Die approbierte Originalversion dieser Dissertation ist in der Hauptbibliothek der Technischen Universität Wien aufgestellt und zugänglich.
http://www.ub.tuwien.ac.at

The approved original version of this thesis is available at the main library of the Vienna University of Technology.
http://www.ub.tuwien.ac.at/eng

TECHNISCHE UNIVERSITÄT WIEN

DISSERTATION

Theory of Distribution-Constrained Optimization Problems

Ausgeführt zum Zwecke der Erlangung des akademischen Grades einer Doktorin der Naturwissenschaften unter der Leitung von

Prof. Dr. Uwe Schmock

E105
Institut für Stochastik und Wirtschaftsmathematik
Forschungsbereiche Finanz- und Versicherungsmathematik

eingereicht an der Technischen Universität Wien
Fakultät für Mathematik und Geoinformation

von

M.Sc. Christiane Elgert

Matrikelnummer: 01428965
Wiedner Hauptstraße 8–10, A-1040 Wien

Wien, am 11. Februar 2019

Kurzfassung der Dissertation

Diese Dissertation befasst sich mit finanz- und versicherungsmathematischen Produkten, deren Auszahlungen von stochastischen Prozessen bestimmt werden. Der Zeitpunkt der Auszahlung ist zufällig und wird daher durch eine Stoppzeit modelliert, die Werte in einem vorbestimmten Zeitbereich annimmt. Diese Stoppzeit soll einer bestimmten Verteilung folgen und kann von dem zugrunde liegenden Auszahlungsprozess abhängen. Diese vorgegebene Verteilung enthält zusätzliche Informationen, die bekannt sind oder auf die man Zugriff hat. Das Ziel ist es, eine Abschätzung für den schlechtesten Fall, also der Worst-Case-Situation, abzuleiten. Dies ergibt sich aus dem Supremum der zu erwarteten Auszahlung über alle Stoppzeiten, die die angegebene Randbedingung erfüllen. Es liegt im besonderem Interesse, eine optimale Stoppzeit zu finden, die diesen Maximalwert annimmt. Dieses Problem soll als OptStop^{τ} bezeichnet werden. Eine Erweiterung besteht in der Verwendung von adaptierten zufälligen Wahrscheinlichkeitsmaßen anstelle von Stoppzeiten. Das dazugehörige Problem wird mit OptStop^{γ} bezeichnet. Aus mathematischer Sicht ist das betrachtete Problem eine spezielle und erweiterte Version eines optimalen Stoppproblems. Zum ersten Mal wurden die Probleme OptStop^{τ} und OptStop^{γ} in [33] betrachtet, was der Ausgangspunkt dieser Arbeit war. In [33] werden drei Hauptannahmen an den stochastischen Auszahlungsprozess gestellt, um ein wohldefiniertes Problem zu garantieren. Diese drei wichtigsten Annahmen sind die fast-sichere Endlichkeit des Supremums der Beträge der Elemente des Prozesses, die Endlichkeit des Erwartungswertes und die gleichgradige Integrierbarkeit des Prozesses. Die hier vorliegende Arbeit beinhaltet eine Verallgemeinerung und berücksichtigt sensitivere Bedingungen. Diese Bedingungen machen sich die Struktur und die Informationen, die sich aus der Verteilungseinschränkung an die Stoppzeiten oder adaptierten zufälligen Wahrscheinlichkeitsmaßen ergeben, tatsächlich zunutze. Man kann das Problem aus verschiedenen Anwendungsbereichen motivieren und es gibt eine weitere Betrachtungsmöglichkeit. Für diesen Ansatz wird die Aufgabe als optimales Transportproblem neu formuliert und aus Sicht des Transports von Massen betrachtet. Dieses Problem wird dann als OptStop^{π} bezeichnet.

Im ersten Teil dieser Arbeit werden wir die Ergebnisse für die adaptierten Abhängigkeiten in diskreter Zeit herleiten, d.h. wir betrachten eine vollständig geordnete, abzählbare Indexmenge. Es werden die verschiedenen Optimierungsprobleme OptStop^{τ}, OptStop^{γ} und OptStop^{π} eingeführt und deren Zusammenhänge schrittweise erarbeitet und beschrieben. Der Schwerpunkt dieses Teils liegt auf dem Problem OptStop^{γ} und damit auf adaptierten zufälligen Wahrscheinlichkeitsmaßen. Ein wichtiger Aspekt bei der Betrachtung solcher Probleme ist die Frage nach einer optimalen Strategie. In dieser Arbeit wird die Existenz einer solchen optimalen Strategie in diskreter Zeit für den verallgemeinerten Ansatz bewiesen, der es erlaubt, eine größere Anzahl möglicher Prozesse zu berücksichtigen.

Diese optimale Strategie ist im Allgemeinen nicht eindeutig, wie anhand einiger Beispiele erläutert wird. Darüber hinaus werden in dieser Dissertation einige Schranken angegeben. Einige davon gelten für das Problem im Allgemeinen, während andere von der Struktur des zugrunde liegenden Prozesses abhängen. Für bestimmte Klassen stochastischer Prozesse ist es möglich, eine optimale Strategie und somit den daraus resultierenden Wert des Optimierungsproblems zu bestimmen. Dazu gehören beispielsweise Prozesse, die als Produkt eines Martingals und einer deterministischen Funktion oder im Binomialmodell gegeben sind. Zusätzlich wird eine Anwendung des Problems im Bereich der fondsgebundenen Lebensversicherungen diskutiert, in der die Modellierung des Vertrags ohne Annahme der Unabhängigkeit zwischen biometrischen Risiken und Finanzmarktrisiken erfolgt. Der letzte Abschnitt dieses Teils behandelt dann das Problem OptStop$^\pi$. Wir fomulieren dazu das Problem als optimales Transportproblem und zeigen die Existenz einer optimalen Strategie mithilfe der Theorie des optimalen Transports. Es werden auch hier Beispiele betrachtet.
Im zweiten Teil dieser Arbeit werden wir die Ergebnisse für die adaptierten Abhängigkeiten in kontinuierlicher Zeit herleiten. Um das Problem OptStop$^\gamma$ in kontinuierlicher Zeit zu betrachten, müssen die adaptierten zufälligen Wahrscheinlichkeitsmaße durch stochastische Übergangskerne ersetzt werden. Es wird eine diskrete Approximation angegeben, mit deren Hilfe die in diskreter Zeit gefundenen Ergebnisse übertragen werden können. Außerdem werden Ergebnisse für den Spezialfall hergeleitet, in dem die Prozesse als Produkt eines Martingals und einer deterministischen Funktion gegeben sind. Der Hauptabschnitt dieses Teils befasst sich jedoch mit dem Problem OptStop$^\pi$. Unter Verwendung der Methoden und Techniken aus der optimalen Transporttheorie erhalten wir die Existenz einer optimalen Stoppzeit einer Brown'schen Bewegung mit vorgegebenen Randverteilungen. Dazu muss der Kostenprozess jedoch mindestens messbar und angemessen adaptiert sein. Gewisse Stetigkeitssannahmen garantieren dann die Existenz von Lösungen des betrachteten Problems. Des Weiteren werden Ideen und Konzepte aus dem optimalen Transport (und seiner Martingalvariante) angepasst, um eine geometrische Beschreibung der optimalen Strategie zu erhalten. Die Methoden sind auf eine große Klasse an Kostenprozessen anwendbar und es wird gezeigt, dass für viele Kostenprozesse eine Lösung durch die erste Trefferzeit einer Barriere in einem geeigneten Phasenraum gegeben ist. Die Ergebnisse dieses Abschnitts der Arbeit sind bereits in [10] veröffentlicht.

Abstract

This thesis deals with financial and actuarial products whose payouts are driven by stochastic processes. The time point of the payouts is random and is therefore modeled by a stopping time that is taking values in a predetermined time domain. This stopping time should follow a given distribution and may depend on the underlying process modeling the payouts. The given distribution contains additional information that is known to us or to which we have access. Our target is to deduce the estimation of the worst-case situation. This results from the supremum of the expected payout over all stopping times satisfying the given marginal. It is of particular interest to find an optimal stopping time that yields this maximal value. This problem is denoted by $\textsc{OptStop}^{\tau}$. An extension is the use of adapted random probability measures instead of stopping times. The problem involved is called $\textsc{OptStop}^{\gamma}$. From a mathematical point of view, the problem being considered is a special and extended version of an optimal stopping problem. For the first time the problems $\textsc{OptStop}^{\tau}$ and $\textsc{OptStop}^{\gamma}$ were introduced in [33], which was the starting point of this work. In [33] three main assumptions are made of the stochastic payout process to guarantee a well-defined problem. These three main assumptions are that the supremum of the absolute values of the elements of the process is almost surely finite, that it has finite expectation, and that the process is uniformly integrable. This thesis contains a generalization and takes more sensitive conditions into account that really take advantage of the structure and information resulting from the distributional restriction of the stopping times or adapted random probability measures. We can motivate the problem from different application areas and there is another way to describe the problem. For this approach, the task is reformulated as an optimal transport problem and discussed from a mass transport perspective. The problem is denoted as $\textsc{OptStop}^{\pi}$.

