by cathodoluminescence (CL) imaging. With increasing load, this rosetta grows in its size, but doesn't change its characteristic shape. An example of the dislocation pattern originating from Vickers indentation aligned with the crystal symmetry planes of free standing (0001) GaN is shown in figure 5.2 (top, center) (see Ratschinski et al. 2011). Comparing the dislocation lines, which appear dark in the figure, with the directions of the crystal symmetry of GaN, the dislocation lines are aligned with the set of equivalent symmetric crystal directions $\langle 2\bar{1}\bar{1}0\rangle$ and therefore of the intersection line between the equivalent symmetric $\{01\bar{1}0\}$ faces with the (0001) surface as shown in figure 5.2 (top,right). Dislocation lines are also shown by means of a CL image of the $(01\bar{1}0)$ crosssection after indentation as shown in figure 5.2 (top, left). The black region contains a high number of dislocation in which no single dislocations can be identified. Beneath this region, three types of dislocation lines that are characterized by their orientation, are visible. The inclined lines (1), the vertical lines (2) and the horizontal lines (3) are aligned with the intersection of the $\{\bar{1}\bar{1}22\}$, $\{01\bar{1}0\}$ and $\{0001\}$ slip-planes with the $(01\bar{1}0)$ cross-section, as shown in figure 5.2 (bottom, right).

From the same indentations, slip lines are shown in figure 5.2 (bottom, center) at the surface of the indent. The slip lines are of the type $\langle 01\bar{1}0\rangle$ and therefore of

the intersection of $\{\overline{1}\overline{1}22\}$ faces or with the (0001) surface. Extensive studies on the orientation of the indenter have been carried out in the literature which conclude, that the orientation of the indenter has only minor influence on the dislocation pattern, which therefore follows the characteristic planes and directions of the crystal structure independent of the indenter orientation (see Ratschinski et al. 2011, 2012, 2013). Slip-lines on the surface occur independent of the indenter orientation in three characteristic directions, although the interconnection of slip-lines in the indent area differs. The characteristic pattern of dislocations and slip-lines was also observed for epitaxially grown (0001) GaN on sapphire (see Ratschinski et al. 2011). From a comparison of foursided Vickers indentations with three-sided Berkovich indentations, it is also obvious that the indenter geometry has only minor influence on the pattern of dislocations and slip-lines (see Ratschinski et al. 2013). Only for the use of a cube-corner indenter, slip-lines are observed outside in the pile-up in addition to the characteristic slip-lines in the area of the indent (see Ratschinski et al. 2013).

The formation and movement of dislocations is expected to consume considerable amount of the work transferred from the indenter to the GaN stack. The dissipated energy

$$W_{\mathrm{diss}} = \int_{v_0}^{v_{\mathrm{max}}} F(v)\mathrm{d}v - \int_{v_{\mathrm{max}}}^{v_{\mathrm{end}}} F(v)\mathrm{d}v \tag{5.2}$$

is evaluated from the slope of the indenter load F with respect to the displacement v, over the range of indentation depth from the contact point v_0 to the maximum indentation depth v_{max} and back to the indentation depth v_{end} at which no more

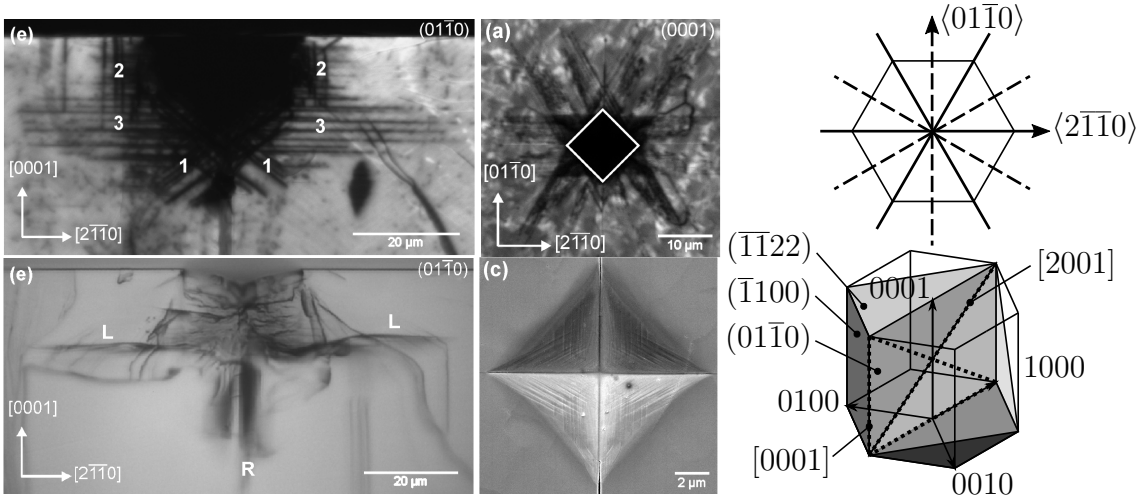

Figure 5.2: Crosssection (top, left) and top-view (top, center) of cathodoluminescence (CL) images of dislocations, cross-section transmission optical microscopy (OM) image (bottom, left) and secondary electron contrast (SE) image (bottom, center) of slip-lines in GaN from Vickers indentation (Reprinted from Ratschinski et al. 2011, fig. 2, 3). Slip-planes and characteristic directions of the crystal structure of GaN (right).

elastic reaction from the specimen is observed ($F = 0$N). An example for a study of the deformation mechanisms in a thin film of GaN loaded by nanoindentation from the literature is provided in figure 5.3 (see Tsai et al. 2008). The large area between the loading and unloading curve provides the dissipated energy, which is obvious to take a significant amount compared to the loading work.

In addition to the previously discussed deformation mechanisms, phase transformation may occur during nanoindentation (see Tsai et al. 2008). The Wurtzite structure of GaN transforms into the rocksalt structure at a hydrostatic stress of more than 50GPa (see Tsai et al. 2008). During nanoindentation testing, neither phase transformation nor cracking occurs at least up to a load of 300mN in Berkovich indentations of a 2μm thick GaN film deposited on sapphire (see Tsai et al. 2008), although in other experiments (see Weyher et al. 2001), phase transformations from

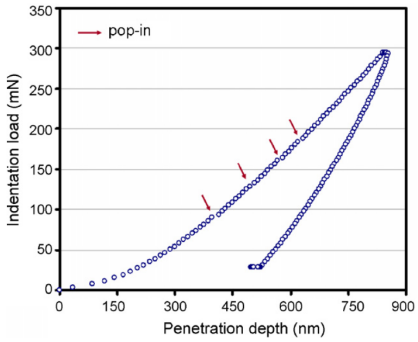

Figure 5.3: Typical load displacement curve from Berkovich nanoindentation (Reprinted from Tsai et al. 2008, fig. 2)

Vickers indentation at loads up to 2N are observed.

5.2.2 Crack nucleation and propagation

At higher indentation loads, crack seeds are formed, from which macroscopic cracks originate and emmanate outwards from the indenter. Pile-up of dislocations on one ore more glide-planes are prominent examples of models for the formation of crack seeds (see Ratschinski et al. 2011). Several basic crack types are distinguished and are shown in figure 5.4 by example of a Vickers indentation (see Cook & Pharr 1990). These include radial cracks, which emmanate radially outwards from the indenter edges and secondary radial cracks in directions different from the edges. Radial cracks that reach from the center of the indent to the edges of the indent are referred to as median cracks and those reaching further outwards penny-cracks. Lateral cracks have orientations close to that of the surface. Shallow cracks show also orientation close to that of the surface and grow from the area of the indent into the material and reach back to the surface in some distance to the indent.

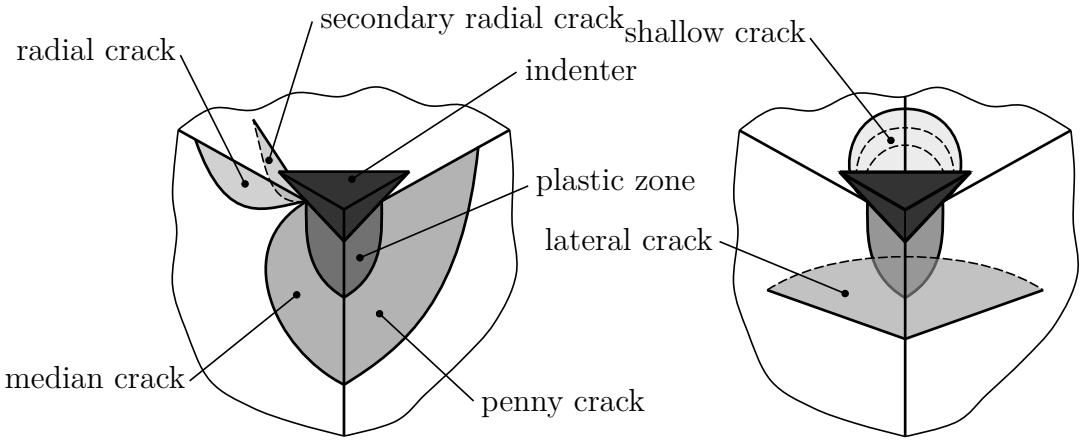

Figure 5.4: Illustration of crack types (see Cook & Pharr 1990)

Compared to the dislocation arrangement, the crack system is predominantly determined by the orientation and symmetry of the indenter, and therefore by the strain-field, generated in conjunction with the mechanical and crystallographic properties of the material (see Ratschinski et al. 2011). The crack types emerging in free-standing (0001) GaN, epitaxially grown (0001) GaN on sapphire and bulk GaN are mainly lateral and radial (by means of a category containing median and penny-wise cracks as well) (see Ratschinski et al. 2011, 2012, 2013). Radial and lateral cracks can form at different loads at the different corners of the indenter, even if the Berkovich indenter is used, showing the same symmetry as the crystal structure of the indented material. This might be subjected to the in-grown dislocation density (see Ratschinski et al. 2011) but also on the type of indenter and orientation. A comparison of the crack-pattern from Berkovich, cube-corner and Vickers indentation in free-standing (0001) GaN can also be found in the literature (see Ratschinski et al. 2013). The indent and crack pattern from Berkovich and Vickers indentation are

shown in figure 5.5. Radial cracks emmanate from all edges of a Berkovich indenter with its edges being aligned with the $\{01\bar{1}0\}$ cleavage planes of GaN, whereas for the indenter being aligned with the $\{2\bar{1}\bar{1}0\}$ planes, only one crack emerges, showing the strong influence of the orientation, subjected to inhomogenities from in-grown dislocations in the material or asymmetries in the indentation process. A similar influence of indenter orientation is shown for Vickers indentation in figure 5.5. In addition to radial cracks in Berkovich indentations, Vickers indentations show also lateral cracks. The type of crack-system, namely lateral or radial types of cracks, are related to an equivalent cone angle (see Ratschinski et al. 2013). Berkovich indentations show a high equivalent cone angle and therefore lead to radial cracks, whereas the Vickers indenter shows a low equivalent cone angle, leading also to lateral cracks. A comparision of the fracture behavior of free-standing GaN and GaN on sapphire substrate, was carried out by Ratschinski et al. (2011), who showed, that lateral cracks can be predominant in Vickers indentation in GaN on sapphire substrate and radial cracks only emerge at highest loads. This behavior is subjected to the mismatch in the cleavage planes of GaN. The loads required for considerable crack growth vary from 0.3N to 2.3N in the literature (see Ratschinski et al. 2011, 2012, 2013). The load depends on the type of indenter used, the structure of the specimen (e.g. free-standing GaN or GaN on a foreign substrate) and the crack system, with radial cracks emerging at lower loads than lateral cracks in tests which exhibit both crack types.

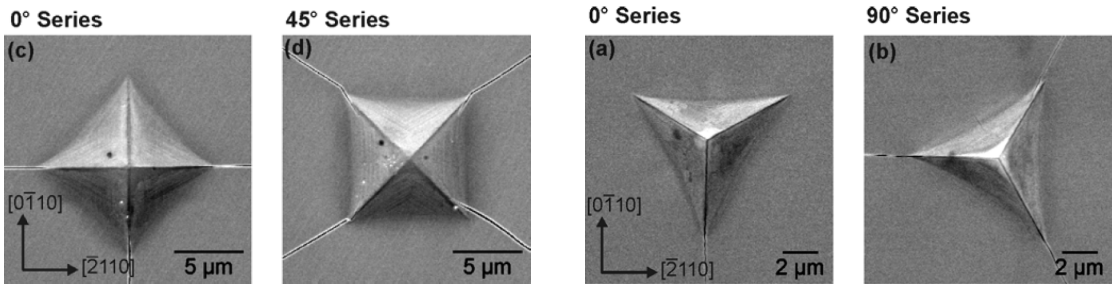

Figure 5.5: SE image of cracks in Vickers (left) and Berkovich (right) indentations in free-standing (0001) GaN (Reprinted from Ratschinski et al. 2013, fig. 2).

5.3 Summary

In this chapter, the indentation procedure has been introduced, which the prototype problem follows. In addition, important aspects of fracture in GaN stacks loaded by nanoindentation are discussed. The strong contribution of dissipative mechanisms such as dislocation generation and slip (plasticity) is evident. The confinement of dislocations and slip to characteristc glide-planes can be utilized in the formulation of a model for plasticity or for considerations regarding the onset of plasticity from stresses in these planes. The predominant crack modes proved to be lateral and radial cracks of which only radial cracks emerge in Berkovich indentations allowing some simplifications in the modelling approach by confining crack growth in radial planes. Typical loads at which cracks emerge are in the range of 0.2N and 2.2N

6 Modelling approach

The findings from the literature review on fracture in GaN stacks in chapter 3 and their fracture mechanical characterization by nanoindentation testing in chapter 5 are used to develop a modelling approach for the prediction of crack initiation and propagation based on the concepts of continuum mechanics and LEFM as introduced in chapter 4, and by application of the FEM method. Although the focus in the development of the modelling approach is to predict crack initiation and propagation in stacks loaded by nanoindentation, the major parts of it may be also valid for many other load cases including thermal loads, loads during processing and inertia during handling operations. Attributes of the stacks and mechanisms that can have significant influence on crack growth in GaN stacks based on the literature study are summarized in section 6.1. From the full complexity, that a modelling approach has to deal with in general stacks, the focus in this thesis lies only on a subset of it. Therefore some additional approximations and assumptions are made in section 6.2. In section 6.3 fracture mechanical methods that incorporate the concepts of LEFM in a FEM environment are assessed for their suitability to predict crack initiation

and propagation in a model based on the approximations and assumptions. The cohesive zone method, which is the method assumed to be best suited to predict crack propagation in the posed problem, is discussed in detail in section 6.4. To complete the modelling approach, reliable thermoelastic and fracture mechanical material properties are required, which are assessed from a literature study in section 6.5 for Si, GaN, AlN and AlGaN. The application of the modelling approach with highly brittle materials provide some challenges, that are addressed by preliminary studies, carried out in section 6.6.

6.1 Requirements to a general modelling approach

The modelling approach is based on the assumption that initiation and propagation of macroscopic cracks can be described consistently and to the required level of accuracy by continuum mechanics and application of LEFM. The theory of LEFM is assumed to be applicable, because the group III-nitrides and substrate materials are crystal materials, and theory of LEFM is assumed to be valid for many crystal materials (see Gross & Seelig 2016). In addition, from phenomenological observations, materials show sharp crack-fronts during nanoindentation, as it is characteristic for materials with a localization behavior, to which LEFM applies. Further arguments include the high fracture strength and low fracture toughness of the materials. The continuum mechanics approach is valid if measures such as stresses and strains are related to length scales that are far greater than those at which these measures

change significantly. This is the case, if the length of the process zone, which is assumed to be the smallest length scale of the model relevant to the fracture process, is larger than the length scales of the defect density and therefore even larger than the lattice constants of the crystal structure within each material.

Within the theories of the modelling approach, the relevant mechanisms and relevant level of abstraction have to be identified which significantly influence the position of crack initiation and path of crack propagation on the macroscopic scale. Based on the cohesive zone model, a crack initiates in a region where the stresses reach the fracture strength and in addition, enough deformation energy is provided by the surrounding area to fully separate the material. The relevant mechanisms to be included in the modelling approach are therefore those, having significant influence on the distribution of the components of the deformation tensor as well as on the distribution and anisotropy of the thermoelastic and fracture mechanical material properties. The following mechanisms and attributes identified from the literature review may therefore have significant influence and are required to be taken into account:

- stacking sequence
- layer thicknesses
- thermo–elastic and fracture–mechanical properties as well as anisotropy of the materials
- interface fracture

- indenter shape, load and orientation
- friction between indenter and surface
- plasticity
- phase transformation

Based on the literature study and confinement to a special case studied in this thesis, some of these mechanisms and attributes are neglected in the next section.

6.2 Approximations and assumptions

Many transistor stacks of practical use are deposited on the hexagonal plane of e.g. sapphire or the (111) plane of silicon. Inherent to epitaxial growth, deposited materials share crystal orientation with the material of the subsequent layers. Therefore it is not a strong confinement for practically relevant applications to assume only stacks of layers with their hexagonal [0001] axis, or the [111] axis in the case of silicon, aligned with the direction perpendicular to the wafer surface and in addition, aligning the main cleavage planes of the materials.

From the set of possible indenters, the modelling approach is confined to the use of a Berkovich indenter, which shows three-fold symmetry similar to the considered stacks and therefore considerably reduces the complexity of the problem. Berkovich indentation has many advantages for practical applications, as discussed in section 5.1, and the assumption therefore preserves validity for many practical applications.

The indenter edges are assumed to be aligned with the main vertical cleavage planes, inducing highest crack loads of mode I in planes of lowest resistance to fracture of mode I. Therefore, radial cracks are assumed to initiate and propagate only in the main cleavage planes, as it is observed in the literature for Berkovich nanoindentation in free-standing GaN, as discussed in section 5.2. As the modelling approach shall not only account for single layers, but stacks of different materials, the initiation of lateral cracks and interface cracks due to induced shear stresses must be critically assessed. Motivated by the wide cone angle of the Berkovich indenter, the influence of friction is neglected.

Each layer is assumed to be homogeneous, making the constitutive law and fracture mechanical material properties indenpendent of the position in each leayer. The influence of inhomogenity of in-grown dislocations from which screw dislocations are especially present close to the interface, as discussed in section 5.2, is assumed to be small. A perfect bond is assumed at the interfaces between the layers.

The deformation behavior of each material is assumed to be elastic. This assumption is contradictory to the strong plastic deformation observed during nanoindentation in the literature, as discussed in section 5.2, but chosen to separate the influence of elastic deformation and stack structure from that of plasticity. In addition a suitable plasticity model is still subject of ongoing research and exceeds the scope of this thesis. Because cracks emmanate from indenter edges outwards, whereas plasticity is confined to the region under the indenter, as discussed in section 5.2, a description of crack propagation by means of LEFM is assumed to be still valid, although it

cannot be excluded, that crack initiation is driven by interaction of slip-planes at the boundary of an elasto-plastic region, where the fracture strength or toughness may be reduced, compared to regions without plasticity. Because the radial extension of cracks are assumed to be far greater than the indent, even the shape of the cracks may be similar, although the strong dissipative effect of plasticity may raise the load at which cracks initiate, change the functional relation between load and crack length and may have an important stabilizing effect on crack growth. Although materials have a hexagonal or cubic crystal structure, which lead to anisotropic deformation behavior, an isotropic constitutive law is used for each material to separate the influence of anisotropy from that of the influence of the constraint at the interfaces. The influence on the crack growth from anisotropy in the hexagonal plane is assumed to be small because indenter, material and crack planes share the three-fold symmetry and are aligned. The distribution of stress and strain components is therefore assumed to also have three-fold symmetry which preserves the directions in which radial cracks propagate. Anyhow, the magnitude of the stresses in these directions may be changed according to the anisotropy. Differences in the elastic tensor components relating wafer-plane strain components to stress components perpendicular to the wafer plane and vice versa, compared to wafer-plane components are also assumed to be of minor importance. Anisotropy can be subsequently incorporated in the modelling approach while preserving the alignment of the indenter edges with the cleavage planes without further adaption of the modelling approach. On the basis of the discussed approximations and assumptions, the problem shows three-fold

symmetry. In the next section, the fracture mechanical methods are assessed for their suitability to predict initiation and propagation of cracks on the basis of the assumptions and approximations.