In the first part of this thesis we will derive the results for the adapted dependence in discrete time, i.e., we consider a totally-ordered countable index set. The different optimization problems $\textsc{OptStop}^{\tau}$, $\textsc{OptStop}^{\gamma}$ and $\textsc{OptStop}^{\pi}$ will be introduced and the various connections between them are gradually worked out and described. The focus of this part is on the problem $\textsc{OptStop}^{\gamma}$ and therefore on adapted random probability measures. An important aspect in considering this problem is the question of an optimal strategy. In this thesis, the existence of such an optimal strategy in a discrete time setting is proven for the generalized approach, which allow us to consider a much larger set of possible processes. This optimal strategy is not unique in general as illustrated by some examples. In addition to this, some bounds are derived in this dissertation. Some of them apply to the problem in general, while others depend on the structure of the underlying process. For certain classes of stochastic processes, it is possible to find an optimal strategy and the resulting value of the optimization problem. These include, for example, processes that are the product of a martingale and a deterministic function or in

the binomial model. In addition, an application of the problem in the area of unit-linked life insurance is discussed in which the modeling of the contract takes place without assuming independence between biometric and financial market risks. The last section of this part deals with the problem $\textsc{OptStop}^{\pi}$. We will formulate the problem as an optimal transport problem and show the existence of an optimal strategy by using the theory of optimal transport. Examples are also considered here.
In the second part of this thesis we will derive the results for the adapted dependence in continous time. To view the problem $\textsc{OptStop}^{\gamma}$ in continuous time, the adapted random probability measures have to be replaced by stochastic transition kernels. A discrete approximation is given, with the aid of which the results found in discrete time can be transferred. In addition, results are derived for the special case in which the processes are given as the product of a martingale and a deterministic function. However, the main section of this part deals with the problem $\textsc{OptStop}^{\pi}$. Using the methods and techniques of optimal transport theory we obtain the existence of optimal stopping times of a Brownian motion with given marginal. For this, the cost process must be at least measurable and appropriately adapted. Certain continuity assurances then guarantee the existence of solutions to the considered problem. Furthermore, ideas and concepts from the optimal transport (and its martingale variant) are adapted to obtain a geometric description of the optimal strategies. The methods work for a large class of cost processes and it is shown that for many cost processes a solution is given by the first hitting time of a barrier in a suitable phase space. As a by-product we recover classical solutions of the inverse first passage time problem / Shiryaev's problem. The results of this section of the thesis are already published in [10].

Acknowledgments

There is a nice metaphor about knowledge whose author is unknown to me. I want to reproduce it in my own words: Your own knowledge is like a ball in the space of all knowledge. If you increase your knowledge, the ball will grow und thus its volume. At the same time also the surface of your ball is growing, but with it more and more points of contact arise to the knowledge that you do not have. It is becoming clearer to you that you basically do not know anything, which is a bit paradox. Happenstance determines what knowledge we acquire as well as how it affects our lives. The time of my PhD was marked by a lot of randomness, especially the last year, and not only in the sense that I studied this subject from the mathematical point of view. This thesis could not have been finalized without the support of many people. Thus, I would like to take this opportunity to express my gratitude to everyone who supported me throughout this period, no matter on which way.

First of all I would like to thank my supervisor Uwe Schmock for introducing me to the subject of this thesis, sharing his knowledge and ideas and teaching me some useful lessions. He also taught me the value of working under my own responsibility. I also would like to thank Mathias Beiglböck, Manu Eder and Gudmund Pammer for very helpful discussions about the part of the thesis dealing with optimal transport and the joint work. Also, I want to thank my third referee Jan Palczewski for taking the time to review my thesis.

Furthermore, I would like to extend my thanks to the people who introduced me to the world of mathematics and showed me the beauty of it: my school teachers Kerstin Häfke and Kerstin Kral. Without them and their support I would have never studied mathematics in the first place.

A special thanks goes to all my colleagues and friends at FAM, for making the past years so much more enjoyable. In particular, I want to thank on the one hand Sandra Trenovatz, for always taking time when I needed help, and on the other hand Tilmann Blümmel and Cetin Gülüm.

Further I would like to thank the *OeNB's Anniverary Fund* for financial support under grant no. 16549, *arithmetica Consulting GmbH* under the project “Approximation von aggregierten Risiken unter Nebenbedingungen” and the *DAAD* for the scholarship.

Finally, I am *extremely* thankful to my closest friends and all the awesome people who I met here in Vienna for making my life what it is. I know I can always count on the important ones. Especially I also want to say thank you to my family.

Es gibt Fehlschläge, die uns kämpferischer machen, und solche die uns weiser machen. Und dann gibt es Fehlschläge, die uns offen machen für Neues.

Contents

1 Introduction

There are many situations in financial and actuarial mathematics where independence is assumed for two stochastic components. It is often questionable whether this assumption is always justified. This work presents a general framework to show how some of these situations can be handled without the assumption of independence.

We want to deal with distribution-constrained optimization problems and the corresponding theory. But what is that? Consider financial and actuarial products whose payoffs are determined by a stochastic process. The time point of the payouts is random and is therefore modeled by a stopping time that takes values in a predetermined time domain. This stopping time should follow a given distribution and may depend on the underlying process modeling the payoffs. The given distribution contains additional information that is known to us or to which we have access. Then we are concretely interested in the deduction of estimations for the worst-case situation. This results from the supremum of the expected payout over all stopping times satisfying the given marginals. It is of particular interest to find an optimal stopping time that yields this maximal value. This problem is denoted by OptStop$^{\tau}$. From a mathematical point of view, the problem being considered is a special and extended version of an optimal stopping problem. This is only one possible description of the problem. But before we look at the others, we want to motivate it more closely with an example. Let us take a look on unit-linked life insurances, for more details see Section 3.6.

First, we consider a single unit-linked life insurance contract with payoff at the end of the year of death of the insured x-year old person or at the end of the contract. We assume this contract runs for 30 years and a payout can always be made at the end of the year, such that our predetermined time interval is given by $I = \{0, \ldots, 30\}$. Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \in I}, \mathbb{P})$ be a filtered probability space and let $(Z_t)_{t \in I}$ be the process of the payouts. The family of the insured person will get the insurance benefit Z_τ at an random time point τ after paying advance premiums. In this case, the random time point τ is an stopping time and the insurance companies are interested in the maximal expected value $\mathbb{E}[Z_\tau]$ of the payouts of the contract. This would be a classical optimal stopping problem with value $V_{\mathcal{T}}(Z)$. However, we have more information. The stopping time τ is modeled as the minimum of the maturity $T = 30$ and the future lifetime T_x of the insured person, where x indicates the age at conclusion of contract. Thus this stopping time has the distribution which is given through the life table or through the termination of a contract. That is how we would like to get our distributional constraint ν from the life table. With the given distribution ν, we