6.3 Assessment of numerical fracture mechanical methods

Several methods are available that incorporate the concepts of LEFM in a FEM-environment. They are assessed for their ability to predict crack growth in GaN stacks based on the approximations and assumptions discussed in the previous section. A suitable method is required to predict the location of crack initiation within the cleavage planes and the path of crack propagation of radial cracks with crack fronts of arbitrary shape. In addition, it should be able to account also for lateral cracks at the interfaces. The occurence of plasticity in the model should be considered. The limitations of different numerical fracture mechanical methods in fulfilling the requirements are assessed and the requirements to the mesh and the computational effort are qualitatively discussed. From the assessment, the best suited method is selected and incorporated in the modelling approach.

The concept of crack propagation in continuum mechanics is a change of the boundary value problem by continuous generation of new surface, the crack faces, and change of the boundary conditions on that surface. Based on this concept, the implementation of crack propagation in a FEM environment is the stepwise adaption

of the spatial discretization and change in boundary conditions (see Kuna 2008). This stepwise adaption of the spatial discretization can be achieved by the separation of nodes, the modification of elements the incorporation of crack tip elements shifted with the crack tip either as part of the mesh or in a submodelling approach, adaptive remeshing, and the incorporation of elements resembling a material model in their Ansatz (see Kuna 2008). The modification of elements can be subclassfied in element splitting, element elimination and stiffness degradation of elements. A common material model used to model crack initiation and propagation and which is also capable to model crack nucleation, is the cohesive zone model. It models the material separation in the thin and small separation zone at the crack tip after reaching the fracture strength by a continuous reduction of transferrable stresses until a critical separation is reached.

From the listed methods, the cohesive zone method not only fulfills the requirements but is also the most promising of the methods due to several reasons. Crack propagation techniques such as the node separation, element splitting and element elimination represent crack growth by a discontinuous process. All three methods change the node and element configuration when a fracture criterion is fulfilled. In the case of node splitting, nodes are duplicated and forces between nodes set to zero. The element splitting technique is also based on a duplication of nodes, but in addition changes the number of elements, as one element is split in two elements. With element elimination, the highest loaded elements are deleted and therefore their contribution to the stiffness of the body is eliminated. Although the discon-

tinuous character is often mitigated by gradually decreasing nodal forces or element stiffness, this process doesn't resemble the behavior of a local material model. The dissipative mechanism of fracture is also neglected. In the case of element elimination, the mass conservation is violated as well. In contrast to these methods, the cohesive zone method not only represents crack propagation as a continuous process of separation, but directly resembles a local material mechanical fracture model in its element Ansatz. Mass is conserved, because elements have zero thickness in their initial state and energy is conserved, while also resembling the dissipative mechanism of fracture. Another advantage of methods that resemble a material mechanical fracture model in their Ansatz is that no additional methods are required to evaluate the fracture criterion at the crack tip and not even the representation of an existing crack-tip by the mesh is required. Such methods therefore circumvent to generate meshes for all possible initial crack configurations and instead enable the initiation and propagation of cracks anywhere, where these kind of elements are incorporated in the overall mesh. The cohesive zone method also circumvents the necessity of excessive remeshing, as it is the case for methods, which require to resolve the stress and displacement solution close to the crack tip for an improvement in accuracy of the crack initiation analysis, although a suitable mesh resolution must be achieved to preserve convergence and numerical stability.

Another important aspect is the spatial discretization of the crack path by the different methods. Techniques such as node separation, element separation, element elimination and also cohesive zone elements arranged at the boundaries of standard

finite elements require to know the crack path beforehand, to align element edges with the crack path. In contrast to these methods, element splitting method or extended finite elements, which resemble the discontinuity in their Ansatz based on cohesive zone law can resemble the actual path of a crack without dependence on the alignment of the standard elements. As the planes of crack growth are known beforehand, this aspect is not of importance under the current assumptions, but may be of importance when allowing also radial cracks not aligned with the symmetry planes or shallow cracks. Within the plane of crack propagation, the spatial discretization plays still an important role, because the results from methods such as the node separation, element elimination, element stiffness degradation depend on the size of elements which have to be carefully chosen. The mesh affects the crack increment when using the node separation, element splitting method, element elimination and element stiffness degradation. In addition it affects the accuracy by which the fracture criterion can be evaluated, when additional crack initiation methods are required, as it is the case for all methods except the cohesive zone method or elements based on damage mechanics.

From the above considerations, interface elements or extended finite elements, which resemble the cohesive zone model in their Ansatz not only fulfill the requirements, but are the most promising. In the next section, the cohesive zone method is introduced in more detail and important aspects when modelling indentation fracture are discussed incorporating findings from similar modelling approaches in the literature.

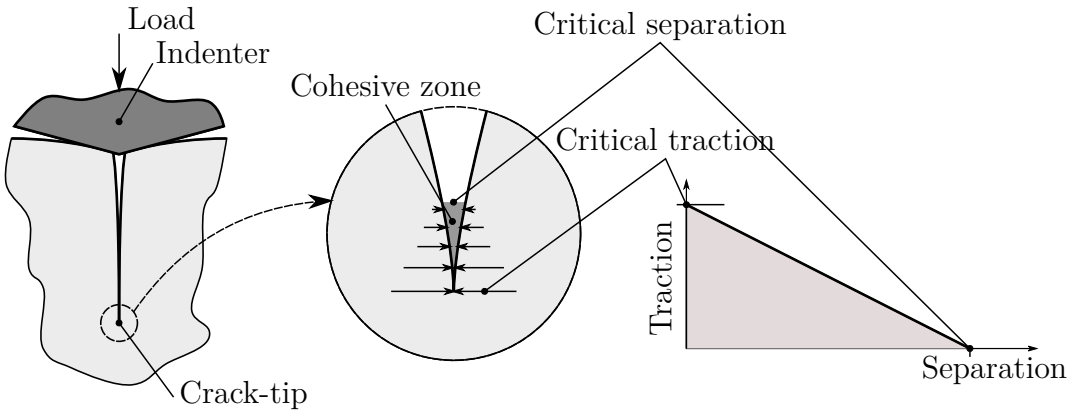

Figure 6.1: Indentation model with pre-existing crack and crack tip (left), cohesive zone (center) and traction–separation law (right).

6.4 Cohesive zone method

The method, that seems most suited for the posed problem at a reasonable amount of computational time and low additional programming effort is the cohesive zone method (CZM). The method is based on the cohesive zone model. In section 4, the process zone was introduced in which material separation is assumed to take place close to the crack tip. The cohesive zone model assumes that the process zone is a thin strip-shaped region attached to the crack tip and confined by an assumed extension of the crack-flanges, as shown in figure 6.1 (left, center). The separation process in this zone is assumed to be determined by a cohesive law. The zone is therefore referred to as a cohesive zone, a term first introduced by Barenblatt (1962), who developed the cohesive zone model to describe brittle fracture in a linear elastic body (see Kuna 2008, Gross & Seelig 2016). The cohesive law describes separation as a continuous damage process based on the relation of stresses and separations, which

is referred to as a traction–separation law. The shape of the traction–separation law is determined by the mechanisms that lead to material separation in the cohesive zone. An exponential traction–separation law is often used for materials, in which pure brittle fracture is the relevant fracture mechanism, because the model resembles the potential of atomic bonding (see Barenblatt 1962). It is mentioned here, because tractions are expressed in closed form as a continuous differentiable function of the separations. Due to this simplicity it is used to study numerical stability mechanism in the literature (see Gao & Bower 2004). A linear traction–separation law is widely used to model fracture in elasto-plastic materials, where it resembles the reduction of transferrable tractions due to void formation, but also in brittle materials, where it resembles the reduction of transferrable tractions due to the formation of micro cracks. In this thesis, the separation process of GaN and its related alloys is also described by a linear traction–separation law, which is shown in figure 6.1 (right). It is based on the assumption, that when a stress criterion is fulfilled, material starts to separate. For mode I loading of a crack in an isotropic material, the stress criterion

$$\sigma_{\mathrm{n}} = \sigma_{\mathrm{c}} \tag{6.1}$$

is fulfilled when the stress normal to the crack-extension direction σ_{n} equals the fracture strength σ_{c}. When the separation increases, transferrable tractions decrease

acoording to the traction–separation law, expressed for the normal traction by

$$\sigma_{\mathrm{n}} = \begin{cases} \sigma_{\mathrm{c}} \cdot (1 - \beta) & \text{if } \beta \leqslant 1 \\ 0 & \text{else} \end{cases} \tag{6.2}$$

which defines a linear decrease of the traction acoording to a damage variable β. The damage variable for mode I loading of a crack in an isotropic material

$$\beta = \begin{cases} \delta_{\mathrm{n}}/\delta_{\mathrm{nc}} & \text{if } \delta_{\mathrm{n}} \leqslant \delta_{\mathrm{nc}} \\ 0 & \text{else} \end{cases} \tag{6.3}$$

depends solely on the ratio of the normal separation to the critical normal separation. When a critical separation δ_{nc} is exceeded, no more stresses can be transferred and the interface is fully separated. When the material separates, work is dissipated. This work is calculated by integrating the traction with respect to the separation. For full separation of the material, this work is expressed by

$$\mathcal{G}_{\mathrm{c}} = \int_0^{\delta_{\mathrm{nc}}} \sigma_{\mathrm{n}}(\delta_{\mathrm{n}}) \mathrm{d}\delta_{\mathrm{n}} \tag{6.4}$$

and is identified as the critical energy release rate $\mathcal{G}_{\mathrm{c}}$ (see Kuna 2008). The behavior of the model is therefore determined by the fracture strength and the fracture toughness, from which the length of the critical separation can be evaluated. Compared to a crack described by the LEFM, the cohesive zone model introduces an

additional length scale either expressed by the critical separation length or the cohesive zone length, which is determined by the critical separation, the structure, the load, the fracture toughness and the fracture strength. The cohesive zone length in a homogeneous isotropic center crack tension (CCT) specimen

$$d = \frac{\pi}{8}\left(\frac{K_{\mathrm{Ic}}}{\sigma_{\mathrm{c}}}\right)^2 \quad \text{(see Johanns et al. 2014, eq. (6))} \tag{6.5}$$

can be evaluated for a given fracture strength and fracture toughness. From this relation it is also obvious, that the length of the cohesive zone increases with the fracture toughness and decreases with increasing fracture strength (see Johanns et al. 2014).

To make use of the concept of the cohesive zone model in a finite element formulation, interface finite elements are widely used with the traction–separation law as the governing material law. Simulating crack growth with such elements is referred to as the cohesive zone method. The simulation of crack nucleation and propagation in an unflawed structure by use of this method provides some challenges that have to be accounted for. At crack nucleation, the reduction of transferrable stresses can lead to a snap-back instability, which leads to dynamic nucleation of a crack of finite length in conjunction with the propagation of deformation waves from the crack-tip in the material, if further crack propagation is stable. The same phenomenon can occur for the initiation of an existing crack with the initiation of the crack and propagation by a crack increment of finite length. These phenomena are related to the

properties of the material and the structure as well as the load case and not to the discretization. The discretized FEM model resembles the same instability mechanism, although the discretization can modify the instability behavior, as observed e.g. by insufficient spatial resolution of the process zone, which can destabilize crack growth and even lead to numerical artefacts. In this thesis, the interest lies on the magnitude of the load required for crack nucleation, the required load for crack nucleation and the stability and path of further crack propagation. The interest lies not on the dynamic behavior between a possible instability at crack nucleation and the regaining of stability after a crack is formed. The modelling approach is therefore not required to follow and resolve this unstable path in the regime of small crack lengths. Modifications are therefore applied to the cohesive zone model to keep simulation time low. To keep predictions of the crack propagation consistent, the influence of these modifications has to be small, which has to be verified from results. Suitable estimates before simulation help to find appropriate parameters for the modifications.

Unstable crack propagation can be stabilized, when the traction–separation law is modified and when viscous forces are induced. In addition, a sufficient spatial resolution of the mesh in the process zone and around it, prevents destabilization from inducing artificial stiffness of the mesh and artificial brittleness in the process zone. A common modification of the linear traction–separation law as shown in figure 6.2 (left) is the introduction of an initial stiffness k, which leads to a bilinear traction–separation law, as shown in figure 6.2 (center). This modification is nec-

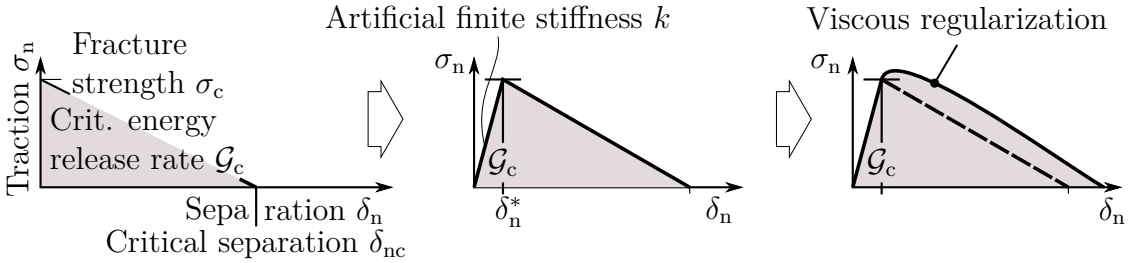

Figure 6.2: Traction–separation law motivated by material mechanics (left) with added artificial stiffness (center) and with viscous regularization (right).

essary, because with an infinite initial stiffness, no convergence can be found due to the instability mechanism. In the literature it is observed, that with decreasing stiffness in an exponential traction separation law, crack initiation and propagation can be stabilized in a simple 2D problem similar to figure 6.3 (see e.g. Gao & Bower 2004). As a downgrade, the induced artificial compliance enables the cohesive zone to separate elastically up to the maximum artificial elastic separation δ_n^* at which the fracture strength is reached and the damage regime starts, and therefore to store elastic energy in the zone. To prevent spurious results, e.g. meaningless results, influenced by the discretization and not by physical mechanisms, the initial stiffness has to be carefully chosen. Estimates for a suitable choice of the initial stiffness are typically based on the comparison of the compliance of the structure compared to that of the cohesive zone of a uniaxially loaded body with an interface, as shown in figure 6.3. The ratio of the stiffness of the body compared to that of the interface

$$\Lambda = \frac{E_{\text{eff}}/l_0}{k} = \frac{E_{\text{eff}}\delta_n^*}{2\sigma_c l_0} \ll 1 \tag{6.6}$$

has to be minimized and depends not only on material properties, but also on the

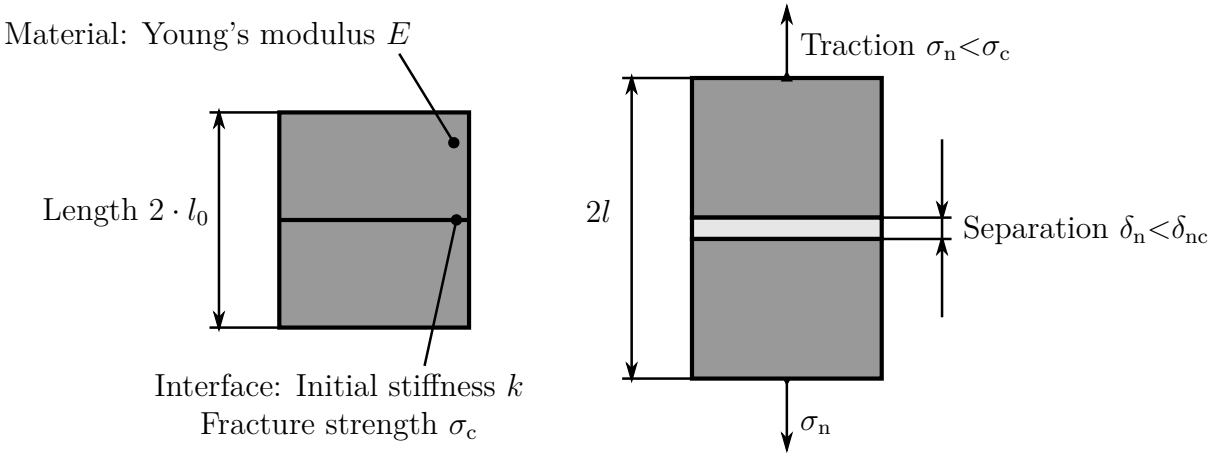

Figure 6.3: 2D Interface model in deformed (left) and undeformed (right) state.

length of the body. To get estimates for a more complex problem from equation (6.6), a length scale l_0 and effective modulus have to be chosen in a way, that the simple 2D problem resembles a similar interface behavior as the more complex problem. Considering the nucleation of three radial cracks close to the indenter tip, equation (6.6) yields

$$l_0 = d\pi/6 \tag{6.7}$$

with the length scale chosen as the arc length intersecting the midpoints of an assumed process zone. Another estimate is gained from the requirement, that the ratio of the critical separation δ_{nc} compared to that at maximum stress in the cohesive zone σ_c

$$\alpha = \frac{\delta^*}{\delta_c} = \frac{\delta^* \sigma_c}{2\mathcal{G}_c} \ll 1 \tag{6.8}$$

has to be minimized. Expression (6.8) is equivalent to the the requirement, that the elastic energy stored in the cohesive zone with respect to the consumption of

energy due to damage in the process zone is small. Although the introduction of an artificial stiffness may improve convergence in many cases, the material properties may confine the possible choice of artificial stiffness in a way, that stabilization is still not possible. An estimate for the onset of instability is given by

$$\frac{{\sigma_\mathrm{c}}^2 l_0}{\mathcal{G}_\mathrm{c} E_\mathrm{eff}} > \frac{1}{2} \quad \text{(see Gao \& Bower 2004)} \tag{6.9}$$

for the 2D problem with an exponential traction–separation law. The stability condition (6.9) may be difficult to fulfill with the expected high fracture strength and low fracture toughness of GaN.

Because the nucleation and propagation of cracks may not be stabilized with the introduced initial stiffness, viscous forces have to be induced. These viscous forces are widely related to the change rate of the separations. These forces dissipate energy in the cohesive zone and therefore stabilize crack nucleation and crack propagation. In the case of the simple 2D problem shown in figure 6.3, the stabilization is always successful if a suitable time step is chosen (see Gao & Bower 2004). Several implementations of such viscous forces exist. The implementation used in this thesis is based on the formulation

$$\sigma_\mathrm{n} = \sigma'_\mathrm{n} + \eta \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\delta_\mathrm{n}}{\delta^*_\mathrm{n}}\right) \quad \text{(see Gao \& Bower 2004)} \tag{6.10}$$

by which the actual traction of the cohesive zone element σ_n is calculated from the tractions evaluated from the original traction-separation formulation σ'_n which are

superimposed by viscous tractions, that depend on the rate of change of separations $\delta_{\mathrm{n}}/\delta_{\mathrm{n}}^{*}$ and scale with the viscosity η. By inducing a viscosity, a characteristic time scale is added to the problem, the stress relaxation time t_0, which is the time, necessary for the stresses to relax after crack nucleation, when keeping the load constant. In the case of the simple 2D problem, the relaxation time is given by

$$t_0 = \frac{\eta l_0}{E_{\mathrm{eff}}\delta_{\mathrm{nc}}} \quad \text{(see Gao \& Bower 2004).} \tag{6.11}$$

Although crack nucleation and propagation may be stabilized by any viscosity, the numerical solution algorithm, such as the Newton–Raphson algorithm require to stay within a radius of convergence. In the case of the simple 2D problem, the Newton–Raphson algorithm converges if the time increment Δt is lower than the relaxation time

$$\Delta t \ll t_0 \quad \text{(see Gao \& Bower 2004).} \tag{6.12}$$

When viscous forces are induced, the actual behavior of the cohesive zone deviates from that of the traction–separation law, as shown in figure 6.2 (right). The cohesive zone may transfer higher stresses than the fracture strength and damage at a higher energy than the fracture toughness. This behavior is referred to as overloading in the literature. Overloading can slow down the crack growth, which therefore lags behind the actual crack growth during proceeding increase of the load. To minimize these artificial effects, the time of unstable crack nucleation or propagation has to be far smaller than any other characteristic time scale, such as the loading time

step. An estimate for a sufficiently small viscosity is again provided for the simple 2D problem by

$$\Theta = \frac{\eta l_0 V}{E_{\text{eff}} {\delta_{\text{nc}}}^2} \ll 1 \quad \text{(see Gao \& Bower 2004)} \tag{6.13}$$

with the characteristic time scale of the model as the ratio of velocity of loading V with respect to the critical separation δ_{nc} For more complex problems, it may be difficult to estimate the characteristic length and time scales and therefore verification is necessary from the results of the simulation. A convergence study with respect to the viscosity may be very costly in terms of computation time. To circumvent this problem, viscous forces can be estimated from a comparison of the work of the cohesive zone

$$W_{\text{coh}} = \int_0^{\delta_{\text{n}}} \sigma_{\text{n}} \mathrm{d}\delta_{\text{n}} \tag{6.14}$$

with the theoretical work of the cohesive zone

$$\int_0^{\delta_{\text{n}}} \sigma'_{\text{n}}(\delta_{\text{n}}) \mathrm{d}\delta_{\text{n}} \tag{6.15}$$

based on the evaluation of the normal traction σ'_{n} from the separation result. This comparison should yield

$$\overline{\Theta} = \frac{W_{\text{coh}} - W'_{\text{coh}}}{W_{\text{coh}}} \ll 1 \tag{6.16}$$

along the interface.

The high fracture strength and low fracture toughness of GaN require a very low viscosity as indicated by the criterion (6.13). Therefore a small time step is

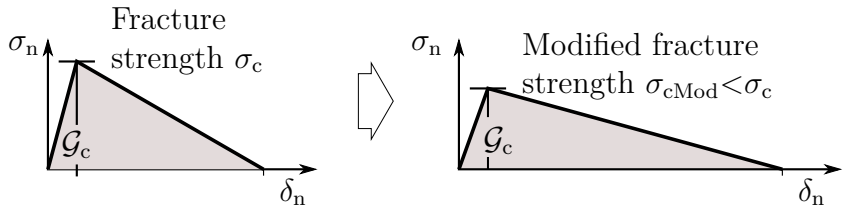

Figure 6.4: Bilinear traction–separation law with original fracture strength (left) and with reduced fracture strength (right) at the same critical energy release rate.

required as indicated by the criterion (6.12) for the relaxation time (6.11). As each time step requires to solve the nonlinear system of equations, the computational time can be very high. In addition, the cohesive zone should contain at least 2-3 elements (see Camanho & Hallett 2015). The resulting high number of elements and therefore degrees of freedom (DOFs) increase the simulation time of each time step and also the requirement in main memory during the simulation steps and hard-disk drive to gather the simulation results. The requirements in main memory, hard-disk drive may therefore exceed the ability of nowadays computer systems and the required simulation time may exceed any reasonable bounds. Another workaround is therefore necessary to circumvent the computational limitations. For crack propagation problems, that are determined by the fracture toughness and not by the fracture strength results should not depend so much on the element strength, but in leading order on the dissipated energy of the element (see Turon et al. 2007). This consideration motivates to reduce the fracture strength while keeping the critical energy release rate constant, as shown in figure 6.4 (right). With the reduction of the fracture strength, less viscosity may be necessary to stabilize

the propagation of cracks and in addition, a larger viscosity and larger time step can be chosen because the relaxation time is increased. If the artificial compliance is not limited by the compliance of the surrounding material, but by the ratio of artificial elastic energy to dissipated energy, the steep initial slope of the traction–separation law can be reduced as well. The application of this method requires to study convergence with respect to the used fracture strength.

With the introduced modifications of the traction–separation law and the induced viscous regularization, phenomenological parameters η, k and σ_{cMod} are added to the problem. Although physical analogies can be found, they are primarily used to gain a positive tangent stiffness matrix of the FEM model and therefore achieve convergence. Although suitable estimates for bounds for these parameters to achieve convergence while preserving consistency can be evaluated from a simple 2D problem, more complex problems, as addressed in this thesis require preliminary studies to find a suitable combination of these parameters. These preliminary studies are carried out in section 6.6. To address these preliminary studies, suitable material properties as input for the cohesive zone model and for the elastic or elasto-plastic material model are necessary and therefore discussed in the next section with a focus on the group III nitrides and the silicon substrate.

6.5 Material data

To predict crack nucleation and propagation in GaN based stacks, material data is required for the thermo–elastic, elasto–plastic and fracture mechanical properties of each layer. With the chosen assumptions, the plastic material data is neglected. Material data depends to a certain degree of accuracy on the analysis method, as each method can be based on different approximations and assumptions. Experimental testing methods can also be carried out on different length scales and on material of different flaw densities, as it is the case for heteroepitaxially grown GaN compared to bulk GaN platelets from synthesis (see Drory et al. 1996). Especially for fracture mechanical properties it is therefore necessary to use in-situ material properties to predict the behavior of stacks that are applied to a device and on the proper length scale (see Gross & Seelig 2016). As this thesis aims more on the differences in the crack behavior with and without induced residual stresses than on the actual behavior of a certain stack that is manufactured and applied to an electronic device, the material data for Si, GaN, AlN and AlGaN are taken from the literature.

In this thesis, the contribution of elastic anisotropy of each materials is neglected. The elastic behavior can therefore be quantified solely by the Young's modulus E and the Poisson ratio ν of the isotropic approximation of the actual anisotropic material behavior. The elastic properties found in the literature are evaluated from either theoretical models, experimental testing methods based on a deformation response from nano-indentation, beam bending and torsion loads or based on the wave velocity response from acoustic excitation. The reported properties differ significantly

depending on the method applied. In the case of experimental testing methods, they also depend on the the material or structure, the method is applied on and on the crystal orientation of the material in this structure. The analysis methods may simplify the material behavior to a certain degree of anisotropy or apply averaging over anisotropic directions. Similar to the focus of this thesis, the focus of the following discussion lies on material data evaluated from nano indentation tests in thin films grown on a substrate.

For silicon, the evaluated Young's modulus can vary from 130GPa to 188GPa depending on the chosen crystal direction and the Poisson ratio from 0.048 to 0.4 depending on the chosen crystal plane and it is therefore required to take the load case into consideration when choosing a suitable degree of anisotropy for which the properties are evaluated (for a summary and detailled discussion see Hopcroft et al. 2010). In this thesis, a Young's modulus of 130GPa is used and a Poisson ratio of 0.27.

The Young's modulus and Poisson ratio can also vary for the hexagonal GaN and AlN depending on the chosen crystal direction and plane, when considering the full anisotropic elastic tensor, as reported in the literature (see Reeber & Wang 2001). For GaN reported values of the Young's modulus vary from 283GPa to 337GPa and for AlN from 283GPa to 350GPa depending on the analysis method (for a summary of literature values for the Young's modulus see Amar et al. 2014). For the Young's modulus a value of 280GPa and a Poisson ratio of 0.26 are used and for AlN, a Young's modulus of 308GPa and a Poisson ratio of 0.25. Similar values can be

found in the literature, as e.g. 290 for GaN and 332 for AlN from nano indentation in thin films grown on sapphire and a Si doped GaN substrate by MOCVD (see Yu et al. 1998, Jian et al. 2010).

For application of the cohesive zone method, a critical energy release rate and a fracture strength have to be chosen. In the literature either the critical energy release rate is reported, or the critical stress intensity factor, from which a critical energy release rate can be evaluated from equation (4.9) and a suitable effective modulus from equation (4.13). The fracture toughness found in the literature is based on different experimental testing methods, such as the single cantilever bending test, the clamped beam bending test, the double cantilever compression and the pillar splitting test. For Si, a critical energy release rate of $3.49\mathrm{J/m^2}$ is used, which was evaluated from a critical energy release rate of $0.7\mathrm{MPa}\sqrt{\mathrm{m}}$ and plane strain assumption. It is similar to the value of a monocrystal disk with the vertical axis in [111] direction, evaluated from double torsion tests, as reported in the literature (see Anstis et al. 1981). It differs slightly from the value of $0.8\mathrm{MPa}\sqrt{\mathrm{m}}$, as obtained from nano indentation (see Anstis et al. 1981). For wafers with the vertical axis in [001] direction a fracture toughness of $0.8\mathrm{MPa}\sqrt{\mathrm{m}}$ is reported in the literature as a mean of the results of five different experimental testing methods (see Jaya et al. 2015). The critical energy release rate is an anisotropic property and differs significantly for the hexagonal materials GaN and AlN between the basal (0001) c-planes and the vertical $(11\bar{2}0)$ a and $(1\bar{1}00)$ m planes as reported from first principle energy calculations (Dreyer et al. 2015). The focus in this thesis lies on fracture in the

vertical planes, which show only differences of a few percent. In the literature, values of $2.2\mathrm{J/m^2}$, evaluated from a fracture toughness of $0.8\mathrm{MPa}\sqrt{\mathrm{m}}$ (see Drory et al. 1996) and a Young's modulus of 287GPa are reported for bulk GaN platelets loaded by Vickers nano indentation and $3.88\mathrm{J/m^2}$ and $4.81\mathrm{J/m^2}$ (see Dreyer et al. 2015) for m-plane GaN and AlN, evaluated from first principle surface energy calculations. To clearly separate the fracture toughness of the layers, a critical energy release rate of $0.53\mathrm{J/m^2}$ is chosen for GaN and $37.287\mathrm{J/m^2}$ for AlN. The values are evaluated from critical stress intensity factors of $0.4\mathrm{MPa}\sqrt{\mathrm{m}}$ for GaN and $3.5\mathrm{MPa}\sqrt{\mathrm{m}}$ for AlN with plane strain assumption.

Only few references were found regarding measures for the strength of a thin-film of GaN grown on a foreign substrate. Therefore, also measures from different material structures within the material group of single crystal GaN materials are taken into consideration. Due to the different defect densities and defect structures, as well as different analysis methods or load cases, the strength measures differ significantly. The influence of the defect structure gets especially evident when comparing that of a thin film of GaN grown on different substrates by heteroepitaxy, or comparing the defect structure of a thin film of GaN with that of a GaN nanowire, which can be completely pure of crystal defects. The measures of the fracture strength can also depend on the load case, considering the anisotropy of GaN and anisotropy of the defect structure. The properties found in the literature range from a bending fracture strength of 350MPa for a 1.5µm thick freestanding dry etched GaN films grown on (111) Si (see Zimmermann et al. 2006) to a tensile

strength between $4 - 7.5$GPa of GaN nanowires in a thickness range from 150nm to 450nm and a length in the range of 4µm to 14µm. In addition to measures for the fracture strength, measures for the onset of elasto–plastic deformation or measures for failure when plastic deformation is involved, are taken into consideration. When GaN is in a material structure and a stress state, that allows plasticity, normal tractions in the crack planes may be limited due to the onset of plastic deformation. Considering that slip occurs when a certain shear stress is reached, an upper limit for the fracture strength of GaN may be estimated from the relation between the maximum shear stresses and the maximum normal tractions in the crack planes. For Berkovich and Vickers nano indentation, normal tractions in the crack planes typically don't exceed a third of the yield strength when plasticity is incorporated in the model based on the von Mises equivalent stress (see Lee et al. 2012). The yield strength of bulk GaN primary shaped from a self seeding process, reveals a yield strength of 15GPa when loaded by nano indentation with a spherical indenter (see Nowak et al. 1999). Micropillar compression tests on GaN prisms grown on a sapphire substrate by MOCVD, reveal a yield stress of 7.5GPa from which a resolved critical shear stress of 2.8GPa is evaluated, as only a single slip system in a single plane is initiated (see Wheeler et al. 2013). This thesis aims rather on developing a modelling approach than testing a specific GaN stack. Therefore a rather high fracture strength of 7GPa is chosen to assess the limits of the approach. With the modifications discussed in section 6.4, the fracture strength may be reduced in subsequent steps to achieve convergence, if crack nucleation converges in the limit

Table 6.1: Summary of the material data, used in the FEM-model.

		Material		
Material property	Unit	Si	GaN	AlN
Fracture toughness for mode I K_{Ic}	(MPa$\sqrt{\mathrm{m}}$)	0.7	0.4	1.25
Young's modulus E	(GPa)	130	280	308
Poisson ratio ν	(1)	0.27	0.26	0.25
Effective modulus E_{eff} (plane strain)	(GPa)	140	300	329
Critical energy release rate $\mathcal{G}_{\mathrm{c}}$ (plane strain)	(N mm^{-1})	0.00349	0.000533	0.00476
Fracture strength σ_{c}	(GPa)	——— 7 ———		
Critical separation δ_{nc}	(nm)	0.998	0.152	1.36

of reaching the target fracture strength and crack propagation is determined by the fracture toughness rather than by the fracture strength. The material properties X of $\mathrm{Al}_x\mathrm{GaN}_{1-x}$ are evaluated from the properties of AlN and GaN based on the rule of mixture

$$X_{\mathrm{Al}_x\mathrm{GaN}_{1-x}} = x \cdot X_{\mathrm{AlN}} + (1-x) \cdot X_{\mathrm{GaN}} \tag{6.17}$$

which is also referred to as the Vegard's relationship when used to evaluate the Young's modulus in the literature (see Amar et al. 2014). The material properties are summarized in table 6.1.

6.6 Preliminary studies

Preliminary studies are carried out to address single aspects of the modelling approach. They include aspects of the numerical implementation, such as inconsistencies that can arise when interface elements are only placed in part of the crack plane and the handling of negative interface separation, as well as aspects of the choice of suitable phenomenological properties.

6.6.1 Artificial interface deformation and element distortion

The incorporation of an artificial elastic stiffness in the traction–separation law in combination with the use of an interface formulation by which negative tractions and positive tractions are treated similar, as well as the placement of interface elements in only a subset of the crack-plane, as it is done to keep the number of DOFs low, can lead to inconsistent, non-physical behavior of the elements and the model.

The incorporation of an artificial stiffness in the traction separation law, as discussed in section 6.4, allows elastic separations until the fracture strength is reached. These elastic separations may be positive or negative depending on the sign of the interface tractions if no additional modifications are made. This behavior may be tolerable when positive tractions are larger than negative tractions along the crack plane. In that case, elements damage under positive separations first, while negative separations may not exceed the elastic regime. When negative tractions along the crack plane are larger than positive tractions, elements damage under negative tractions. Large negative separations can be the consequence, which is an inconsistent

non-physical behavior of the model.

Interface elements are typically placed along the whole crack plane, confined by the model boundaries. When interface elements are placed only in a subset of the crack-plane, with constraints applied to the remaining crack-plane, inconsistent deformation behavior can be the result. This inconsistent behavior may be subjected to the discontinuity at the transition line between these areas. The result can be strong element distortion and alternation of positive and negative node displacements at the interface, which may again lead to non-physical element damage under negative tractions.

Before modifications of the modelling approach are discussed to circumvent the inconsistent behavior while keeping the modifications of the traction–separation law and the confinement of interface elements to a subset of the crack-plane, the inconsistencies that arise are demonstrated by example of the behavior of the interface in a 2D indentation problem. The problem resembles the pressure under the indenter and the stress inversion from pressure to tension along the interface similar to the 3D indentation problem. The model is shown in figure 6.5. The layer has a thickness of $l_1 = 100\mu m$, a length of $l_2 = 80\mu m$ and interface elements are confined to a length of $l_3 = 30\mu m$ under the indenter. Two different interface properties are chosen for the simulation, which are summarized in table 6.2. Besides equal fracture toughness, the artificial stiffness ratio of model A is by a factor of 1000 greater than that of model B mainly due to its different elastic separation ratio. The model is simulated by use of continuum and interface elements with a linear element Ansatz.

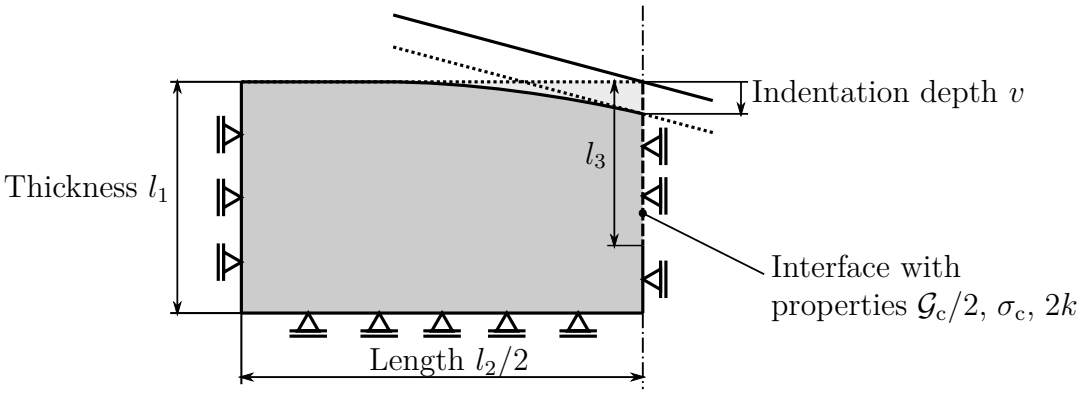

Figure 6.5: Two dimensional indentation model

Table 6.2: Properties of cohesive-zone elements

Property	Unit	Model A	Model B
Critical energy release rate $\mathcal{G}_c$	(N/mm)	—— $8.12 \cdot 10^{-3}$ ——	
Fracture strength σ_c	(MPa)	2000	700
Elastic separation ratio α		$5 \cdot 10^{-8}$	$5 \cdot 10^{-6}$
Stiffness ratio Λ (length scale $l_0 = 0.08$mm)		$6 \cdot 10^{-10}$	$5.8 \cdot 10^{-7}$
Young's modulus E	(GPa)	——— 280 ———	
Viscosity η	(sN/mm^2)	——— $1 \cdot 10^4$ ———	
Element size at interface	(µm)	0.2	0.1
Interface length	(µm)	30	80

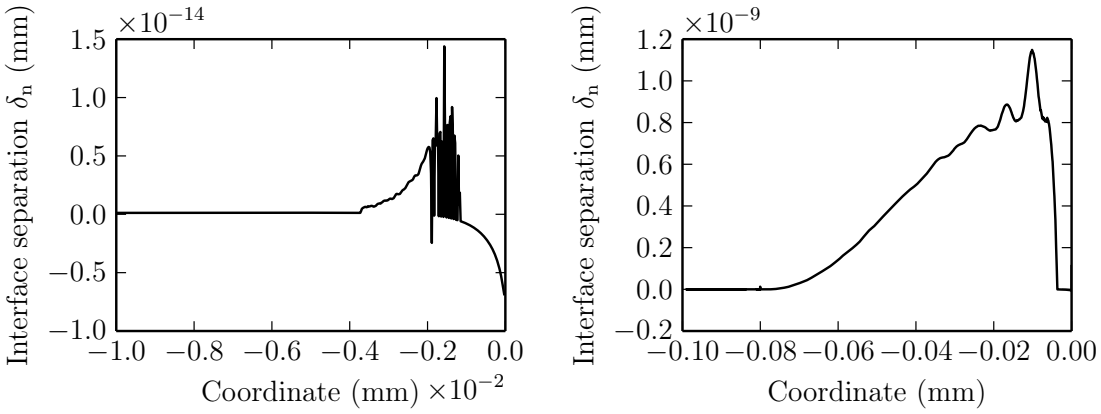

Figure 6.6: Spurious results of the interface separation are plotted against the interface coordinate for model A (left) and model B (right).