consider then a distribution-constrained optimal stopping problem OptStop$^\tau$ with value $V_{\mathcal{T}}^{\nu}(Z)$.
As already described in [33], in the praxis for unit-linked life insurances it is normally assumed that financial and biometric risks are independent, see e.g. [43]. Under the assumption of independence between Z and τ we will get the value $V_{\text{ind}}^{\nu}(Z)$. Examples show that the two values $V_{\text{ind}}^{\nu}(Z)$ and $V_{\mathcal{T}}^{\nu}(Z)$ can differ greatly from each other.
If surrender of the contract is allowed, this reason for dropping out, which also leads to a payoff, should not be set independent of the financial market. It is possible that a downturn in the economy, which is often followed by high unemployment rates, leads to more lapses for an insurance company. It is equally conceivable that a flu epidemic could influence the financial markets. In the technical specifications of the long-term guarantees assessment (LTGA, [49]) or the fifth quantitative impact study (QIS5, [60]) for Solvency II there are assumptions about a positive correlation between financial and biometric risks used to compute the solvency capital requirement. Similar ideas are followed in current research. In [19], worst-case scenarios for pricing and reserving life insurance products are considered where a mutual dependence between interest rates and mortality is allowed. In [47], a valuation framework is presented with a given correlation between the dynamics of mortality and interest rates. Further upper and lower bounds for the value of a guaranteed annuity option are found using comonotonicity theory. Variable annuities are very flexible, long-term tax-deferred investments whose design matches features of unit-linked life insurance contracts that package several types of options and guarantees, at the policyholder's discretion. In [7], a quite general valuation model for variable annuities, with death and survival guarantees and state-dependent fee structure, along the lines of [6], is defined and numerically analysed the interaction between fee rates, death/survival guarantees, fee thresholds and surrender penalties under alternative model assumptions and policyholder behaviors, thus getting also some interesting insights into the model risk. In [6], it is shown that in some situations, namely when the guarantee concerns the choice of the post-retirement income, policyholder preferences significantly affect the value of the guarantee.
Let us now consider the unit-linked life insurance of a married couple or a group of persons. In addition to the assumption of independence of financial risks and biometric risks, in this case, the independence of the physical and emotional health of the partners from each other is often assumed. Then we calculate the expected values for each individual person by means considered above and add them together. But with some common sense, it is clear that this is not the case. The couple live in the same environment and is strongly connected. For example, both can get be injured in a possible car accident. Furthermore, the broken heart syndrome is also known and studied since a long time in medicine, see [26]. They found that the mortality rate of bereaved close relatives is much greater within a year of bereavement compared with a control group. As a consequence, health can drastically deteriorate when one's partner dies. Therefore, it is not reasonable to assume independence of the times of death of either partner. Current research, that is concerned with this, is for example [50].
To model a portfolio of similar contracts in a discrete time setting we have to use adapted random probability measures instead of stopping times. From a mathematical point of view, the problem being considered then is a special and extended version of an optimal stopping problem and is called OptStop$^\gamma$. The problems OptStop$^\tau$ and OptStop$^\gamma$ were introduced in [33] for the first time. If a stopping time τ is used for modeling a life

insurance contract for one person, then the adapted random probability measure can be used to model a married couple.
It is also possible to use this setting for health insurance contracts. These are often modeled in a similar way as life insurance contracts. The payoff for these contracts, called claims amount per risk in this setting, is normally a deterministic number, corresponding to the value the insurer expects to pay, and based on historical data. Using the setting of this article, such claims amount per risk can be modeled stochastically. This is more appropriate, since it is influenced by many factors, such as modern techniques in health care, the status of the corresponding country (social turmoil, peace or war, ...) and political decisions. These factors also influence the probability of occurrence of an insured event. Improvements in the medical system will guarantee that people are cured more rapidly and that the probability of a relapse declines.
The dependence between severe medical diseases and crises or catastrophes in the surroundings of patients is a matter of paramount interest to medical research. One especially interesting work with regard to this thesis is about the impact of the socioeconomic crisis in Greece on acute myocardial infarction [51]. In [51] the authors found that the financial crisis may have led to a higher incidence of acute myocardial infarction in the population of Messinia and assert the need for an analysis of this phenomenon for the entire Greek population. In [44] and [54] the aftermath of the earthquake in Japan in March 2011 on coronary syndromes is analyzed. Both studies seem to demonstrate that the stress of this disaster has increased the number of hospitalized patients. Similarly, an alteration in the pattern of acute myocardial infarction onset followed in the wake of hurricane Katrina in New Orleans. This is discussed in [59].
As seen, this issue has a lot of applications in the field of financial and actuarial risk management. We have only introduced the two problems $\textsc{OptStop}^{\tau}$ and $\textsc{OptStop}^{\gamma}$ so far, but there is another way to describe the problem. For this approach, the task is reformulated as an optimal transport problem and discussed from a mass transport perspective. The problem is then denoted as $\textsc{OptStop}^{\pi}$. IIf we deal with the theory of optimal transport, we come into contact with the two common basic concepts: cyclical monotonicity and Kantorovich duality. The cyclical monotonicity is a geometric property. An optimal plan should be c-cyclically monotone, i.e., it is concentrated on a c-cyclically monotone set and you can not improve the cost by rerouting mass along some cycle. It is impossible to perturb it and get something more economical. Informally, a c-cyclically monotone plan is a plan that cannot be improved. The converse property is considerably less obvious, i.e., c-cyclically monotone plan should be optimal. Maybe it is possible to get something better by radically changing the plan as only rerouting mass along some cycle. In this work we will see that it holds true under certain conditions. The Kantorovich duality is used to show the existence of an optimal strategy.
In the first part of this thesis we will derive the results for the adapted dependence in discrete time. The different optimization problems $\textsc{OptStop}^{\tau}$, $\textsc{OptStop}^{\gamma}$ and $\textsc{OptStop}^{\pi}$ will be introduced and the various connections between them are gradually worked out and described. The focus of this part is on the problem $\textsc{OptStop}^{\gamma}$ and consequently on adapted random probability measures. An important aspect in considering this problem is the question of an optimal strategy. In this thesis, the existence of such an optimal strategy in a discrete time setting is proven for the generalized approach, which allows us to consider a much larger set of possible processes. This optimal strategy is not unique in general as illustrated by some examples. In addition to this, some bounds are derived in

this dissertation. Some of them apply to the problem in general, while others depend on the structure of the underlying process. For certain classes of stochastic processes, it is possible to find an optimal strategy and the resulting value of the optimization problem. These include, for example, processes that are the product of a martingale and a deterministic function or in the binomial model. In addition, an application of the problem in the area of unit-linked life insurance is discussed in which the modeling of the contract takes place without assuming independence between biometric and financial risks. The last section of this part deals with the problem $\textsc{OptStop}^{\pi}$. We will formulate the problem as an optimal transport problem and show the existence of an optimal strategy by using the theory of optimal transport. Examples are also considered here.
In the second part of this thesis we will derive the results for the adapted dependence in continous time. To view the problem $\textsc{OptStop}^{\gamma}$ in continuous time, the adapted random probability measures have to be replaced by stochastic transition kernels. A discrete approximation is given, with the aid of which the results found in discrete time can be transferred. In addition, results are derived for the special case in which the processes are given as the product of a martingale and a deterministic function. However, the main section of this part deals with the problem $\textsc{OptStop}^{\pi}$. Using the methods and techniques of optimal transport theory we obtain the existence of optimal stopping times of a Brownian motion with given marginals. However, the cost process must be at least measurable and appropriately adapted. Certain continuity assurances then guarantee the existence of solutions of the considered problem. Furthermore, ideas and concepts from the optimal transport (and its martingale variant) are adapted to obtain a geometric description of the optimal strategies. The methods work for a large class of cost processes and it is shown that for many cost processes a solution is given by the first hitting time of a barrier in a suitable phase space. As a by-product we recover classical solutions of the inverse first passage time problem / Shiryaev's problem. The results of this section of the thesis are already published in [10].

I.

Adapted Dependence in Discrete Time

2

The Problem

We now introduce our problem of study and create a connection to the classical optimal stopping problem. The distribution-constrained optimization problem, which we consider, is a modified version of an optimal stopping problem.
We deal with financial and actuarial products, whose payoffs taking value during a certain time interval are determined by an stochastic process. The time point of the payoff is modeled by a stopping time. This stopping time or adapted random probability measure follows a given distribution and can depend on the underlying process of payoff. With other words we want to consider distribution-constrained optimal stopping problems. Our target is to deduce the estimation of the worst-case situation. That means, the supremum of the expected payoff over all stopping times satisfying the given marginals. It may happen that this problem is not well-posed and therefore does not have a solution.
There are different views on these restricted optimization problem. One possibility is that we replace the stopping times by adapted random probability measures and prove the existence of an optimal strategy for these problems by using functional analysis, for more details see Chapter 3. Another possibility is that we formulate the problem in terms of an optimal transport problem and prove the existence of an optimal strategy for these problems by using the theory of optimal transport, see Chapter 4.

First of all, we introduce the notational conventions for this part. We will see that the problem is a modified version of an optimal stopping problem.

Notation 2.0.1. Throughout this part, we consider a discrete-time setting and stick to the following notation.

(a) Let $I \neq \emptyset$ denote a countable, i.e., a finite or countably infinite, totally-ordered index set; for simplicity the reader may assume that $I \subseteq \mathbb{R} \cup \{-\infty, \infty\}$. Let $T := \sup(I)$ and $\bar{I} := I \cup \{T\}$.

(b) For $t \in I$ we define the set $I_{<t} = \{s \in I \,|\, s < t\}$ of all times before t, the set $I_{\leq t} = \{s \in I \,|\, s \leq t\}$ of all times up to t, the set $I_{\geq t} = \{s \in I \,|\, s \geq t\}$ of all times from t on, and the set $I_{>t} = \{s \in I \,|\, s > t\}$ of all times after t. The same holds for $\bar{I}$.