The solution of the interface separation of Model A is shown in figure 6.6 (left) at an indentation depth of $v_{\text{ind}} = 2.61\mu\text{m}$ and of Model B in figure 6.6 (right) at an indentation depth of $v_{\text{ind}} = -5.8004\mu\text{m}$. Model A, which has the lower initial stiffness shows strong negative separation of interface elements under the indenter and strong oscillations close to the depth at which the interface normal stress turns from pressure to tension. Although model B shows a smoother separation distribution, negative separations and oscillations are still visible under the indenter. In both models, the oscillations lead to strong element distortion and hence break the simulation. It seems, that increasing the artificial stiffness is not a suitable solution to prevent negative separations, because when the maximum pressure at the interface exceeds the maximum tensile stress, the interface will always damage under pressure as well.

To avoid negative interface separations, interface elements are overlayed with a contact model. The properties of the contact model are chosen in a way, that

negative separations are always far smaller than the separation at which the fracture strength is reached. Therefore, the artificial stiffness in tension and compression are successfully decoupled. The introduction of a transition zone, in which the stiffness of contact elements and interface elements is gradually increased to the boundaries of the interface zone proved to avoid convergence problems due to element distortion and or interface oscillations from the inconsistent stiffness jump.

6.6.2 Considerations on crack emergence, a critical crack length and stability

In section 6.4, the modification of the fracture strength was introduced to broaden the process zone and therefore make simulations accessible to available computer systems. This modification is only valid if the fracture process is insensitive to the fracture strength. Within the chosen modelling approach, this can only be assessed by a convergence study, as it may not be possible to resolve the process zone originating from the actual fracture strength.

Therefore, another approach is discussed that can be used to predict crack nucleation under the indenter in a 3D problem and which is used in this section to predict the load at which a crack emerges under a stress concentration in a 2D beam bending problem as well as the length of this initiated crack, which is then used as a length scale to estimate a suitable element size in the 3D problem. The approach is based on a hybrid energy–stress criterion, that is proposed in the literature to predict the emergence of a crack in an unflawed continuum in the vicinity of a stress

concentration as it is the case under an indenter, or at a stress singularity as it is the case at the front of an existing crack (see Leguillon 2002). The hybrid energy–stress criterion is based on the consideration that the initiation and propagation of a crack in unflawed continuum requires two mechanisms. First, a stress criterion must be fulfilled over the assumed crack increment Δa to ensure the existence of a flaw of this size, based on the assumption that the flaw size is proportional to the stress state. The stress criterion is given for a mode I load by

$$\sigma_{\mathrm{n}} = \sigma_{\mathrm{c}}. \tag{6.18}$$

Second, the formed crack increment can only initiate if also the crack criterion is fulfilled at the same load with this flaw size. The fracture criterion for pure mode I is given by

$$\mathcal{G} = \mathcal{G}_{\mathrm{c}}. \tag{6.19}$$

The crack increment for which both criteria are equally fulfilled at a certain load, is referred to as the critical crack length a_{c}. The associated load is referred to as the threshold load (see Johanns et al. 2014). The hybrid crack criterion consistently predicts crack initiation not only in the large crack regime when the crack behavior can be described by the LEFM, but also in the small crack regime when the process zone dominates the crack behavior. Therefore size effects are modelled accurately by this criterion.

The critical crack length can also be interpreted as the length that separates

two regimes of failure mechanisms. If an actual or hypothetic crack in a structure is smaller than the critical crack length, the structure fails due to the fracture strength. If on the other hand, the crack is greater than the critical crack length, the structure fails due to the initiation and propagation of this crack through the material driven by the energy release rate based on the fracture criterion. The critical crack length depends on the structure, the load case, elastic material properties and the fracture toughness and fracture strength.

With the introduced two–step model, the critical crack length and threshold load can be evaluated by the meeting point of two curves, the evolution of the critical load based on the stress criterion with respect to the length or area over which the stress criterion is fulfilled and the evolution of the critical load based on the fracture criterion with respect to the crack length. For simple models, evolutions can be drawn from analytical results. For more complex models, such as the indentation model, the evolutions can be drawn from results of FEM simulations with different initial assumed and actual crack sizes.

The approach is demonstrated by the example of a 2D beam bending problem, shown in figure 6.7. The end of the beam is subjected to a bending angle φ and cracks are expected to emerge in a predefined plane at the center of the beam. To evaluate the curve of the critical load with respect to the crack nucleation length based on a stress criterion ${\varphi_c}^{\sigma_c}$, a nonlinear simulation is carried out and for each load increment, the length of the region which is subjected to higher stresses than the fracture strength is evaluated, as shown in figure 6.8 (left, top). Neglecting

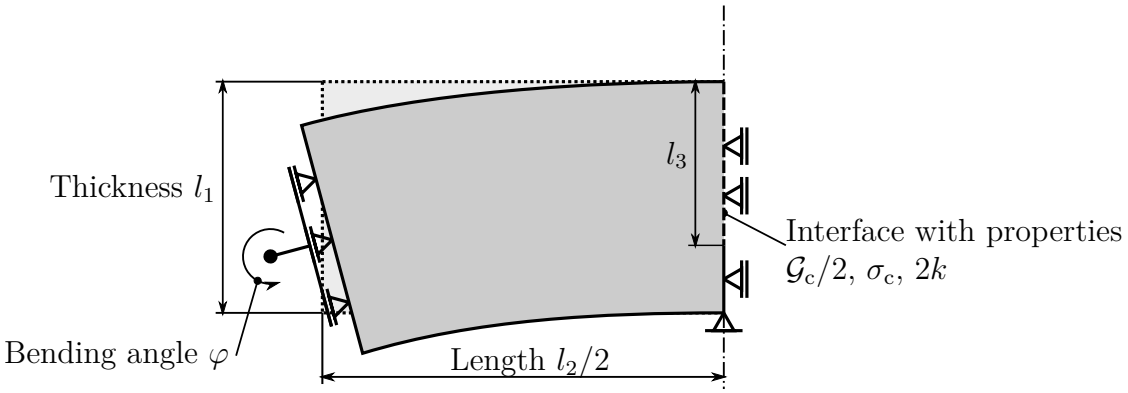

Figure 6.7: Two dimensional beam bending model.

nonlinear effects, an analytical solution can be derived from Euler-Bernoulli beam theory with the stress distribution of a cantilever beam of length $l_1/2$ subjected to a bending moment M given by

$$\sigma_n = \frac{M}{I}(l_1/2 - a) \tag{6.20}$$

with the second order area moment I. Substitution of the normal traction by the stress criterion

$$\sigma_n = \sigma_c \tag{6.21}$$

yields the critical moment

$$M_c{}^{\sigma_c}(a) = \frac{\sigma_c I}{l_1/2 - a} \tag{6.22}$$

In a second step, cracks of various lengths are incorporated in the crack plane and for each crack length, the bending angle has to be increased until the fracture

criterion is fulfilled at the crack tip, as shown in figure in figure 6.8 (left, bottom). Again, neglecting nonlinearities, simulations are not carried out until the fracture strength is reached, but the critical load is evaluated from the difference quotient of the overall potential

$$(\mathcal{G}_\mathrm{c}/2) = \frac{\Pi(a+\Delta a) - \Pi(a)}{\Delta a} \tag{6.23}$$

considering that only half of the critical energy release rate is sufficient and with the total potential of the symmetric model given by

$$\Pi = \frac{1}{2}\Pi_\mathrm{ext} = \frac{1}{2}\left[M\varphi\right](a) = \frac{1}{2}\frac{1}{k(a)}M^2(a) \tag{6.24}$$

with the stiffness

$$k(a) = \frac{M}{\varphi} \tag{6.25}$$

and taking into account that the model is loaded by a dead load, yields

$${M_\mathrm{c}}^{\sigma_\mathrm{c}}(a) = \sqrt{\frac{2(\mathcal{G}_\mathrm{c}/2)\Delta a}{\left[\frac{\varphi}{M}\right](a+\Delta a) - \left[\frac{\varphi}{M}\right](a)}} \tag{6.26}$$

similar to the solution reported in the literature for loading by an external force (see Gross & Seelig 2016).

The critical bending angle with respect to the stress criterion ${\varphi_\mathrm{c}}^{\sigma_\mathrm{c}}$ and the critical bending angle with respect to the fracture criterion ${\varphi_\mathrm{c}}^{\mathcal{G}_\mathrm{c}}$ are plotted against the crack length, as shown in figure 6.8 (right). From the intersection, the critical crack length and threshold bending angle φ_th are evaluated.

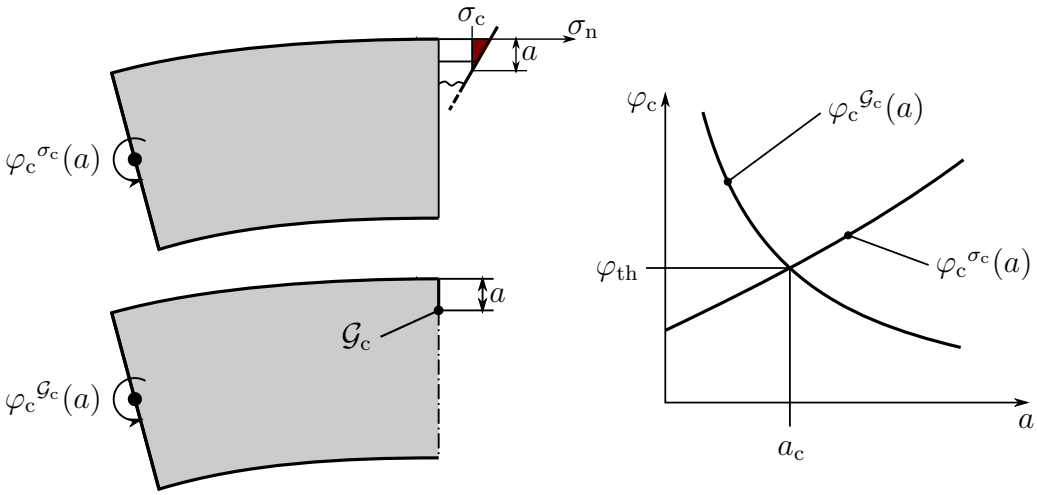

Figure 6.8: Evaluation of the nucleation and propagation curve for the two dimensional beam bending model.

Table 6.3: Material properties

Property	Unit	Steel	GaN
Young's modulus E	(MPa)	210000	280000
Poisson ratio		0.3	0.26
Critical energy release rate $\mathcal{G}_c$	(N/mm)	10.83	$5.328 \cdot 10^{-4}$
Fracture strength σ_c	(MPa)	250	7000

The nucleation and propagation curves are evaluated for 3 different property sets. Model 1 and 2 use a steel-like material. The models are 80mm thick, but differ in their length, as Model 1 is 800mm and Model 2 is 320mm long. Model 3 consists of GaN and is 80µm thick and 160µm long. All models are studied under the assumption of a plane-strain state. The material properties are summarized in table 6.3 and geometric properties in table 6.4.

The critical loads are plotted against the crack lengths in figures 6.9 and 6.10 for the steel cantilever beams and in figure 6.11 for the GaN cantilever beam. Comparing

Table 6.4: Model properties

Property	Unit	Model 1	Model 2	Model 3
Material		——— Steel ———		GaN
Length l_2	(mm)	800	320	$320 \cdot 10^{-3}$
Thickness l_1	(mm)	——— 80 ———		$80 \cdot 10^{-3}$

the critical crack-length of the two models with steel-like material behavior, no significant change of the critical crack length for the two different lengths of the models is observed. In both cases, the critical crack length is about $a_{\text{crit}} = 7\text{mm}$. In the literature, a critical crack length of $a_{\text{crit}} = 10\text{mm}$ is reported for a pressure vessel by (see Ashby et al. 2013,), which is of the same order as found in this thesis. The GaN model shows a critical crack length of approximately $a_{\text{crit}} = 1\text{nm}$. The lattice constants of the Wurtzite structure of GaN are $a_{\text{epi}} = 3.214\text{Å} \approx 0.3\text{nm}, c_{\text{epi}} = 5.119\text{Å} \approx 0.5\text{nm}$, as reported by Xiong et al. (2010). Therefore, the critical crack-length is about the same as the higher lattice-constant. All models have a descending crack propagation curve, which indicates that a crack propagates unstable through the material once initiated.

The models are simulated again applying the cohesive zone modelling approach with interface elements in the crack plane. The evaluated critical crack lengths are used as a reference for a suitable element length. An element length of a tenth of the critical crack length is used to simulate the beams of steel-like material to ensure that the process zone is well enough resolved. The GaN beam is simulated for a modified fracture strength of 500MPa and 1000MPa. The element length is chosen

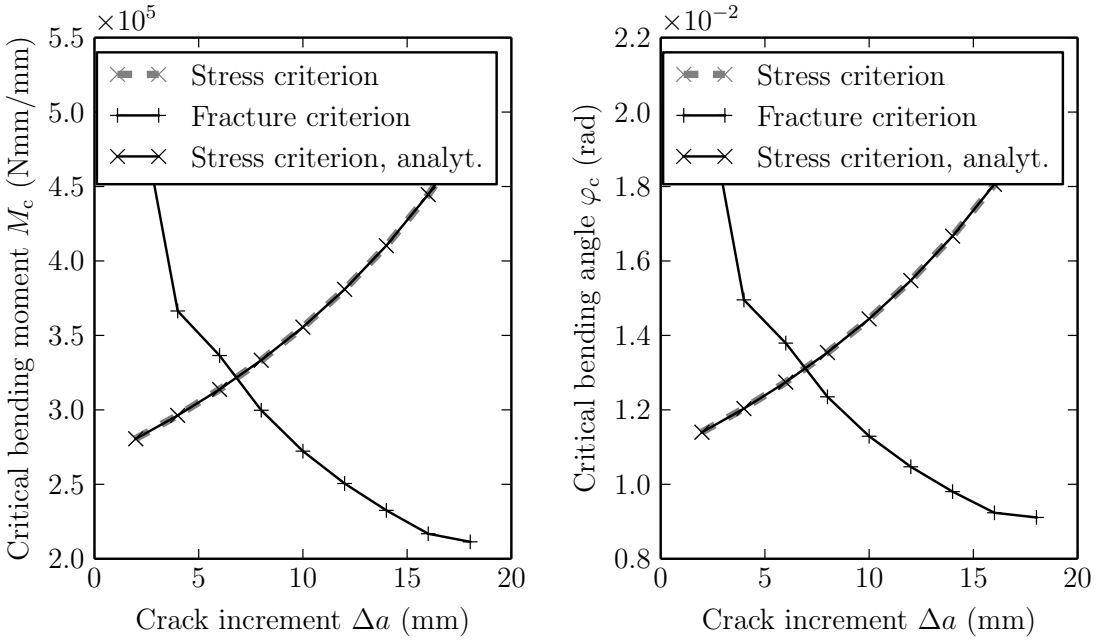

Figure 6.9: Crack nucleation and crack propagation curves for the short beam of steel-like material (Model 1).

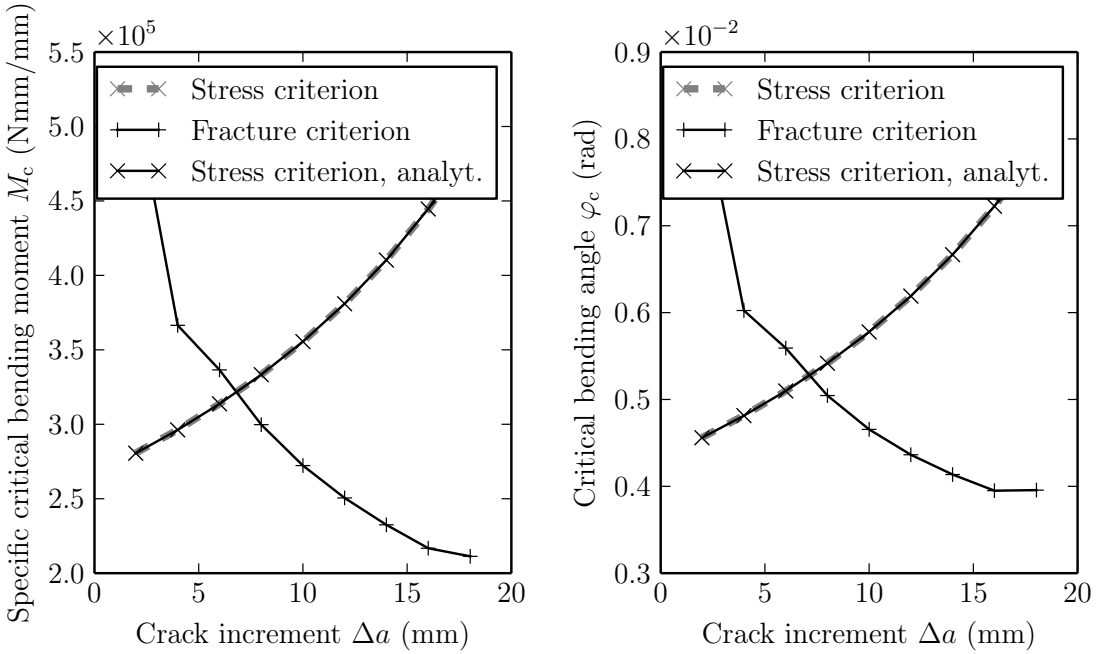

Figure 6.10: Crack nucleation and crack propagation curves for the long beam of steel-like material (Model 2).

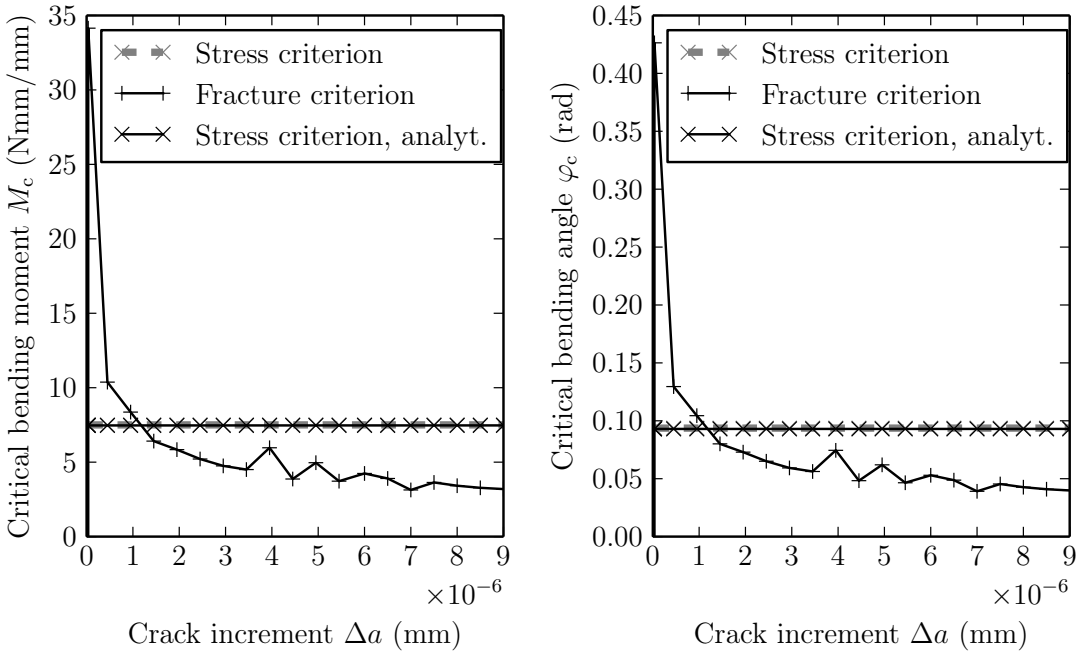

Figure 6.11: Crack nucleation and crack propagation curves for the GaN (Model 3).

larger than the critical element length according to the fracture strength reduction. The artificial and phenomenological properties of the cohesive zone elements are summarized in table 6.5.

With the chosen element length and properties evaluated from estimates discussed in section 6.4, numerical convergence could be ensured. The bending moment and the separation work of the full model are plotted against the bending angle in figures 6.12, 6.13 and 6.14. In each case, reaching a critical bending load, a steep decrease in the load-displacement curve is observed, accompanied by a steep increase of the work of cohesive tractions W_{coh} (plotted for the full model) due to crack growth. This phenomenon is an indication for rapid or even unstable crack-growth, as already predicted from the hybrid approach. The threshold load depends therefore on the

Table 6.5: Model properties

Property	Unit	Model 1	Model 2	Model 3	Model 4
Material		Steel		GaN	GaN
Mod. Fracture strength	(MPa)	250		500	1000
Element size at interface	(mm)	1		$3 \cdot 10^{-2}$	$5 \cdot 10^{-3}$
Viscosity η	(sN/mm^2)	$1 \cdot 10^{-7}$			
Elastic separation ratio α		$1 \cdot 10^{-1}$		$1 \cdot 10^{-5}$	
Stiffness ratio Λ		$9 \cdot 10^{-3}$	$2 \cdot 10^{-2}$	$2 \cdot 10^{-4}$	$3 \cdot 10^{-4}$

chosen fracture strength, as shown in figure 6.14 for the GaN model. In this case, a reduction of the fracture strength is not valid, at least not before a crack of considerable length is formed.

6.7 Summary

A set of fracture mechanical methods applicable in a FEM environment is assessed for its suitability to model crack initiation and propagation in GaN stacks. The cohesive zone method was found to be well suited for the posed problem as it allows the prediction of the initiation and propagation of complex crack pattern from a crack–free continuum. The theory of the method as well as its limitations were discussed. Suitable estimates for the phenomenological properties of the model from less complex problems were presented to keep consistency while preserving convergence when the method is applied to brittle materials. The highly brittle material properties turned out to pose some challenges in the application of the method,

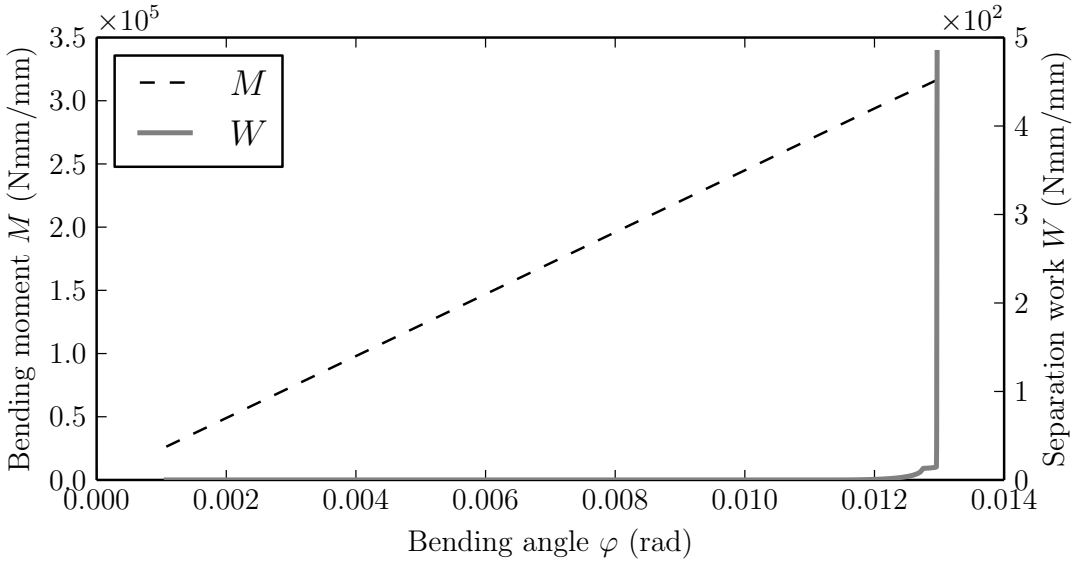

Figure 6.12: Load–displacement curve for the long steel-like beam (Model 1).

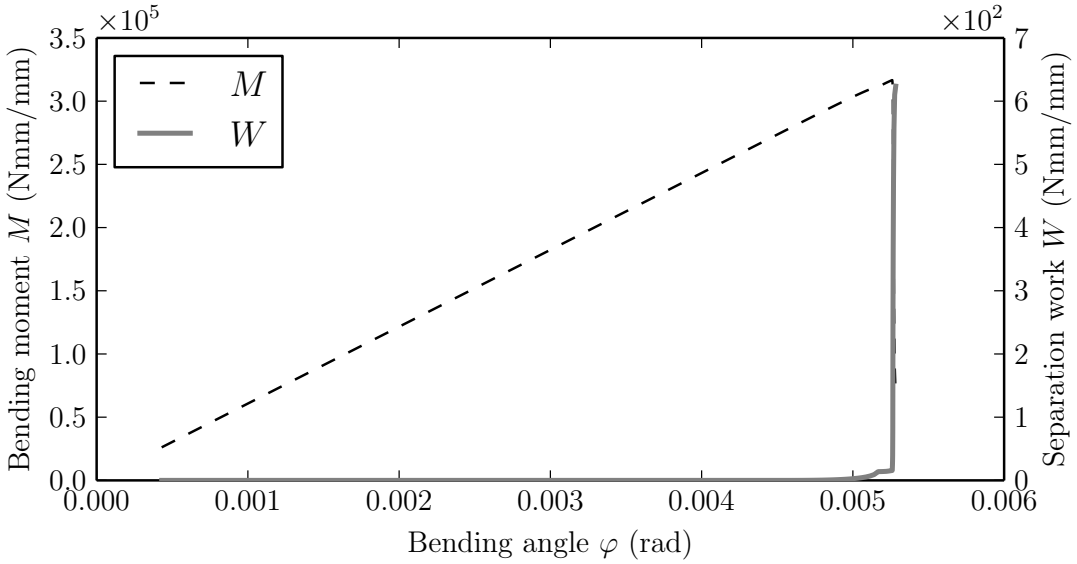

Figure 6.13: Load–displacement curve for the short steel-like beam (Model 2).

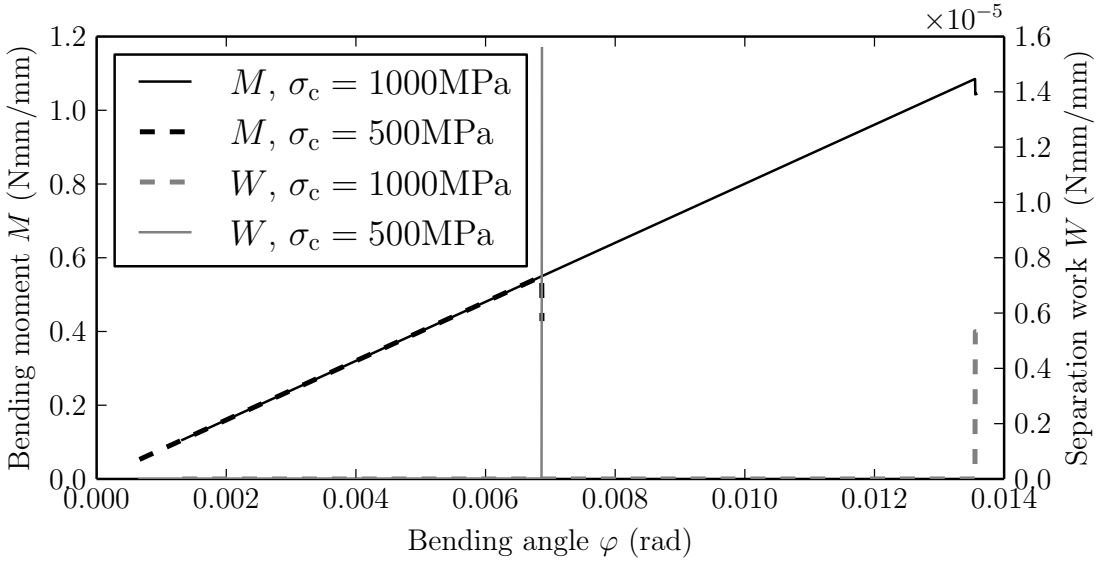

Figure 6.14: Load–displacement curve for a GaN beam (Model 3 and 4).

due to the small process zone, which has to be resolved. To address these challenges, a suitable modification of the method by reduction of the fracture strength was presented and the boundaries of their validity demonstrated by application to a less complex problem. The preliminary studies demonstrated the necessity for a critical assessment of the validity of the modification either based on a convergence study which may leave uncertainty, as the fracture strength cannot be reached, or by application of the hybrid approach prior to the application of the cohesive zone method, assessing crack nucleation and stability at all stress concentrations in the model. In any case, if the cohesive zone method is applied and crack propagation is found to be unstable by assessment of the energy balance and when the viscosity is very high to achieve convergence, the modification is invalid or lead to a considerable reduction of the threshold load and stabilization of the crack propagation.

7 Crack growth in a gallium nitride wafer loaded by nanoindentation and influence of residual stresses

The proposed modelling approach is tested on a realistic prototype problem, the nano indentation testing of a layer of GaN on a Si substrate. The influence of residual stresses on the crack initiation and propagation behavior is studied by simulation of a model with and a model without residual stresses. The properties of the models are discussed in section 7.1. The load–displacement results are discussed in section 7.2 and the crack propagation in section 7.3. The influence of the modifications of the traction–separation law are discussed in more detail in section 7.4 from the results of a convergence study with respect to the fracture strength. As the results differ in the order of magnitude of dissipation compared to experimental results from the literature, the stress state is discussed in section 7.5 to assess the onset of plastic deformation, which may be one possible reason for the differences.

Table 7.1: Thickness of the layers for each stack in bottom-up sequence

	Dimension	Stack 1	Stack 2
Stacking sequence	(µm)	7 (GaN) 12 (Si)	6 (GaN), 0.46 ($Al_{0.5}GaN_{0.5}$), 0.2 (AlN), 12 (Si)
Wafer radius	(µm)	100	

7.1 Model properties

A 1/6 model is used, as the problem shows three-fold symmetry with each fold being symmetric itself under the approximations and assumptions, discussed in section 6.2. The GaN layer is 7µm thick and is deposited on a 12µm thick Si layer. The radius of the wafer is chosen to be 100 µm. Although typical Si substrates have larger dimensions in thickness and radius, these dimensions are chosen, as stresses are assumed to be sufficiently decayed at the model boundaries. The modelling approach is tested on an additional stack with 0.200µm thick AlN and a 0.460µm thick AlGaN interlayer deposited between the GaN and Si layers. The geometric properties are summarized in table 7.1

The Si–GaN stacks are subjected to two different residual stress states, as summarized in table 7.2. The first model remains without residual stresses and to the second model, a non-equilibrium biaxial stress state in the GaN layer with 400MPa in each direction is applied, from which an equilibrium stress state is reached after stress rearrangement. All models are loaded to a maximum indentation depth of 1.7µm, which is higher than the indentation depth in experimental tests, discussed

Table 7.2: Thickness of the layers for each stack in bottom-up sequence

Property	Dimension	Model 1	Model 2	Model 3
Stack		Stack1	Stack 2	Stack 1
Indentation depth v	(µm)	——— 1.7 ———		
Residual stress components $\sigma_{xx}^{res}, \sigma_{yy}^{res}$	(MPa)	0		400

Table 7.3: Material properties

Property	Dimension	Si	GaN	AlN	$Al_{0.5}$ $Ga_{0.5}N$
Young's modulus E	(10^3 MPa)	130	280	349	294
Poisson ratio ν	(1)	0.27	0.26	0.25	0.255
Crit. stress intensity factor K_{Ic}	(MPa$\sqrt{m}$)	0.7	0.4	3.5	1.95
Mod. fracture strength σ_{cMod}	(MPa)	—— 1000 ——		295	648
Crit. energy release rate $\mathcal{G}_c$	(10^{-3} N/mm)	3.4945	0.5328	37.287	12.093

in section 5.2.2, but still of the same order and is therefore expected to lead to a large crack pattern compared to the cohesive zone length, but sufficiently confined to not interact with the model boundaries.

The material properties are chosen from the literature study in section 6.5 with the critical energy release rate derived from a plane strain state assumption. A modified fracture strength of 1/7 of the material fracture strength is used. The material properties are summarized in table 7.3.

The parameters, such as the viscosity and the elastic separation ratio (6.8) of the cohesive zone model are summarized in table 7.4. An element size of 0.3µm was found to be suitable, although it is considerably larger than the estimate of the

Table 7.4: Mesh properties and model parameters.

Property	Dimension	Model 1	Model 2	Model 3
Element size L	(10^{-3}mm)	———	0.3	———
Viscosity η	(sN/mm^2)	———	$1 \cdot 10^{-1}$	———
Elastic separation ratio α	(1)	0.5	0.5	10^{-5}

Table 7.5: Material properties

Property	Si	GaN	AlN	$Al_{0.5}$ $Ga_{0.5}N$
Stiffness ratio Λ $l_0 = 1.3$µm, $\alpha = 0.5$	0.38	0.1	100	7
Relaxation time ratio Θ $l_0 = 1.3$µm, $V = 1.7$µm/s	0.03	0.6	$1 \cdot 10^{-5}$	0.0004

cohesive zone size based on the CCT specimen from equation (6.5).

From the parameters of the cohesive zone method, the stiffness ratio (6.6) and the relaxation time ratio (6.13) are evaluated with the characteristic length scale chosen as a third of the perimeter of the radius of the assumed process zone length. Based on the observation, that an element length of 0.3µm is sufficient to have at least two elements in the process zone, a length scale of $l_0 = 1.3$µm is used for this purpose. The results are summarized in table 7.5. It should be noted, that the stiffness ratio decreases when the radius of the expected penny crack increases and therefore even tough the artificial compliance may be too high at the beginning, it may be well chosen when the crack is well formed.

Simulations are carried out with the described properties and parameters. Load–

displacement curves are evaluated, contour plots are extracted from the cohesive zone state variables to study the crack propagation. The tractions and separations are integrated over indentation time to estimate the amount of the stored artificial elastic energy and dissipated viscous energy. The results are discussed in the following sections.

7.2 Load–displacement curve

Figure 7.1 shows the load–displacement curves of stack 1 and stack 2 without residual stresses. The differences between stack 1 and stack 2 in the load-displacement behavior are negligible. The AlN and AlGaN interlayers are too thin to provide considerable contribution to the overall load-displacement behavior of the stack. The load-displacement curves are characterized by a positive curvature. At the full indentation depth $v = 1.7$µm a reaction force of $F = 1.356$N is evaluated. No difference is visible between the loading and unloading curves. Hence, dissipative effects must be a few orders of magnitude smaller than the work applied. The work applied at full indentation depth is $W = 0.822$Nµm , evaluated from integration of the reaction force F over the indentation depth v. At low indentation depth, the load–displacement curve shows good agreement with measurements from nano-indentation in 2µm thick GaN films grown by metallo-organic chemical vapour deposition (MOCVD) found in the literature (see Tsai et al. 2008). Values from this thesis and from the literature are provided in table 7.6. With increasing indentation

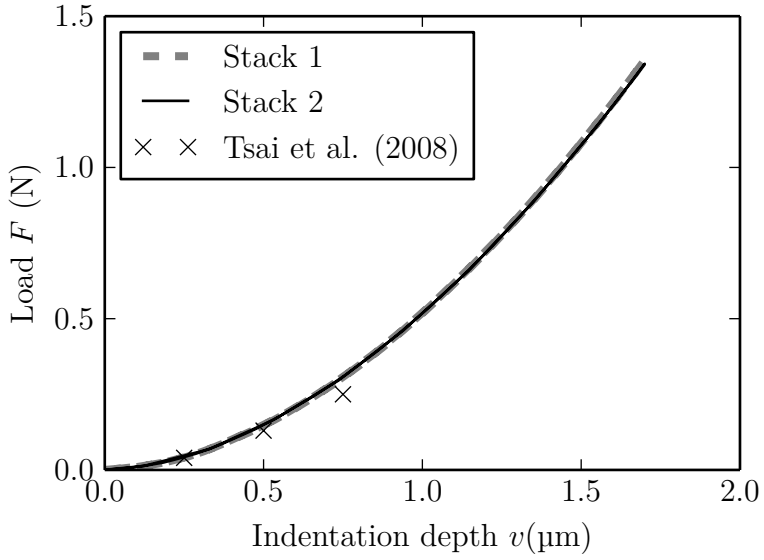

Figure 7.1: The force of the indenter is plotted against the indentation depth.

depth, the results deviate significantly. The experimental results show lower values for the indentation force than the results from the simulation in this thesis. The model behaves stiffer than the experiment, which may be subjected to the spatial discretization, which induces artificial stiffness, or neglected dissipative effects such as friction or plasticity. The pop-in events that are found in the literature, discussed in section 5.2 are absent in figure 7.1, because they are subjected to dislocation nucleation and propagation, which is not included in the modelling approach.

7.3 Crack propagation

The spatio-temporal propagation of the crack is discussed by means of a contour plot of the critical separation of the interface-elements, as shown in figure 7.2 for stack 1 and in figure 7.3 for stack 2 without residual stresses as well as in figure 7.8

Table 7.6: Comparison of load-displacement values with values from the literature.

Indentation depth v (nm)	Reaction force F(mN)	
	This work	Tsai et al. (2008)
250	43	40
500	153	130
750	307	250
1000	518	–

for stack 1 with residual stresses.

For stack 1 in the absence of residual stresses, a crack initiates at low indentation depth compared to the layer thickness directly under the indenter, as shown in figure 7.2 (left). With increasing indentation, the crack grows towards the interface. At an indentation depth of $v = 0.6$µm, the crack reaches the interface and has a radial extent at the interface that is similar to the thickness of the GaN layer, as shown in figure 7.2 (middle). The radial extent at the surface is only a fourth of the radial extent at the interface. The crack propagates into the Si layer from under the indenter. At the interface, the crack propagation in the Si layer lags behind the crack propagation in the GaN layer due to the higher compliance and fracture-toughness. When the highest indentation depth of $v = 1.7$µm is reached, a „penny-wise“ crack in the Si layer has formed with equal radial and vertical extent, as shown in figure 7.2 (right). Compressive stresses lead to closing of the crack under the indenter at full indentation depth.