(c) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in I}$. If $T \notin I$, we define $\mathcal{F}_T = \sigma\Big(\bigcup_{t \in I} \mathcal{F}_t\Big)$.

(d) The given probability measure on I will be denoted by $\nu = (\nu_t)_{t \in I}$, its support by supp(ν) := $\{t \in I \mid \nu_t > 0\}$. For $t \in I$ we define $\nu_{<t} = \sum_{s\in I_{<t}} \nu_s$ and $\nu_{>t} = \sum_{s\in I_{>t}} \nu_s$, as well as $\nu_{\leq t} = \nu_{<t} + \nu_t$ and $\nu_{\geq t} = \nu_{>t} + \nu_t$. Relations like $\nu_{\leq t} + \nu_{>t} = 1$ will be used without mentioning it. For $(\gamma_t)_{t\in I}$ as in Definition 3.1.1 below, we use a similar notation.

(e) $\mathcal{T}_I$ denotes the set of all stopping times $\tau : \Omega \to \bar{I}$ with $\mathbb{P}(\tau \in I) = 1$ and $\mathcal{T}_I^{\nu}$ its subset of all τ with distribution ν, i.e., $\mathcal{L}(\tau) = \nu$. Note that we use in this thesis another definition as [33, Definition 2.4].

Typical examples for an infinite index set I are $\mathbb{N}$, $\mathbb{Z}$ or $\mathbb{Q}$ and for a finite $\{1,\dots,N\}$ for some $N \in \mathbb{N}$. Our considered index set I is especially a directed set, which satisfies the following definiton.

Definition 2.0.2 (Directed set)**.**
A directed set is a partially ordered set with the additional property that two elements in the set have a common upper bound in the set, respectively

Furthermore, note the difference between a mapping $\sigma : \Omega \to I$ such that $\{\sigma = t\}$ is $\mathcal{F}$-measurable for each $t \in I$ and a stopping time. A stopping time τ with respect to the filtration $(\mathcal{F}_t)_{t\in I}$ is a mapping $\tau : \Omega \to I$ which satisfies the measurability property $\{\tau = t\} \in \mathcal{F}_t$ for each t in I. Corresponding to each stopping time τ there is a τ-field denoted by $\mathcal{F}_\tau$ and defined as

$$\mathcal{F}_\tau := \{A \in \mathcal{F} \mid A \cap \{\tau = t\} \in \mathcal{F}_t \text{ for all } t \in I\}.$$

We now assume that the values we are interested in exist and are finite. If we want to make sure that the values exist, we can assume that $\mathbb{E}[\sup_{t\in I} |Z_t|] < \infty$ or that we are given an adapted process Z in $L^1(\mathbb{P})$ with $\mathbb{E}[\sup_{t\in I} Z_t^+] < \infty$ or $\mathbb{E}[\sup_{t\in I} Z_t^-] < \infty$. Note that such assumptions are very strong and do not use the given information about the distribution on I.
The value of a classical optimal stopping problem, which we will denote by $V_{\mathcal{T}}(Z)$, is given by

$$V_{\mathcal{T}}(Z) := \sup_{\tau\in\mathcal{T}_I} \mathbb{E}[Z_\tau].$$

Note that $\mathcal{T}_I$ is defined slightly different as in [33] and therefore this value too, cf. equation (2.5) in [33]. This value coincides for a non-negative process Z with the value of a standard American option without any hedging possibilities. The pricing of American options or optimal stopping problems are well known problems in the literature. An example of the calculation via Snell envelope can be found in Section 3.6.
Keeping such classical optimal stopping problems in mind, the difference to the following one is our assumption that we have considered some information about the distribution of the stopping times. Then the distribution-constrained optimal stopping problem is given in the following way:

Problem (OptStop$^\tau$)**.** Consider a real-valued and $\mathbb{F}$-adapted stochastic process $Z = (Z_t)_{t\in I}$ such that $\mathbb{E}[Z_\tau^+]$ is finite for all $\tau \in \mathcal{T}_I^{\nu}$. Find sufficient conditions such that among all stopping times $\tau \in \mathcal{T}_I^{\nu}$ there exists a maximizer τ^* solving

$$\mathbb{E}[Z_{\tau^*}] = \sup_{\tau\in\mathcal{T}_I^{\nu}} \mathbb{E}[Z_\tau].$$

$\mathcal{T}_I^{\nu}$ and $V_I^{\nu}(Z)$ are again defined slightly diffferent as in [33], cf. equation (2.7) there. In general we can not expect a maximizer τ^* to be unique. As example, consider a uniformly integrable martingale Z with index set $I = \mathbb{N}$. By Doob's optional stopping theorem, every stopping time $\tau \in \mathcal{T}_{\mathbb{N}}$ gives the same value for $\mathbb{E}[Z_\tau]$.
It may happen, in particular on a finite space Ω, that OptStop$^\tau$ is not well posed and therefore does not have a solution because the filtration is so small that $\mathcal{T}_I^{\nu} = \emptyset$. This happens for example when there exists a $t \in I$ such that no event $A \in \mathcal{F}_t$ satisfies $\mathbb{P}(A) = \nu_t$, cf. Example 2.0.3 below, or when there is an $A_t \in \mathcal{F}_t$ with $\mathbb{P}(A_t) = \nu_t$ for each $t \in I$, but it is impossible to have $A_s \cap A_t = \emptyset$ for all $s, t \in I$ with $s \neq t$, cf. Example 2.0.4 below.

Example 2.0.3. Cf. [33, Example 2.21]: Given a one-period model with $I = \{0,1\}$, we consider the probability space $\Omega = \{0,1\}$ with $\mathcal{F} = \mathcal{P}(\Omega) = \{\emptyset, \{0\}, \{1\}, \Omega\}$ and $\mathbb{P}(\{\omega\}) \in (0,1)$ for all $\omega \in \Omega$. Let the filtration be given by $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_1 = \mathcal{F}$. For every probability distribution ν on I with $\nu_0 \in (0,1)$ there does not exist a stopping time τ with $\{\tau = 0\} \in \mathcal{F}_0$ and $\mathbb{P}(\tau = 0) = \nu_0$. Consequently OptStop$^\tau$ is not a well-posed problem and cannot be solved.

Example 2.0.4. Given a two-period model with $I = \{0,1,2\}$, we consider the probability space $\Omega = \{0,1,3\}$ with $\mathcal{F} = \mathcal{P}(\Omega)$ and $\mathbb{P}(\{\omega\}) = \frac{1}{3}$ for all $\omega \in \Omega$. Let the filtration be given by $\mathcal{F}_0 = \mathcal{F}_1 = \{\emptyset, \{0\}, \{0,1\}, \Omega\}$ and $\mathcal{F}_2 = \mathcal{F}$. For the probability distribution ν on I with $\nu_0 = \nu_1 = \nu_2 = \frac{1}{3}$ there does not exist a stopping time τ with $\{\tau = 0\} \in \mathcal{F}_0$ and $\{\tau = 1\} \in \mathcal{F}_1$ satisfying $\mathbb{P}(\tau = 0) = \mathbb{P}(\tau = 1) = \frac{1}{3}$, because there is only one event in $\mathcal{F}_1$ with the correct probability. Thus OptStop$^\tau$ is not a well-posed problem and can not be solved.

Assume that the expected values of interest are well-defined. Then the connection between the standard and distribution-constrained optimal stopping problem is given for a process Z by

$$V_{\mathcal{T}}^{\nu}(Z) := \sup_{\tau \in \mathcal{T}_I^{\nu}} \mathbb{E}[Z_\tau] \leq \sup_{\tau \in \mathcal{T}_I} \mathbb{E}[Z_\tau] =: V_{\mathcal{T}}(Z). \tag{2.0.5}$$

Note that we set $\sup_{\tau \in \mathcal{T}_I^{\nu}} \mathbb{E}[Z_\tau] = -\infty$ in the case $\mathcal{T}_I^{\nu} = \emptyset$. It implies the existence of the value $V_{\mathcal{T}}^{\nu}(Z)$ whenever the corresponding classical optimal stopping problem is well-defined. For further information about the value $V_{\mathcal{T}}(Z)$ we refer to the corresponding literature about optimal stopping problems.
If we assume that the adapted process Z and the stopping time $\tau \in \mathcal{T}_I^{\nu} \neq \emptyset$ are independent, then we receive by using Corollary 3.5.1 below that

$$V_{\text{ind}}^{\nu}(Z) := \mathbb{E}[Z_\tau] = \sum_{t \in I} \mathbb{E}[Z_t \mathbb{1}_{\{\tau = t\}}] = \sum_{t \in I} \mathbb{E}[Z_t] \nu_t$$

with $\nu_t := \mathbb{P}(\tau = t)$ for all $t \in I$. If such an independent stopping time $\tau \in \mathcal{T}_I^{\nu}$ does not exist, then we set $V_{\text{ind}}^{\nu}(Z) = -\infty$.
Like in [33], it is easy to see that $V_{\text{ind}}^{\nu}(Z) \leq V_{\mathcal{T}}^{\nu}(Z) \leq V_{\mathcal{T}}(Z)$. If $I \subseteq \mathbb{N}_0$ is a discrete interval with $0 \in I$, the process Z is a uniformly integrable martingale and $\mathcal{T}_I^{\nu} \neq \emptyset$, then $V_{\mathcal{T}}^{\nu}(Z) = V_{\mathcal{T}}(Z)$, because of Doob's optional stopping theorem (given below) that states that $\mathbb{E}[Z_\tau] = \mathbb{E}[Z_0]$ for all stopping times τ. Furthermore, if there exists a stopping time $\tau \in \mathcal{T}_I^{\nu} \neq \emptyset$, which is independent of the uniformly integrable martingale Z, then this stopping time proves $V_{\text{ind}}^{\nu}(Z) = V_{\mathcal{T}}^{\nu}(Z) = V_{\mathcal{T}}(Z)$. Of course these equalities are also true

for all martingales Z and stopping times τ that satisfy the necessary conditions for using Doob's optional stopping theorem. The different conditions on the martingale and the stopping time in Doob's optional stopping theorem are noted, for example Theorem 2.0.10 in [74] or Theorem 2.0.11 below.

Theorem 2.0.6 (Doob's optional stopping theorem).
Cf. [33, Theorem 2.9]:

(a) *Given $I \subseteq \mathbb{N}_0$ with $0 \in I$. Let τ be a stopping time and Z be a supermartingale. Then Z_τ is integrable and $\mathbb{E}[Z_\tau] \leq \mathbb{E}[Z_0]$ in each of the following situations:*

1. *τ is bounded a.s., i.e., for some $N \in I \setminus \{0\} \subseteq \mathbb{N}$ we have $\mathbb{P}(\tau \leq N) = 1$,*
2. *τ is finite a.s. and Z is bounded a.s., i.e., for some $K > 0$ we have $\mathbb{P}(|Z_t| \leq K) = 1$ for all $t \in I$,*
3. *$\mathbb{E}[\tau] < \infty$ and for some $K > 0$ we have $\mathbb{P}(|Z_t - Z_s| \leq K|t-s|) = 1$ for all $s, t \in I \setminus \{0\}$.*

(b) *If $I \subseteq \mathbb{N}_0$ with $0 \in I$, any of the conditions (a1), (a2), (a3) or (a4) holds and Z is a martingale, then $\mathbb{E}[Z_\tau] = \mathbb{E}[Z_0]$.*

(c) *If $I \subseteq \mathbb{Z}$ is a countably infinite index set, Z is a martingale and τ is a bounded stopping time, then Z_τ is integrable and $\mathbb{E}[Z_\tau] = \mathbb{E}[Z_t]$ for all $t \in I$.*

(d) *Given a totally ordered countable set $I \subseteq \mathbb{R}$, let ν be a probability distribution on I and let τ be a stopping time.*

1. *ν has a finite support. Let Z be a supermartingale. Then, for every $\tau \in \mathcal{T}_I^{\nu}$, the random variable Z_τ is well-defined, integrable and satisfies $\mathbb{E}[Z_\tau] \leq \mathbb{E}[Z_t]$ for every $\min_{\omega \in \Omega} \tau(\omega) \geq t$, $t \in I$.*
2. *Let Z be a martingale with $Z^* = \sup_{t \in I} |Z_t| \in L^1$. Then, for every $\tau \in \mathcal{T}_I^{\nu}$, the random variable Z_τ is well-defined, integrable and satisfies $\mathbb{E}[Z_\tau] = \mathbb{E}[Z_t]$ for every $t \in I$.*
3. *Let Z be a supermartingale, $\sum_{t \in I} |t| \nu_t < \infty$ and for some $K > 0$ we have $|Z_t - Z_s| \leq K|t-s|$ a.s. for all $s, t \in I$. Then, for every $\tau \in \mathcal{T}_I^{\nu}$, the random variable Z_τ is well-defined, integrable and satisfies $\mathbb{E}[Z_\tau] \leq \mathbb{E}[Z_t]$ for every $\min_{\omega \in \Omega} \tau(\omega) \geq t$, $t \in I$.*

Before we prove this theorem, let us start with some preliminary considerations. There are a lot of different versions of Doob's optional sampling theorem and different ways to get it.
The paper [74] gives a characterization of the class of stopping times for which the optional sampling theorem is true for all uniformly bounded submartingales indexed by countable partially ordered set. A totally ordered countable index set is a special countable partially ordered set, so that the results in [74] remain also in the special index set I.
Before we come to the main results of this paper, we want to define uniformly boundedness.

Definition 2.0.7 (Uniformly bounded).
A mapping $X : I \times \Omega \to \mathbb{R}^d$, $d \in \mathbb{N}$, is uniformly bounded if there exists a non-negative random variable $X_+ : \Omega \to \mathbb{R}_+$ with finite expectation $\mathbb{E}[X_+]$ such that $\|X_t\| \leq X_+$ for all $t \in I$, where $\|\cdot\|$ is a norm on $\mathbb{R}^d$.

Theorem 2.0.8. *See [74]: For a given pair σ, τ of stopping times such that $\sigma \leq \tau$ on a countable partially ordered set, the optional sampling inequality*

$$X_\sigma \leq \mathbb{E}[X_\tau | \mathcal{F}_\sigma], \quad a.s.,$$

is true for all uniformly bounded submartingales X if and only if τ is reachable from σ.

Example 2.0.9. Cf. [74]: Let $I = \{a, b, c\}$ with the order relation $a \leq b$ and $a \leq c$. Let τ be a random function taking only the value b and c with $\mathbb{P}(\tau = b) = \mathbb{P}(\tau = c) = \frac{1}{2}$. The filtration is defined as

$$\mathcal{F}_a = \{\emptyset, \Omega\} \text{ and } \mathcal{F}_b = \mathcal{F}_c = \{\emptyset, \Omega, \{\tau = b\}, \{\tau = c\}\}.$$

Our considered process is given as

$$X_t = \begin{cases} 0, & t = a, \\ 1, & t \neq \tau, \\ -1, & t = \tau. \end{cases}$$

Then $\mathbb{E}[X_b | \mathcal{F}_a] = \mathbb{E}[X_c | \mathcal{F}_a] = X_a$ and X is a uniformly bounded martingale on I. However, $E[X_\tau | \mathcal{F}_\sigma] = -1 < X_\sigma$ for the choice $\sigma = a$. In this example τ is not reachable from σ and Theorem 2.0.8 fails.

Theorem 2.0.10. *See [74]: If X is a uniformly bounded martingale and if the countable partially orderd index set is directed, then*

$$X_\sigma = \mathbb{E}[X_\tau | \mathcal{F}_\sigma], \quad a.s.,$$

is true for any stopping time σ, τ such that $\sigma \leq \tau$.

It is also possible to use the continuous case. If the process $(X_{t_n})_{n \in \mathbb{N}_0}$ is a discrete-time submartingale with respect to the filtration $(\mathcal{F}_{t_n})_{n \in \mathbb{N}_0}$, where $0 = t_0 < t_1 < t_2 \cdots < \infty$, then $X_t := X_{\gamma(t)}$ and $\mathcal{F}_t := \mathcal{F}_{\gamma(t)}$ with $\gamma(t) := \max\{t_n \mid t_n \leq t, n \in \mathbb{N}_0\}$ for $t \in \mathbb{R}_+$ gives a well-defined right-continous submartingale and the setting of Theorem 2.0.11.

Theorem 2.0.11 (Doob's optional sampling theorem)**.**
Cf. [71, Theorem 4.83]: Let $X = (X_t)_{t \geq 0}$ be a right-continuous $\mathbb{F}$-submartingale, $\sigma : \Omega \to [0, \infty]$ an $\mathbb{F}$-stopping time, and $\tau : \Omega \to [0, \infty]$ an $\mathbb{F}$-stopping time. Then the following holds:

(a) For every $u \geq 0$, $X_{\tau \wedge u}$ and $X_{\sigma \wedge \tau \wedge u}$ are integrable random variables and

$$X_{\sigma \wedge \tau \wedge u} \leq \mathbb{E}[X_{\tau \wedge u} | \mathcal{F}_\sigma], \quad a.s. \tag{2.0.12}$$

(b) If $\mathbb{P}(\tau < \infty) = 1$ and if $(X^+_{\tau \wedge u})_{u \geq 0}$ is uniformly integrable, then X_τ and $X_{\sigma \wedge \tau}$ are a.s. well defined and integrable, and

$$X_{\sigma \wedge \tau} \leq \mathbb{E}[X_\tau | \mathcal{F}_\sigma], \quad a.s. \tag{2.0.13}$$

We need the following two other versions to prove parts of Theorem 2.0.6.