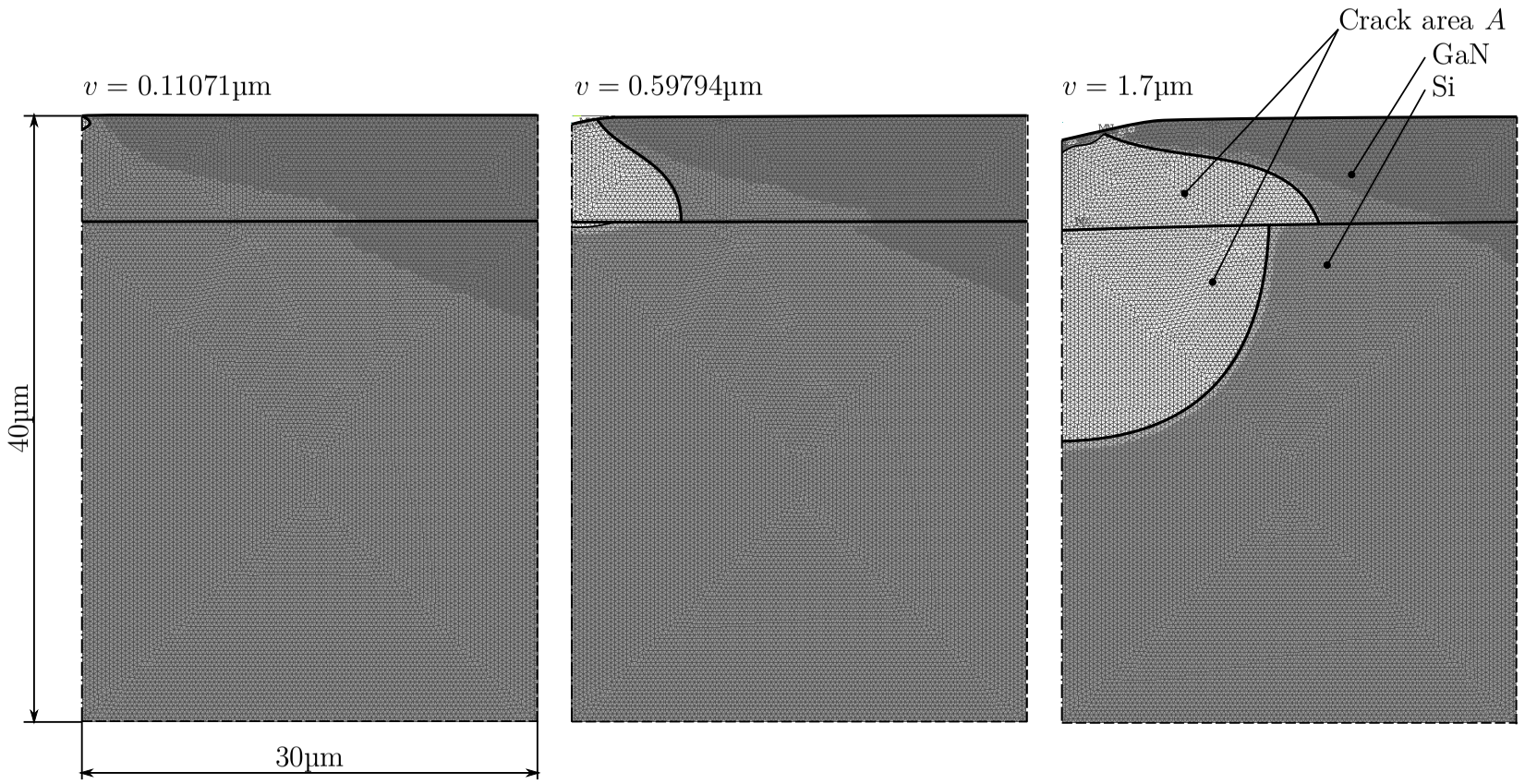

Figure 7.2: Contour plot of the interface separation of stack 1 without residual stresses at low (left), medium (center) and maximum (right) indentation depth with the crack plane (white) surrounded by the process zone (light grey) and artificial elastic tension (grey) and artificial elastic compression (dark grey).

For stack 2, the crack propagation is similar to stack 1, with only low contribution of the thin inter-layers. The similarity is a result of the disability of such small interlayers to drastically change the overall elastic response of the model. Also, the higher fracture toughness of the interlayers cannot be utilized to decrease the overall crack-area. The overall dissipated energy is a product of crack-area and critical energy-release rate. Due to the restriction of crack-propagation to the radial-vertical plane, the available crack-area in the thin interlayers is very small, due to the small thickness compared to the Si and GaN layers. To stop the crack at an interface, crack-meandering to the interface-plane is necessary which requires a stress state that has strong shear or normal components perpendicular to the interface to allow crack-initiation in the interface-plane.

In the radial-vertical plane, the high critical separation of AlN prevents separation of the interface elements even at the greatest indentation depth. In the AlGaN layer, the crack initiates under the indenter and grows radially outwards as in the other layers with increasing indentation depth. The radial extent is smaller than in the GaN and Si layer, as the stiffness of the materials prevent high gradients of the separation of the interface-elements.

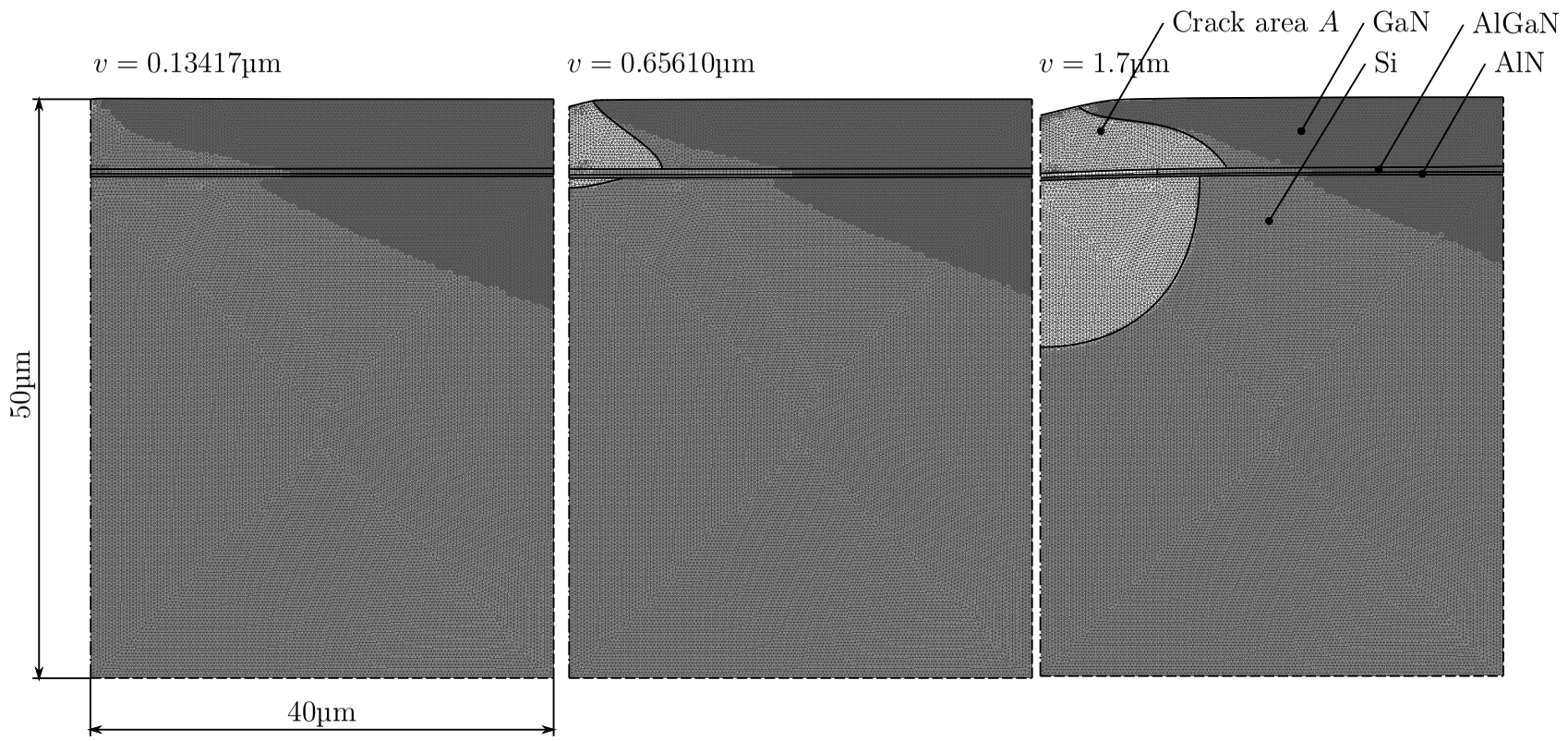

Figure 7.3: Contour plot of the interface separation of stack 2 without residual stresses at low (left), medium (center) and maximum (right) indentation depth with the crack plane (white) surrounded by the process zone (light grey) and artificial elastic tension (grey) and artificial elastic compression (dark grey).

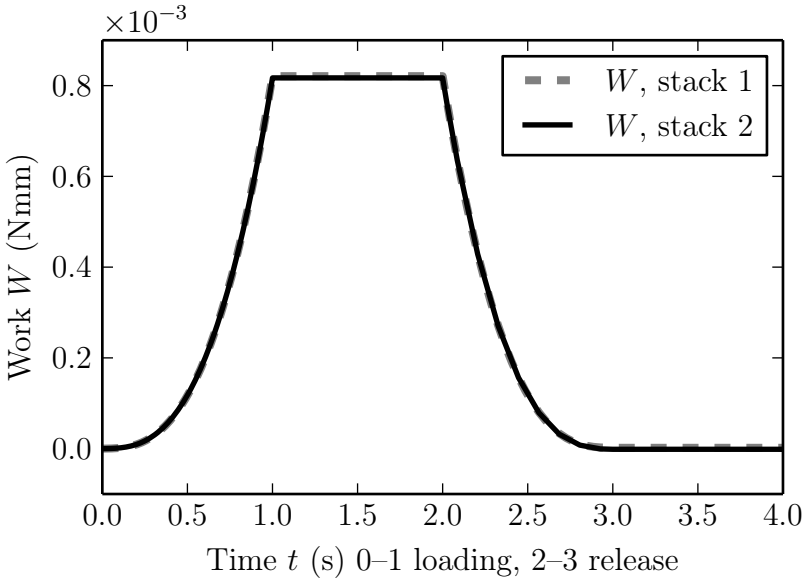

Figure 7.4: The applied work is plotted against the simulation time. Measures for force, energy and work are related to the full 1/1 model.

Considering the dissipated energy due to material separation

$$\begin{aligned} W_{\text{coh}} &= 3\left(\mathcal{G}_{\text{c}}{}^{\text{Si}} \cdot A^{\text{Si}} + \mathcal{G}_{\text{c}}{}^{\text{GaN}} \cdot A^{\text{GaN}}\right) \\ &= 3\left(3.494510^{-3}\text{N/mm} \cdot 158 \cdot 10^{-6}\text{mm}^2 + 5.32810^{-4}\text{N/mm} \cdot 75 \cdot 10^{-6}\text{mm}^2\right) \\ &= 1.776 \cdot 10^{-6}\text{Nmm} \end{aligned} \tag{7.1}$$

it is obvious, that the dissipative mechanism of crack propagation is too low in its order of magnitude compared to the overall work, estimated from the area under the load–displacement curve, which was shown in figure 7.1. For completeness, the evolution of the applied work is shown in figure 7.4.

The chosen modifications of the traction–separation law require a critical discussion of the contribution of artificial energy and viscous work. The evolution of

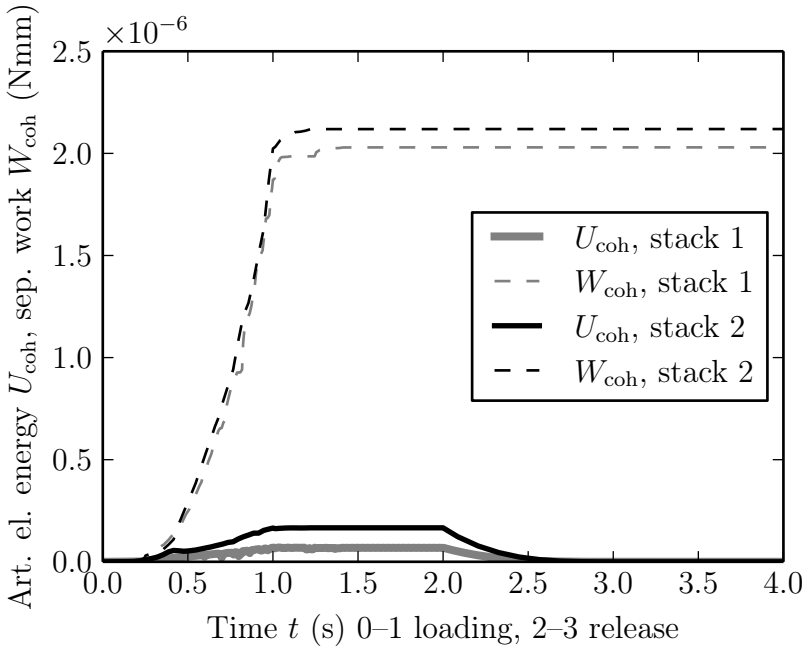

Figure 7.5: The artificial elastic energy U_{coh} and work of cohesive tractions W_{coh} are plotted against the simulation time t. Measures for force, energy and work are related to the full 1/1 model.

the overall artificial elastic energy is plotted in comparison to the work of cohesive tractions against the simulation time in figure 7.5 for both stacks without residual stresses. After reaching the threshold load, the artificial elastic energy increases until it reaches a fraction of about 5% for stack 1 and a fraction of 10% for stack 2 of the work of cohesive tractions is reached. The values differ due to the larger extent of the area with interface elements of stack 2. The separation work of stack 2 as shown in figure 7.5 (dashed, grey) shows some increase even after the full indentation load at $t = 1$, which indicates that crack propagation is suppressed by viscous forces.

The ratio of viscous work compared to the work of cohesive tractions is estimated from integrating equation (6.16) over the crack-plane and over the simulation time. The evolution of the viscous work ratio is plotted against simulation time in figure

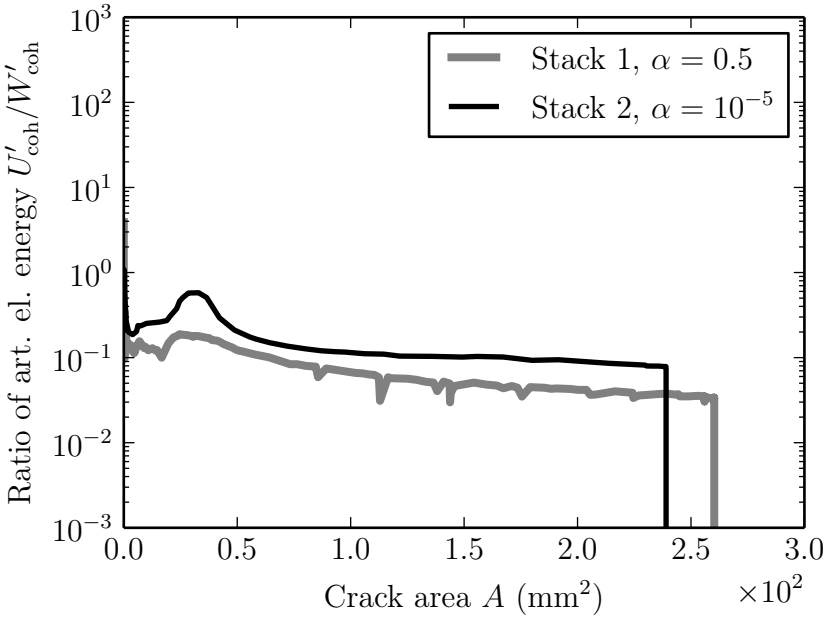

Figure 7.6: The ratio of artificial elastic energy U_{coh} over the work of cohesive tractions W_{coh} is plotted against the simulation time.

7.7. Both stacks show some oscillations and high artificial elastic energy after crack initiation. With increasing crack area, the amount of artificial elastic energy converges to a very low value for stack 1, whereas it keeps quite the same for stack 2.

Figure 7.8 shows the temporal-spatial crack propagation for stack 1 subjected to a residual stress state in the GaN layer. The crack initiates again under the indenter at the surface, as shown in figure 7.8 (left) and a penny crack is formed as shown in figure 7.8 (middle). Compared to the model without residual stresses, the crack reaches the boundary of the confined crack plane at a low indentation depth of $v = 0.1\mu\mathrm{m}$.

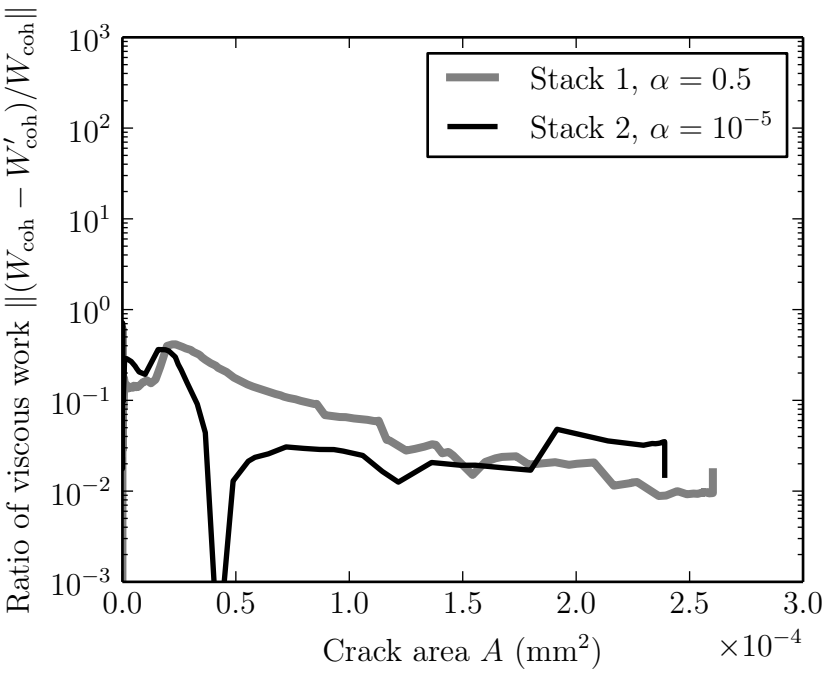

Figure 7.7: The ratio of the work of viscous forces over the work of chohesive tractions is plotted against the simulation time.