Theorem 2.0.14 (Optional sampling, Doob)**.**
See [39, Theorem 6.12]: Let Z be a martingale on some countable totally ordered index set I with filtration $\mathbb{F}$, and consider two optional times σ and τ, where τ is bounded. Then Z_τ is integrable, and $Z_{\sigma\wedge\tau} = \mathbb{E}[Z_\tau|\mathcal{F}_\sigma]$, a.s.

Theorem 2.0.15 (Finite optional sampling for submartingales)**.**
Cf. [71, Lemma 4.57]: Let $(X_t)_{t\in I}$ be a submartingale and let σ, τ be stopping times, where τ attains only finitely many values of I and σ takes values of I such that $\{t \in \sigma(\Omega)|t \le \max\tau(\Omega)\}$ is finite. Then X_τ and $X_{\sigma\wedge\tau}$ are integrable and

$$X_{\sigma\wedge\tau} \le \mathbb{E}[X_\tau|\mathcal{F}_\sigma], \quad a.s.$$

With I is discrete and contains the infimum of every subset which is bounded below, like I finite, $I = \mathbb{N}_0$, $I = \mathbb{Z}$ or $I = \{0\} \cup \left\{\frac{1}{n} \mid n \in \mathbb{N}\right\}$.

Proof of Theorem 2.0.6.

(a) For the special case that $I = \{0, 1, \dots, T\}$ with $T \in \mathbb{N}$ you can find the proof in [75]. Using Theorem 2.0.11 with $\sigma \equiv 0$ and $n \ge 0$ we obtain that $Z_{\tau\wedge n}$ is integrable and

$$\mathbb{E}[Z_{\tau\wedge n}] \le \mathbb{E}[Z_0]. \tag{2.0.16}$$

1. For some $N \in \mathbb{N}$ we have $\mathbb{P}(\tau \le N) = 1$ for all $t \in I$. Furthermore, for $n = N$ we get, using the first considerations, that $Z_\tau = Z_{\tau\wedge N}$ is integrable and $\mathbb{E}[Z_\tau] \le \mathbb{E}[Z_0]$, a.s.

2. Right now τ is finite a.s., i.e., $\mathbb{P}(\tau < \infty) = 1$ and Z is bounded a.s., i.e., for some $K > 0$ we have $|Z_t(\omega)| \le K = 1$ for all $t \in I$ and $\omega \in \Omega$. We can let $n \to \infty$ in (2.0.16) using bounded convergence theorem such that we get the statement.

3. $\mathbb{E}[\tau] = \sum_{t\in I} |t|\nu_t < \infty$ and for some $K > 0$ we have $|Z_t(\omega) - Z_{t-1}(\omega)| \le K = 1$ for all $t \in I \setminus \{0\}$ and for all $\omega \in \Omega$. Assume that we define the index set as $I := \{t_n | n \in \mathbb{N}_0, 0 = t_0\} \subseteq \mathbb{N}_0$. Either there exists an $m \in \mathbb{N}$ such that $\tau \wedge n = t_m$ or we use [33, Remark 3.91], if $\tau \wedge n = n \notin I$. For simplification we assume that there exists an $m \in \mathbb{N}$ such that $t_m = \tau \wedge n$. With

$$|Z_{\tau\wedge n} - Z_0| = \left|\sum_{k=1}^{m} Z_{t_k} - Z_{t_{k-1}}\right| \le K\left(\sum_{k=1}^{m} \underbrace{|t_k - t_{k-1}|}_{>0}\right) = K(\tau \wedge n) \le K\tau$$

   and $\mathbb{E}[\tau] < \infty$, so that dominated convergence justifies letting $n \to \infty$ in (2.0.16) to obtain the answer we want.

(b) Z is a martingale implies that Z is a supermartingale and $-Z$ is an supermartingale. The statement follows from the application of (a) on Z and $-Z$.

(c) It follows from Theorem 2.0.14.

(d) 1. If ν has a finite support, there exists an element $T \in I$ such that $\nu_{\le T} = 1$, i.e., $\mathbb{P}(\tau \le N) = \nu_{\le N} = 1$. If ν has a finite support, it also implies that τ attains only finitely many values of I. If a $t \in I$ with $\nu_{\leq t} = 1$ exists, then the corresponding stopping time is bounded by t a.s. We can use [71, Lemma 3.72] below or [71, Theorem 3.86] and get by the convenient choice with $\sigma = t$ that

$$Z_{\tau \wedge t} \geq \mathbb{E}[Z_\tau | \mathcal{F}_t], \qquad \text{a.s. for all } \max_{\omega \in \Omega} \tau(\omega) \geq t$$

such that

$$\mathbb{E}[Z_t] \geq \mathbb{E}[Z_\tau], \qquad \text{a.s. for all } \min_{\omega \in \Omega} \tau(\omega) \geq t$$

and that Z_τ is integrable.
If we assume that Z is a martingale, the corresponding statement would be follow immediately from Theorem 2.0.14.

2. With $Z^* = \sup_{t \in I} |Z_t| \in L^1$ we have that Z is uniformly bounded.
Using Theorem 2.0.10 we obtain the desired answer.

3. $\sum_{t \in I} |t| \nu_t < \infty$ and for some $K > 0$ we have $|Z_t - Z_s| \leq K|t - s|$ a.s. for all $s, t \in I$.
For all $\omega \in \Omega$ and for all $s, t \in I \setminus \{0\}$: $\exists k, l, m \in \mathbb{N}_0$ such that $t = t_k$ and $t_l = s$ with $s \leq \tau \wedge t = t_m$:

$$|Z_{\tau \wedge t} - Z_s| = \left| \sum_{k=l}^{m} Z_{t_k} - Z_{t_{k-1}} \right| \leq K \left(\sum_{k=l}^{m} |t_k - t_{k-1}| \right) = K|\tau \wedge t - s| \leq \underbrace{K|\tau \wedge t|}_{K|\tau|} + \underbrace{K|s|}_{<\infty}.$$

We get

$$\mathbb{E}[|Z_{\tau \wedge t} - Z_s|] \leq K \cdot \mathbb{E}[|\tau|] + K|s| = K \underbrace{\sum_{t \in I} |t| \nu_t}_{<\infty} + K|s| < \infty,$$

so that dominated convergence justifies letting $k \to \infty$ in $\mathbb{E}[Z_{\tau \wedge t_k} - Z_{t_l}] \leq 0$, $k, l \in \mathbb{N}_0$ $k \leq l$, to obtain the answer we want. ❑

3

Adapted Random Probability Measure

In the previous section we have discussed the problem $\textsc{OptStop}^{\tau}$ and its connection to the standard problem. Now we will replace the stopping times by adapted random probability measures. First we have to define adapted random probability measures and describe the corresponding problem $\textsc{OptStop}^{\gamma}$. Furthermore, in Section 3.2 we will show the connection between the different optimization problems $\textsc{OptStop}^{\tau}$ and $\textsc{OptStop}^{\gamma}$ and how the corresponding values of the problems change. An important aspect in considering this problem is the question of an optimal strategy. In Section 3.3, the existence of such an optimal strategy in a discrete time setting is proven for the generalized approach, which allow us to consider a much larger set of possible processes. This optimal strategy is not unique in general as illustrated by some examples. In addition to this, some general results and bounds are derived in Section 3.4. Some of them apply to the problem in general, while others depend on the structure of the underlying process. For certain classes of stochastic processes, it is possible to find an optimal strategy and the resulting value of the optimization problem, see Section 3.5. These include, for example, processes that are the product of a martingale and a deterministic function or in the binomial model. In Section 3.6, an application of the problem in the area of unit-linked life insurance is discussed in which the modeling of the contract takes place without assuming independence between biometric and financial risks.

Some additional results and proofs that would unnecessarily disturb the flow of reading are outsourced. The reader can find them in Subsection 3.3.3 or in the Appendix A.