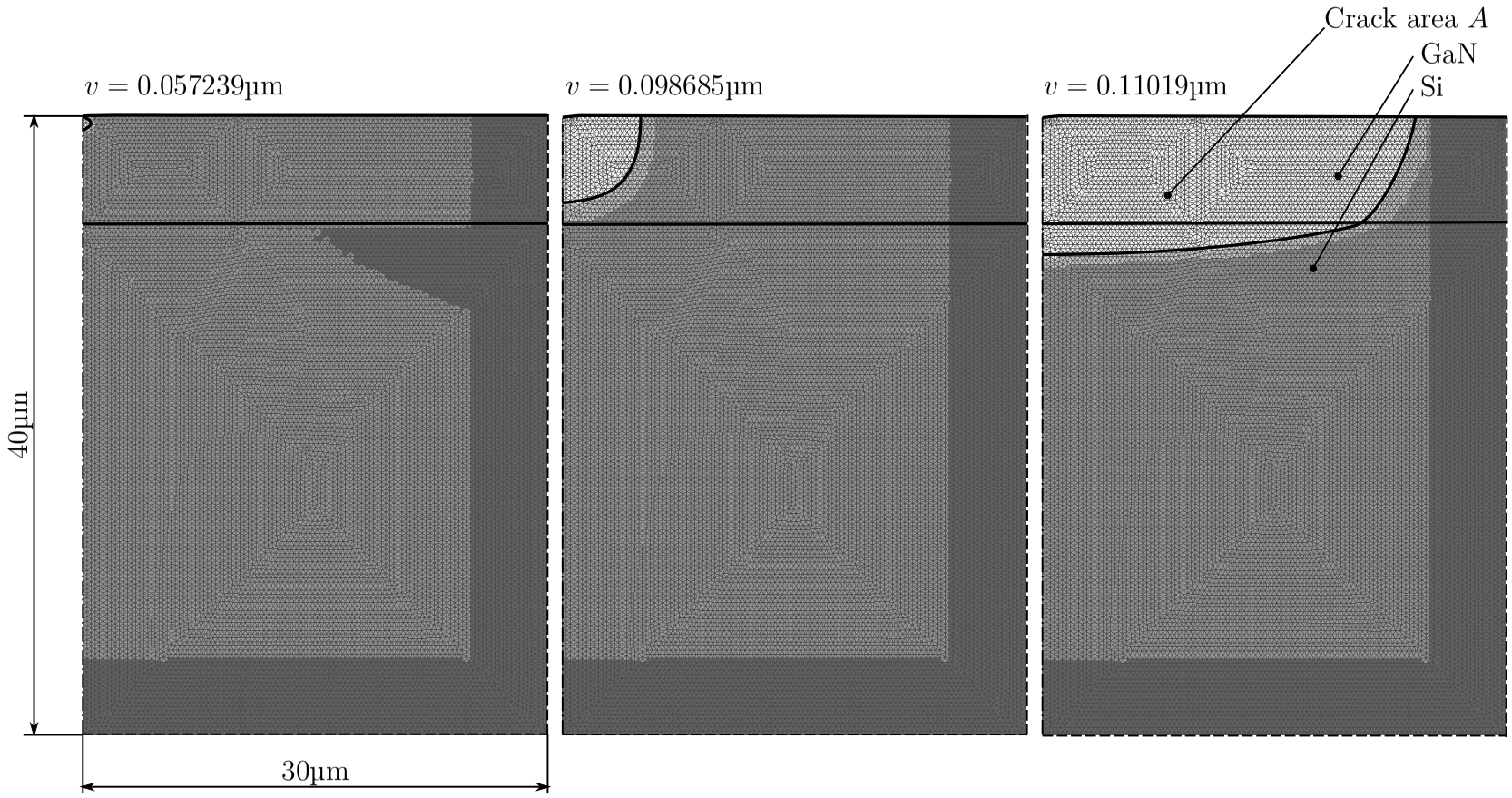

Figure 7.8: Contour plot of the interface separation of stack 1 with residual stresses at low (left), medium (center) and maximum (right) indentation depth with the crack plane (white) surrounded by the process zone (light grey) and artificial elastic tension (grey) and artificial elastic compression (dark grey)

The work of cohesive tractions of stack 1 with and withour residual stresses are compared in figure 7.9, showing the strong increase in the work of cohesive tractions at crack initiation. The contribution of artificial elastic energy and viscous work are plotted for stack 1 with residual stresses compared to stack 1 without residual stresses in figure 7.10 and 7.11. The influence of artificial elastic energy is similar and under 10% for both models. The stack with residual stresses shows a high amount of viscous forces, which are necessary to stabilize the propagation of the crack, an indication for unstable crack propagation.

For a more detailled view on the error due to the modifications of the traction–separation law, a convergence study is carried out, which is discussed in the next sections.

7.4 Influence of modification of the fracture strength

The influence of the modification of the fracture strength is studied by a convergence study on the GaN–Si stack (Stack 1) with three different fracture strength modifications of 500MPa, 1000MPa and 2000MPa. The model with 2000MPa led to the highest computational effort that was acceptable with the available computer system and time. The models differ in element size at the interface, in the artificial stiffness ratio and in viscous regularization to minimize computational time, but ensure convergence. The properties of the model are summarized in table 7.7

The different fracture strength leads also to different stiffness ratio and relaxation

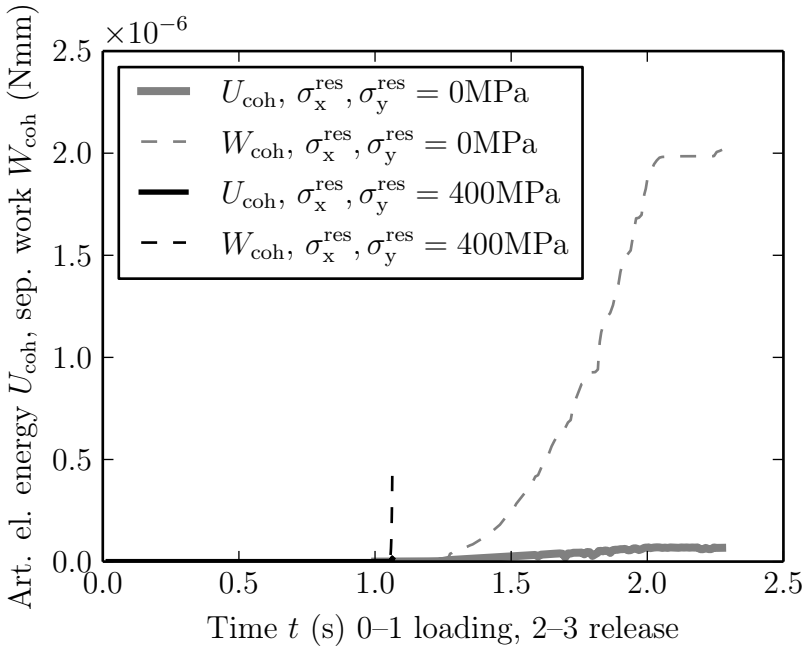

Figure 7.9: The artificial elastic energy U_{coh} and work of cohesive tractions W_{coh} are plotted against the simulation time t. Measures for force, energy and work are related to the full 1/1 model.

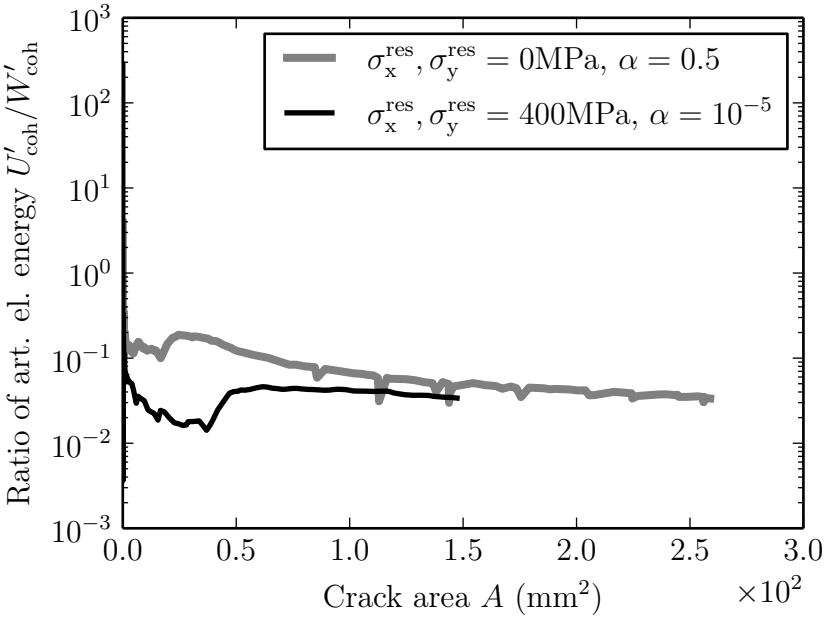

Figure 7.10: The ratio of artificial elastic energy U_{coh} over the work of cohesive tractions W_{coh} is plotted against the simulation time.

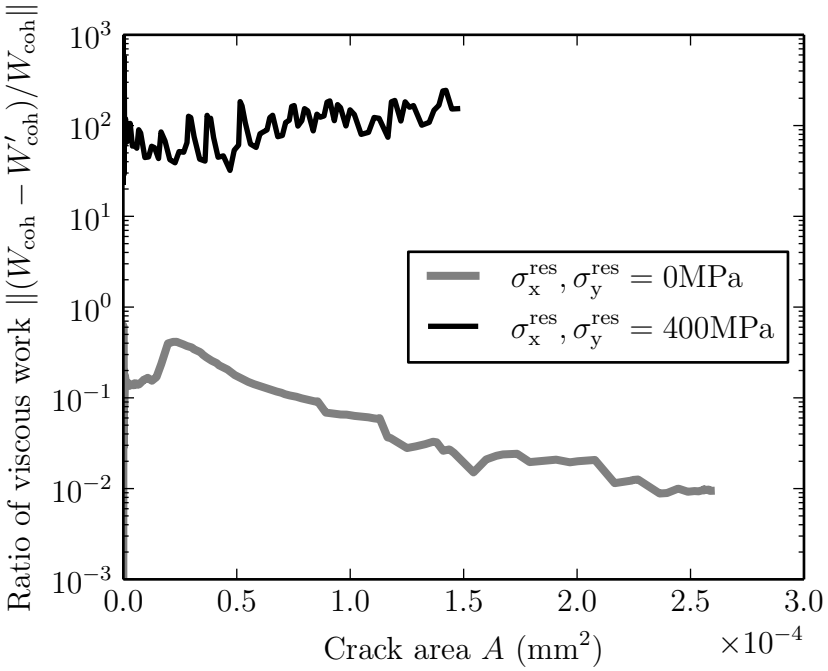

Figure 7.11: The ratio of the work of viscous forces over the work of chohesive tractions is plotted against the simulation time.

Table 7.7: Mesh properties and model parameters.

Property	Dimension	Model 1	Model 2	Model 3
Fracture strength		500	1000	2000
Element size L	(10^{-3}mm)	——— 0.3 ———		0.15

Table 7.8: Consistency measures

Property	Fracture strength (MPa)	Si	GaN
Stiffness ratio Λ, $l_0 = 1.3$µm	500	1.507	0.49
Relaxation time ratio Θ, $l_0 = 1.3$µm		0.0081	0.16
Stiffness ratio Λ, $l_0 = 1.3$µm	1000	0.38	0.12
Relaxation time ratio Θ, $l_0 = 1.3$µm		0.03	0.65
Stiffness ratio Λ, $l_0 = 0.65$µm	2000	0.19	0.062
Relaxation time ratio Θ, $l_0 = 0.65$µm		0.065	1.3

time ratio, which are provided in table 7.8 The influence of the modified fracture strength on the evolution of crack propagation is low, as shown in figure 7.12 by a plot of the work of cohesive tractions against the simulation time. The threshold load is not subjected to a considerable change due to the fracture strength. Differences in the separation work and therefore in the overall crack area are due to the different artificial stiffness ratio, as provided in table 7.8. A more detailled view on the ratio of artificial elastic energy U_{coh} over the work of cohesive tractions W_{coh} is provided in figure 7.13. The evolution of the viscous work ratio is plotted against simulation time in figure 7.14. The models with a modified fracture strength of $\sigma_{\text{cMod}} = 500$MPa and $\sigma_{\text{cMod}} = 2000$MPa show a similar evolution of the viscous work, whereas the model with $\sigma_{\text{cMod}} = 1000$MPa shows a different behavior, due to the suppressed crack growth, which is resolved in the hold state of the indenter. All three models show a long term ratio of viscous work below 10%. The convergence study indicates

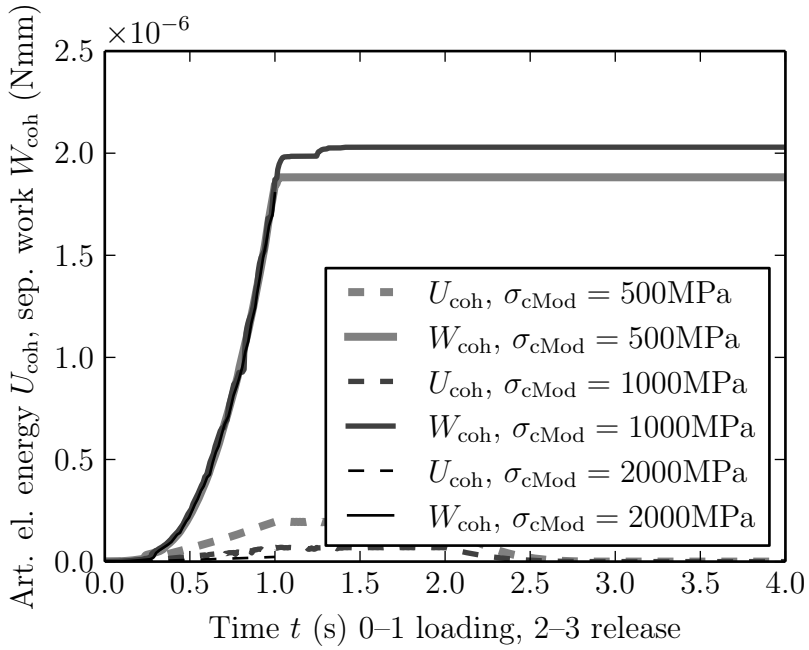

Figure 7.12: The artificial elastic energy U_{coh} and work of cohesive tractions W_{coh} are plotted against the simulation time t. Measures for force, energy and work are related to the full 1/1 model.

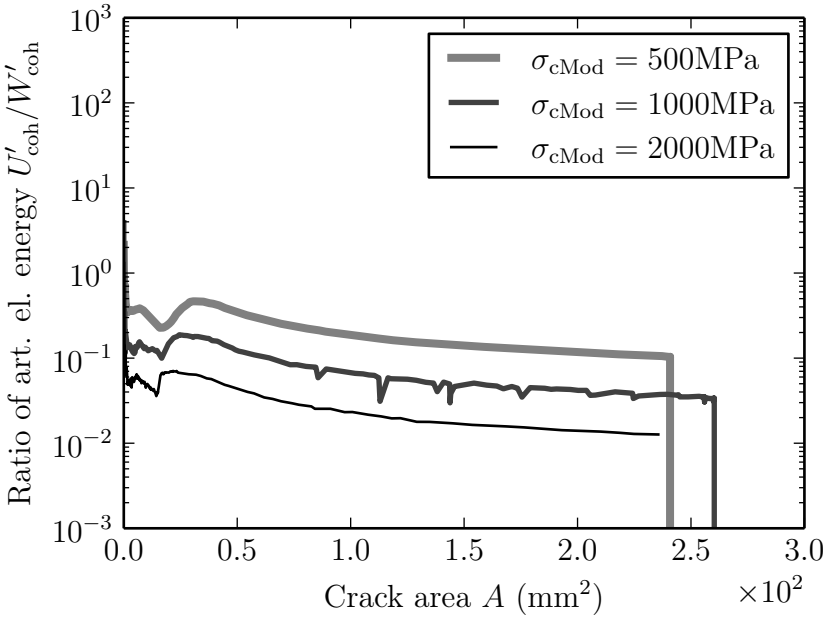

Figure 7.13: The ratio of artificial elastic energy U_{coh} over the work of cohesive tractions W_{coh} is plotted against the simulation time.

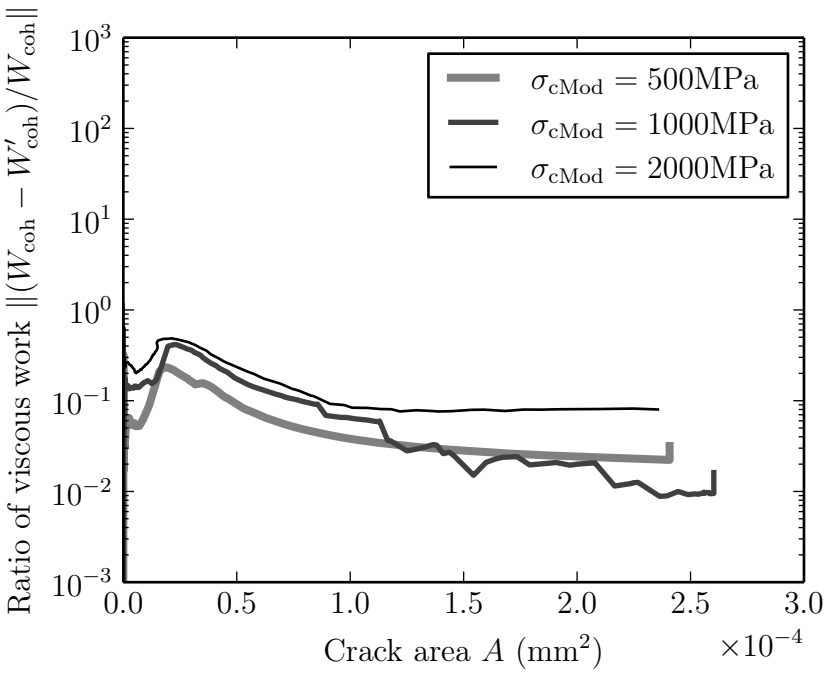

Figure 7.14: The ratio of the work of viscous forces over the work of chohesive tractions is plotted against the simulation time.

a stable behavior, that is not too much influenced by artificial numerical effects. To finally exclude unstable behavior that may be stabilized by viscous forces, would still require the application of the approach discussed in section 6.6.2 or using the original fracture strength and fully resolving the process zone at least in the earlier state of crack propagation until a crack of considerable length with respect to the length of the process zone is formed.

7.5 Elastic stress state and begin of plastic deformation

In section 7.2 it was shown, that the phenomenon of fracture alone cannot explain the strong dissipation observed in experimental nanoindentation testing of GaN in the literature. A possible origin of the strong dissipation is plasticity, which is observed in nanoindentation testing in the literature, as discussed in section 5.2. Although the involved materials are assumed to be brittle and therefore the amount of plastic deformation should be low before fracture, the expected highly compressive region under the indenter may still allow a certain amount of plastic deformation that may contribute by dissipation to the overall energy balance and therefore even have a stabilizing effect on crack growth. The onset of plasticity is therefore assessed by discussion of the elastic stress state from results of the simulation.

To assess the onset of plastic deformation, a yield criterion is required. The plastic deformation behavior of GaN is still topic of research. Experimental micropillar compression tests in the literature give rise to the understanding of plastic slip in GaN. {0001}-oriented GaN prisms grown by MOVPE are observed in the literature to shear off in the $\{11\bar{2}2\}$ plane when a critical resolved shear stress of $\tau_{\mathrm{c}}^{\mathrm{slip}} = 2800\mathrm{MPa}$ is reached (see Wheeler et al. 2013).

The critical resolved shear stress from the literature is used for comparison with two types of measures from the stress results of the simulation. The stress intensity or Tresca equivalent stress, which is twice the maximum shear stress is computed and

shown in figure 7.15 (left). To also take the anisotropy of the hexagonal structure into consideration, the maximum of the resolved shear stress components in a cone that includes the same angle with the $\{0001\}$ axis as the $\{11\bar{2}2\}$ plane is evaluated and shown in figure 7.15 (right). The contour plots show a fold of the symmetry plane on the left side of the axis and a fold of the crack plane on the right side. Contour plots are provided for three different indentation depths.