3.1. The Problem and Main Results

First we will give the definition of adapted random probability measures. Note that we modify [33, Definition 2.12] for our considerations. This modified definition gives us the possibility to avoid the necessity of stopping on the one hand and consider adapted random subprobability measures on the other hand, see Remark 3.1.8. Adapted random probability measures are adapted stochastic processes $\gamma = (\gamma_t)_{t\in I}$ defined as follows:

Definition 3.1.1. For a real-valued process $\gamma = (\gamma_t)_{t\in I}$, we write $\gamma \in \mathcal{M}_I$, if

(a) $\gamma_t \geq 0$ for all $t \in I$,

(b) $\sum_{t\in I} \gamma_t \leq 1$,

(c) $\sum_{t\in I} \gamma_t \geq 1$, a.s.,

(d) γ_t is $\mathcal{F}_t$-measurable for all $t \in I$, i.e., γ is adapted.

Given a probability measure $\nu = (\nu_t)_{t\in I}$ on I, we say that the above stochastic process γ is in $\mathcal{M}_I^\nu$, if in addition,

(e) $\mathbb{E}[\gamma_t] = \nu_t$ for all $t \in I$.

Two adapted random probability measures may be identified if they induce the same probability measure on $\mathcal{F} \otimes \mathcal{P}(I)$.

Example 3.1.2. Cf. [33, Remark 2.18]: Given a $\mathbb{F}$-stopping time τ with $\mathbb{P}(\tau \in I) = 1$, it can be naturally identified with the $\mathbb{F}$-adapted stochastic process $\gamma = (\gamma_t)_{t\in I}$ defined by

$$\gamma_t(\omega) = \mathbb{1}_{\{\tau(\omega)\}}(t), \quad \omega \in \Omega,\ t \in I. \tag{3.1.3}$$

On $\{\tau \in I\}$ this stochastic process γ defines a probability measure on I. If $\mathcal{L}(\tau) = \nu$, then $\gamma \in \mathcal{M}_I^\nu$. The reverse construction, i.e., finding a stopping time τ producing a given $\gamma \in \mathcal{M}_I^\nu$ via (3.1.3), is contained in Theorem 3.2.8 below and relies on an enlargement of the filtration $\mathbb{F}$.

Example 3.1.4. Cf. [33, Remark 2.20]: The set $\mathcal{M}_I^\nu$ is never empty, because it contains the adapted random probability measure γ defined by $\gamma_t = \nu_t \mathbb{1}_\Omega$ for all $t \in I$.

Furthermore, we know that the set $\mathcal{M}_I^\nu$ of adapted random probability measures is convex.

Lemma 3.1.5. *Cf. [33, Lemma 3.1]: Given γ and $\tilde{\gamma}$ in $\mathcal{M}_I^\nu$ and a $[0,1]$-valued random variable Λ which is $\mathcal{F}_t$-measurable for all $t \in I$. If Λ is uncorrelated to γ_t and $\tilde{\gamma}_t$ for all $t \in I$, then also $\Lambda\gamma + (1-\Lambda)\tilde{\gamma} \in \mathcal{M}_I^\nu$. In particular the set $\mathcal{M}_I^\nu$ is convex.*

Proof. Using Definition 3.1.1 it is easy to check that for γ and $\tilde{\gamma}$ in $\mathcal{M}_I^\nu$ we have $\Lambda\gamma + (1-\Lambda)\tilde{\gamma} \in \mathcal{M}_I^\nu$. ❑

The attentive reader might have stumbled upon ‘almost surely’ in Definition 3.1.1(c). This is related to the definition of a stopping time as a map to $\bar{I}$. Example 3.2.4 below shows the necessity. Additionally, note that the Jensen’s inequality for the convex function $x \mapsto |x|^p$ with $p \geq 1$ works for substochastic measures. The definition of adapted random probability measures, Definition 3.1.1, can be modified depending on the interest.

Remark 3.1.6. Cf. [33, Remark 2.16]: Another point of interest could be to assume $\sum_{t\in I}\gamma_t \leq x$, a.s., or $\gamma_t \in [0,y]$, a.s., for $t \in I$, $x,y \in [0,\infty)$. Some of the results, which are shown, can be adjusted to such a problem. Nevertheless, we will concentrate on $x = y = 1$ in the article.

In this chapter we will mainly consider the distribution-constrained optimization problem of the following form:

Problem (OptStop$^\gamma$). Consider a real-valued and $\mathbb{F}$-adapted stochastic process $Z = (Z_t)_{t\in I}$. Find sufficient conditions such that:

(a) For every $\gamma \in \mathcal{M}_I^\nu$, the series $\sum_{t\in I}\gamma_t Z_t$ defining Z_γ, is $\mathbb{P}$-a.s. absolutely convergent in $\overline{\mathbb{R}}$ satisfying $\mathbb{E}[Z_\gamma^+] < \infty$.

(b) There exists a maximizer $\gamma^* \in \mathcal{M}_I^\nu$ solving

$$\mathbb{E}[Z_{\gamma^*}] = \sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}[Z_\gamma]. \tag{3.1.7}$$

Remark 3.1.8. Our framework also includes two special cases which can be of interest. By adding an additional time point $t^\star$ to I, which is an upper bound for I, and setting $Z_{t^\star} = 0$, we can avoid the necessity of stopping on the one hand and consider adapted random subprobability measures on the other hand. By defining $\gamma_{t^\star} = 1 - \sum_{t\in I}\gamma_t$ we construct an adapted random probability measure.

Remark 3.1.9. The definitions of $\mathcal{T}_I$, $\mathcal{T}_I^\nu$, $\mathcal{M}_I$ and $\mathcal{M}_I^\nu$ depend on the underlying filtration $\mathbb{F}$.

Due to $\mathcal{M}_I^\nu \neq \emptyset$, see Example 3.1.4, the filtration $\mathbb{F}$ is not a limiting factor to have a solution for the problem OptStop$^\gamma$ (unlike OptStop$^\tau$). However, as Example 3.3.90 below shows, a process Z, even when it is bounded in L^1 and guarantees (a) of OptStop$^\gamma$, might be growing too fast for an optimizer γ^* to exist. Therefore, we will concentrate on moment conditions for the adapted stochastic process Z.
Thus the value we want to compute under the assumptions of Theorem 3.3.5 or Theorem 3.3.34 for Z, is defined by

$$V_{\mathcal{M}}^\nu(Z) = \sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}[Z_\gamma]. \tag{3.1.10}$$

Note that $\mathcal{M}_I^\nu \neq \emptyset$ and $V_{\mathcal{M}}^\nu(Z) < \infty$ is guaranteed by Example 3.1.4 and Theorem 3.3.34 (or Theorem 3.3.5) , as the assumptions stated there make sure that the problem is well-posed. The value $V_{\mathcal{M}}^\nu(Z)$ is defined and finite for all processes $Z \in \prod_{t\in I}\mathcal{L}^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$ with $\|Z\|_{\nu,p,q} < \infty$, $p \in [1,\infty)$ and $q \in [1,\infty]$, see Lemma 3.3.65 (or $Z \in \prod_{t\in I} L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{R})$ with $\|Z\|_{\nu,p} < \infty$, $p \in [1,\infty)$, see Lemma 3.3.25). Furthermore, the distribution-constrained optimization problem OptStop$^\gamma$ is indeed an enlargement of the problem OptStop$^\tau$ by using (3.1.3).
All relevant considerations, preliminaries and the proof of the existence of an optimal strategy can be found in Section 3.3.

Remark 3.1.11. Cf. [33, Remark 2.10]: All the results stated can also be adjusted to treat the infimum instead of the supremum, because

$$\inf_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}[Z_\gamma] = -\sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}[-Z_\gamma].$$

Example 3.1.12. If we assume that $(Z_t)_{t\in I}$ satisfies $\mathbb{E}[\sup_{t\in I}|Z_t|] < \infty$ as required in [33], our considered problem is well-posed and the values, we are interested in, exist and are finite. For example, $\mathbb{E}[Z_\gamma]$ exists and is finite for every $\gamma \in \mathcal{M}_I$, because of

$$\mathbb{E}[|Z_\gamma|] = \mathbb{E}\Big[\Big|\sum_{t\in I} Z_t\gamma_t\Big|\Big] \leq \sum_{t\in I}\mathbb{E}[|Z_t|\gamma_t] \leq \mathbb{E}\Big[\sup_{t\in I}|Z_t|\underbrace{\sum_{t\in I}\gamma_t}_{\leq 1 \text{ a.s.}}\Big] \leq \mathbb{E}[\sup_{t\in I}|Z_t|] < \infty.$$

This implies that Z_γ is well-defined and integrable. Because of $\mathcal{M}_I^\nu \subseteq \mathcal{M}_I$, we get an analogously definition on the smaller set $\mathcal{M}_I^\nu$.