The highest shear stresses emerge close to the indenter tip. The isolines of equal shear stresses have higher radial coordinates at the surface and at the interface and an hourglass shape in-between. In the silicon layer, the isolines are of circular shape and maximum shear stresses decrease with the radius of this circular shape. At full indentation depth, the Tresca equivalent stress exceeds a value of 10GPa over the full thickness of the GaN layer. Even before crack-initiation, a considerable region of at least 1/10 of the GaN-layer thickness exceeds this value. Considering that the Tresca equivalent stress is twice the maximum shear stress, the maximum shear stress exceeds by far the critical resolved shear stress. A similar behavior is observed for the resolved criterion in figure 7.15 (right), although the topology differs slightly. The results provide an indication not only for plasticity but also for a large area that may be subjected to plasticity similar to the results of nanoindentation tests found in the literature.

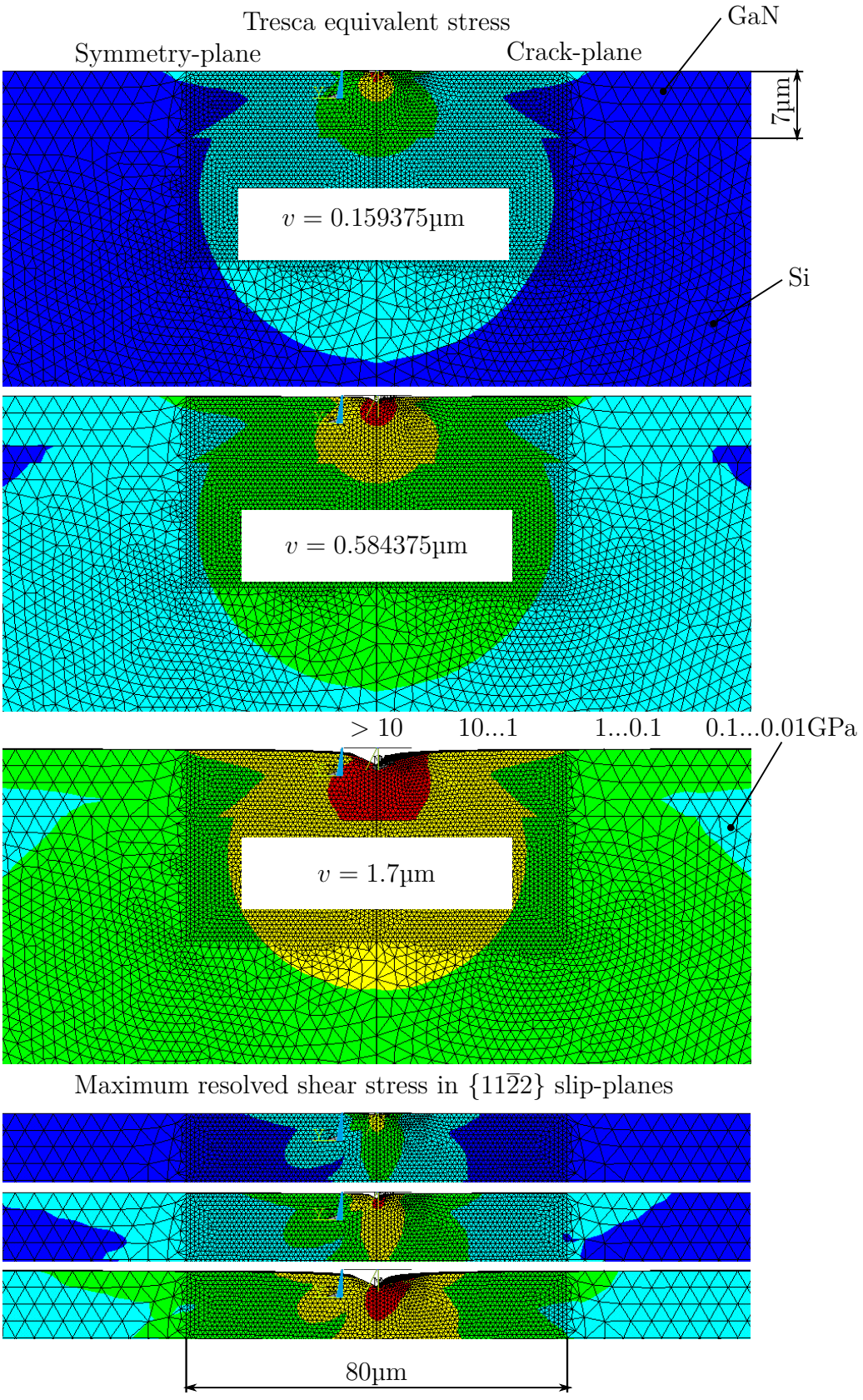

Figure 7.15: A contour plot of the stress intensity (left) and the maximum resolved shear stresses perpendicular to axisymmetric planes with equal angle than the $\{11\bar{2}2\}$ slip planes are shown.

8 Conclusions

A literature review demonstrated the importance to gain a better understanding of the relationship between stack design and fracture liability of GaN stacks, to decrease stack complexity and hence manufacturing cost and increase reliability and therefore allow to tap the full potential of GaN technology based on its advantageous electronic properties. The requirement to assess brittle fracture of GaN stacks in the stack design process is mainly attributed to the growth of the stacks on a foreign substrate in the literature in conjunction with the high deposition temperatures, which lead to strong tensile residual stresses due to lattice and thermo-mechanical mismatch between the materials. Although homoepitaxial growth on GaN substrates is possible, the growth on Si substrate is cheaper and enables larger wafer diameters. As a downgrade, Si has the most unfavorable mismatch in the thermo-elastic properties to GaN, which motivated to choose a Si–GaN wafer as the prototype problem to study in this thesis.

A modelling approach based on the FEM method and fracture mechanical methods within the FEM environment under the assumption of LEFM was developed

and demonstrated to be applicable to predict crack initiation and propagation in a brittle-elastic abstraction of these stacks. The strong driving force of residual stresses, even when they are rather small compared to the fracture strength, was demonstrated on a Si–GaN wafer and shown to lead to failure once a crack is initiated. The challenge in applying the modelling approach to more complex problems lies in the limitations of computational systems to allow the simulation at the required spatial and temporal resolution that is required due to the highly brittle material behavior. The proposed workarounds, such as the modification of the fracture strength, can lead to considerable reduction of the simulation effort, but may not consistently predict fracture if crack propagation is unstable. A workaround is found in the application of the hybrid-approach to model crack nucleation in the flaw-free continuum at all possible stress concentrations and decouple the simulation of crack nucleation from that of crack propagation. The combination of the hybrid approach and the cohesive zone method may be a potential combination to predict crack resistance in stacks where fracture is characterized by an initial period of unstable nucleation and propagation and subsequent stabilization when the crack propagates into areas in which the magnitude of stress components that drive crack growth are lower. Another challenge that lies in the prediction of crack initiation and propagation in real stacks lies in the difficulty to obtain accurate values of the fracture strength which may be influenced by the defect structure in the final stack and also by plastic deformation when the stack is loaded. Another challenge may arise from the deviations in the residual stress state between the wafers of the

same stack sequence, which may restrict predictions to only a single wafer. An important requirement in the reliable prediction is therefore an accurate control and measurement of the residual stress state.

Considerations on the challenges in the application of the modelling approach are based on the assumption of a purely elastic brittle material behavior without further dissipative effects involved. Experimental results from the literature show strong dissipation and plasticity under the indenter. The study of the elastic stress state in this thesis contributes to the understanding by demonstrating the strong shear stresses under the indenter, especially perpendicular to the slip planes of the GaN crystal structure. Future studies may therefore aim on incorporating a plasticity model, which may also reduce the challenges in the application of the proposed modelling approach due to incorporation of dissipation that may stabilize crack propagation.

Bibliography

Amar, A. B., Faucher, M., Brandli, V., Cordier, Y. & Théron, D. (2014), 'Young's modulus extraction of epitaxial heterostructure algan/gan for mems application', *Phys. Status Solidi A* **211**(7), 1655–1659.

Anstis, G. R., Chantikul, P., Lawn, B. R. & Marshall, D. B. (1981), 'A critical evaluation of indentation techniques for measuring fracture toughness: I, direct crack measurement', *Journal of the American Ceramic Society* **64**(9), 533–538.

Ashby, M., Schepp, P. & Jones, D. (2013), *Ingenieurwerkstoffe: Einführung in ihre Eigenschaften und Anwendungen*, Springer Berlin Heidelberg.
URL: *https://books.google.at/books?id=HRqaBwAAQBAJ*

Barenblatt, G. (1962), 'The mathematical theory of equilibrium cracks in brittle fracture', *Advances in Applied Mechanics* **7**, 55 – 129.
URL: *http://www.sciencedirect.com/science/article/pii/S0065215608701212*

Borchardt-Ott, W. (2009), *Kristallographie : Eine Einführung für Naturwissenschaftler*, 7. edn, Springer Berlin Heidelberg, Berlin, Heidelberg.

Camanho, P. & Hallett, S., eds (2015), *Numerical Modelling of Failure in Advanced Composite Materials*, Woodhead Publishing.

Cook, R. F. & Pharr, G. M. (1990), 'Direct observation and analysis of indentation cracking in glasses and ceramics', *Journal of the American Ceramic Society* **73**(4), 787–817.
URL: *http://dx.doi.org/10.1111/j.1151-2916.1990.tb05119.x*

Doronzo, A. & Virk, P. (2016), 'What applications are driving the strongest need for innovations in gan device technology?', *Wireless Design Development* **24**(2), 7.

Dreyer, C. E., Janotti, A. & de Walle, C. G. V. (2015), 'Brittle fracture toughnesses of gan and aln from first-principles surface-energy calculations', *Applied Physics Letters* **106**(21), 212103.
URL: *https://doi.org/10.1063/1.4921855*

Drory, M., Anger, J., Suski, T., Grzegory, I. & Porowski, S. (1996), 'Hardness and fracture toughness of bulk single crystal gallium nitride', *Applied Physics Letters* **69**(26), 4044–4046.

Gao, Y. & Bower, A. (2004), 'A simple technique for avoiding convergence problems in finite element simulations of crack nucleation and growth on cohesive interfaces', *Modelling and Simulation in Materials Science and Engineering* **12**(3), 453–463.

Gross, D. & Seelig, T. (2011), *Bruchmechanik; Mit einer Einführung in die Mikromechanik*, Springer Verlage, Berlin, Heidelberg.

Gross, D. & Seelig, T. (2016), *Bruchmechanik; Mit einer Einführung in die Mikromechanik*, Springer Verlage, Berlin, Heidelberg.

Hay, J. (2009), 'Introducition to instrumented indentation testing', *Society for Experimental Mechanics* pp. 66–72.

Hopcroft, M., Nix, W. & Kenny, T. (2010), 'What is the young's modulus of silicon?', *Journal of Microelectromechanical Systems, April 2010, Vol.19(2), pp.229-238* .

Jackson, K. A. & Schröter, W. (2000), *Handbook of Semiconductor Technology: Electronic Structure and Properties of Semiconductors*, Vol. 1, Wiley-VCH Verlag GmbH.

Jaya, B. N., Kirchlechner, C. & Dehm, G. (2015), 'Can microscale fracture tests provide reliable fracture toughness values? a case study in silicon', *Journal of Material Research* .

Jian, S.-R., Chen, G.-J. & Lin, T.-C. (2010), 'Berkovich nanoindentation on aln thin films', *Nanoscale Research Letters* **5**(6), 935.
URL: *https://doi.org/10.1007/s11671-010-9582-5*

Johanns, K. E., Lee, J. H., Gao, Y. F. & Pharr, G. M. (2014), 'An evaluation of the advantages and limitations in simulating indentation cracking with cohesive zone finite elements', *Modelling and Simulation in Materials Science and Engineering* **22**(1), 015011.
URL: *http://stacks.iop.org/0965-0393/22/i=1/a=015011*

Johnson, W. C., Parson, J. B. & Crew, M. C. (1931), 'Nitrogen compounds of gallium. iii', *The Journal of Physical Chemistry* **36**(10), 2651–2654.
URL: *http://dx.doi.org/10.1021/j150340a015*

Keller, S. (2017), Substrates and materials, *in* M. Meneghini, G. Meneghesso & E. Zanoni, eds, 'Power GaN Devices: Materials, Applications and Reliability', Power Electronics and Power Systems, Springer International Publishing, chapter 2.

Kuech, T. E. & Tischler, M. A. (2000), Epitaxial growth, *in* K. A. Jackson & W. Schroter, eds, 'Handbook of Semiconductor Technology: Processing of Semiconductors', Vol. 2, Wiley VCH Verlag GmbH, chapter 3, pp. 111–176.

Kukushkin, S., Osipov, A., Bessolov, V., Medvedev, B., Nevolin, V. & Tcarik, K. (2008), 'Substrates for epitaxy of gallium nitride: new materials and techniques', *Rev. Adv. Mater. Sci* **17**(1/2), 1–32.

Kuna, M. (2008), *Numerische Beanspruchungsanalsye von Rissen; Finite Elemente in der Bruchmechanik*, Vieweg+Teubner Verlag / GWV Fachverlage GmbH, Wiesbaden, Wiesbaden.

Lee, J. H., Gao, Y. F., Johanns, K. E. & Pharr, G. M. (2012), 'Cohesive interface simulations of indentation cracking as a fracture toughness measurement method for brittle materials', *Acta Materiala* **60**, 5448–5467.

Leguillon, D. (2002), 'Strength or toughness? a criterion for crack onset at a notch', *European Journal of Mechanics / A Solids, 2002, Vol.21(1), pp.61-72* .

Marcon, D. & Stoffels, S. (2017), Gan-on-silicon cmos-compatible process, *in* M. Meneghini, G. Meneghesso & E. Zanoni, eds, 'Power GaN Devices: Materials, Applications and Reliability', Power Electronics and Power Systems, Springer International Publishing, chapter 3.

MarketsandMarkets (2017), 'Gallium nitride semiconductor device market by device type (opto, power, rf), wafer size, application (power drives, supply and inverter, rf, lighting and laser), vertical (telecommunication, consumer, automotive), and geography - global forecast to 2023', http://www.marketsandmarkets.com/Market-Reports/gallium-nitride-gan-semiconductor-materials-devices-market-698.html.

Mishra, U. K. & Guidry, M. (2017), Lateral gan devices for power applications, *in* M. Meneghini, G. Meneghesso & E. Zanoni, eds, 'Power GaN Devices: Materials, Applications and Reliability', Power Electronics and Power Systems, Springer International Publishing, chapter 4.

Morkoç, H. (2009), *Handbook of Nitride Semiconductors and Devices : Materials Properties, Physics and Growth*, Wiley, Hoboken.

Nowak, R., Pessa, M., Suganuma, M., Leszczynski, M., Grzegory, I., Porowski, S. & Yoshida, F. (1999), 'Elastic and plastic properties of gan determined by nano-

indentation of bulk crystal', *Applied Physics Letters* **75**(14), 2070–2072.

URL: *https://doi.org/10.1063/1.124919*

Ratschinski, I., Hartmut, S., Heyroth, F., Fränzel, W., Leibiger, G. & Habel, F. (2012), 'The effect of the indenter orientation on the formation of dislocations and cracks in (0001) gan bulk crystals', *Materials Science Forum* **725**, 67–70.

Ratschinski, I., Leipner, H. S., Heyroth, F., Fränzel, W., Moutanabbir, O., Hammer, R. & Jurisch, M. (2011), 'Indentation-induced dislocations and cracks in (0001) freestanding and epitaxial gan', *Journal of Physics: Conference Series* **281**, 012007.

Ratschinski, I., Leipner, H. S., Heyroth, F., Mook, W., Michler, J., Fränzel, W., Leibiger, G. & Habel, F. (2013), 'Cracks and dislocations at vickers, berkovich and cube corner indentations in (0001) gan single crystals', *Physica stats solidi (c)* **10**(1), 76–79.

Reeber, R. R. & Wang, K. (2001), 'High temperature elastic constant prediction of some group iii-nitrides', *MRS Internet Journal of Nitride Semiconductor Research* **6**.

URL: *http://journals.cambridge.org/article_S1092578300000156*

Tsai, C.-H., Jian, S.-R. & Juang, J.-Y. (2008), 'Berkovich nanoindentation and deformation mechanisms in gan thin films', *Applied Surface Science* **254**(7), 1997 – 2002.

URL: *//www.sciencedirect.com/science/article/pii/S0169433207011932*

Turon, A., Dávila, C., Camanho, P. & Costa, J. (2007), 'An engineering solution for mesh size effects in the simulation of delamination using cohesive zone models', *Engineering Fracture Mechanics* **74**(10), 1665 – 1682.

URL: *http://www.sciencedirect.com/science/article/pii/S0013794406003808*

Ueda, D. (2017), Properties and advantages of gallium nitride, *in* M. Meneghini, G. Meneghesso & E. Zanoni, eds, 'Power GaN Devices: Materials, Applications and Reliability', Power Electronics and Power Systems, Springer International Publishing, chapter 2.

Weber, E. R., Krüger, J. & Kisielowski, C. (2000), New materials: Gallium nitride, *in* K. A. Jackson & W. Schroter, eds, 'Handbook of Semiconductor Technology: Electronic Structure and Properties of Semiconductors', Vol. 1, Wiley-VCH Verlag GmbH, chapter 13, pp. 771–808.

Weyher, J., Albrecht, M., Wosinski, T., Nowak, G., Strunk, H. & Porowski, S. (2001), 'Study of individual grown-in and indentation-induced dislocations in gan by defect-selective etching and transmission electron microscopy', *Materials Science and Engineering: B* **80**(1), 318 – 321.

URL: *http://www.sciencedirect.com/science/article/pii/S0921510700006565*

Wheeler, J., Niederberger, C., Tessarek, C., Christiansen, S. & Michler, J. (2013), 'Extraction of plasticity parameters of gan with high temperature, in situ micro-compression', *International Journal of Plasticity* **40**, 140 – 151.

URL: *//www.sciencedirect.com/science/article/pii/S0749641912001167*

Xiong, J., Tang, J., Liang, T., Wang, Y., Xue, C., Shi, W. & Zhang, W. (2010), 'Characterization of crystal lattice constant and dislocation density of crack-free gan films grown on si(1 1 1)', *Applied Surface Science* **257**(4), 1161 – 1165.
URL: *//www.sciencedirect.com/science/article/pii/S0169433210010159*

Yonenaga, I. (2001), 'Thermo-mechanical stability of wide-bandgapsemiconductors: high temperature hardness of sic, aln, gan, zno and znse', *Physica B* **308-310**, 1150–1152.

Yonenaga, I., Hoshi, T. & Usui, A. (2000), 'High temperature hardness of bulk single crystal gan', *MRS Internet Journal of Nitride Semiconductor Research* **5**(S1), 343–348.
URL: *https://www.cambridge.org/core/article/div-class-title-high-temperature-hardness-of-bulk-single-crystal-gan-div/51CA616DC3BAB0C1AF1CF535A312DF07*

Yu, G., Ishikawa, H., Egawa, T., Soga, T., Watanabe, J., Jimbo, T. & Umeno, M. (1998), 'Mechanical properties of the gan thin films deposited on sapphire substrate', *Journal of Crystal Growth* **189-190**, 701 – 705.
URL: *http://www.sciencedirect.com/science/article/pii/S0022024898002620*

Zimmermann, T., Neuburger, M., Benkart, P., Hernandez-Guillen, F. J., Pietzka, C., Kunze, M., Daumiller, I., Dadgar, A., Krost, A. & Kohn, E. (2006), 'Piezoelectric gan sensor structures', *IEEE Electron Device Letters* **27**(5), 309–312.