Furthermore, it should be also clear that an optimal strategy for our problem $\textsc{OptStop}^\gamma$ satisfies the following definition:

Definition 3.1.13 (Optimal strategy)**.**
Let I be a countable, totally-ordered index set and $(Z_t)_{t\in I}$ a process such that $\mathbb{E}[Z_\gamma^+]$ is finite for all $\gamma \in \mathcal{M}_I^\nu$. If there exists a $\gamma^* \in \mathcal{M}_I^\nu$ such that

$$\mathbb{E}[Z_{\gamma^*}] \geq \mathbb{E}[Z_\gamma], \quad \forall \gamma \in \mathcal{M}_I^\nu,$$

then γ^* is optimal for $(Z_t)_{t\in I}$.

3.2. Connections between the Different Optimal Stopping Problems and Illustrating Examples

In this section we will show the connection between the different optimization problems $\textsc{OptStop}^\tau$ and $\textsc{OptStop}^\gamma$ and how the corresponding values of the optimization problems change. Note that thereby the problems depend on the given distribution, the filtration and the underlying process. In Example 3.1.2 we have already seen that $\mathcal{T}_I^\nu$ can be embedded in $\mathcal{M}_I^\nu$ via (3.1.3). By enlarging the filtration in an eligible way we can also embed the original set $\mathcal{M}_I^\nu$ into a set $\tilde{\mathcal{T}}_I^\nu$ corresponding to an enlarged filtration, if the underlying process Z retains its original measurability. This reverse construction is contained in Theorem 3.2.8 below.
For the computation of the value $V_\mathcal{T}^\nu(Z)$ we assume that the filtration in our model is chosen appropriately. Otherwise it could happen that $\mathcal{T}_I^\nu = \emptyset$, as shown in Example 2.0.3 above, since a set might not exist in $\mathcal{F}_t$ with probability ν_t for some $t \in I$. This is not necessary for the computation of $V_\mathcal{M}^\nu(Z)$ since at least one adapted random probability measure exists in $\mathcal{M}_I^\nu$, see Example 3.1.4. We have seen that the $\textsc{OptStop}^\gamma$ is an enlargement of the problem $\textsc{OptStop}^\tau$, cf. (3.1.3). Therefore, it holds obviously that

$$\sup_{\tau\in\mathcal{T}_I^\nu}\mathbb{E}[Z_\tau] \leq \sup_{\gamma\in\mathcal{M}_I^\nu}\mathbb{E}[Z_\gamma] \leq \sup_{\gamma\in\mathcal{M}_I}\mathbb{E}[Z_\gamma] =: V_\mathcal{M}(Z), \tag{3.2.1}$$

because $\mathcal{T}_I^\nu$ is embedded in $\mathcal{M}_I^\nu$ via (3.1.3) and $\mathcal{M}_I^\nu \subseteq \mathcal{M}_I$.

Remark 3.2.2. Cf. [33, Remark 2.18]: If a stopping time τ is used for modeling a claim, then an adapted random probability measure can be used to model a portfolio of such claims. If a portfolio consists of countable many claims, indexed by the set C, modeled by a series of stopping times $(\tau_j)_{j\in C}$, then the whole portfolio can be modeled using the adapted random probability measure γ given by $\gamma_t(\omega) = \sum_{j\in C} w_j \mathbb{1}_{\tau_j(\omega)}(\{t\})$ for $t \in I$, $\omega \in \Omega$ with non-negative weights $(w_j)_{j\in C}$ defining a probability measure on C. Note that τ_j in $\mathcal{T}_I^\nu$ for every $j \in C$ implies $\gamma \in \mathcal{M}_I^\nu$.

In the study of these two problems OptStop$^\gamma$ and OptStop$^\tau$ we do not suppose that the underlying process and the stopping time or the adapted random probability measure are independent. Hence, we have an adapted dependence between Z and τ or Z and γ.
We already noted that $V_\mathcal{T}^\nu(Z) \leq V_\mathcal{M}^\nu(Z)$ and that the inequality can be strict if $\mathcal{T}_I^\nu = \varnothing$. As we will see in the following example, it is possible that $V_\mathcal{T}^\nu(Z) < V_\mathcal{M}^\nu(Z)$ even in the case $\mathcal{T}_I^\nu \neq \varnothing$.

Example 3.2.3. Cf. [33, Remark 2.22]: We take a look at a one-period model with $I = \{0,1\}$ and $\Omega = \{\omega_0, \omega_1\}$. We assume that a probability distribution ν on I with $\nu_0 \neq \nu_1$ and $\nu_0, \nu_1 \in (0,1)$ is given. Let $\mathcal{F}_0 = \mathcal{F}_1 = \mathcal{P}(\Omega) = \{\varnothing, \{\omega_0\}, \{\omega_1\}, \Omega\}$. Furthermore, we assume $\mathbb{P}(\{\omega_i\}) = \nu_i$ for $i \in \{0,1\}$. Now we know that the only stopping time $\tau \in \mathcal{T}_I^\nu$ with the given distribution ν is given by

$$\{\tau = 0\} = \{\omega_0\}, \quad \{\tau = 1\} = \{\omega_1\}.$$

An adapted random probability measure $\gamma \in \mathcal{M}_I^\nu$ different from the one induced by the stopping time τ is given by $\gamma_i = \nu_i$ for $i \in \{0,1\}$. If the process Z is given by $Z_0(\omega_0) = 0$, $Z_1(\omega_0) = 1$, $Z_0(\omega_1) = 1$ and $Z_1(\omega_1) = 0$, then we have $V_\mathcal{T}^\nu(Z) = 0$, whereas

$$V_\mathcal{M}^\nu(Z) \geq \mathbb{E}[Z_0\gamma_0] + \mathbb{E}[Z_1\gamma_1] = \nu_0\nu_1 + \nu_1\nu_0 > 0\,.$$

Therefore, we have $V_\mathcal{T}^\nu(Z) < V_\mathcal{M}^\nu(Z)$.

The next example explains why it is necessary to consider $\bar{I}$ and how it connects to the condition (c) in Definition 3.1.1.

Example 3.2.4. Let $I = \mathbb{N}$ and the probability space be given by $\Omega = \{0,1\}^\mathbb{N}$, the product σ-algebra and the product measure of the Laplace distribution on $\{0,1\}$. Then $\bar{I} = \mathbb{N} \cup \{\infty\}$. Define the process Z by $Z_t(\omega) = \omega_t$ for all $t \in \mathbb{N}$. Let $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{N}}$ be the natural filtration of Z, i.e., $\mathcal{F}_t = \sigma(Z_1, \dots, Z_t)$ for all $t \in I$. We want maximize the value $\mathbb{E}[Z_\tau]$. It is obvious that the greedy strategy will solve the problem such that an optimal stopping time is given by $\tau = \inf\{t \in I \mid Z_t = 1\}$. For this it holds that $\tau((0,0,\dots)) = \infty$. If we instead of using the convention $\inf\varnothing = \infty$ assume that $\tau((0,0,\dots)) = t$ for some $t \in \mathbb{N}$, we get a contradiction to $\{\tau \leq t\} \in \mathcal{F}_t$. Therefore, the definition of stopping times as maps to $\bar{I}$ is necessary. Furthermore, this implies by using (3.1.3) that the corresponding adapted random probability measure γ satisfies $\sum_{t\in I} \gamma_t \geq 1$ only a.s. Note that $\mathcal{T}_\mathbb{N}^\nu \neq \varnothing$ if and only if $2^t \nu_t \in \mathbb{N}_0$ for all $t \in \mathbb{N}$.

Example 3.2.5. Given $p = 1$, $q \in (1, \infty)$ and a probability measure ν with countably infinite support $\operatorname{supp}(\nu)$, we claim that there is always a process $Z \in \mathcal{X}_{\nu,p,q}$ and a $\gamma \in \mathcal{M}_I^\nu$ such that $\sum_{t \in I} |Z_t| \gamma_t = \infty$ on Ω. Indeed, for an enumeration $(t_n)_{n \in \mathbb{N}}$ of $\operatorname{supp}(\nu)$, define the deterministic process

$$Z_t = \begin{cases} 1/(n\nu_{t_n}) & \text{if } t = t_n \text{ for an } n \in \mathbb{N}, \\ 0 & \text{if } t \in I \setminus \operatorname{supp}(\nu). \end{cases}$$

Using Definition 3.1.1(e), it follows for every $\gamma \in \mathcal{M}_I^\nu$ and $n \in \mathbb{N}$ that

$$(\mathbb{E}[|Z_{t_n}| \gamma_{t_n}])^q = (Z_{t_n} \mathbb{E}[\gamma_{t_n}])^q = (Z_{t_n} \nu_{t_n})^q = 1/n^q,$$

hence $\|Z\|_{\nu,p,q}^q = \sum_{n \in \mathbb{N}} 1/n^q < \infty$. If, as in Example 3.1.4, we take $\gamma_t := \nu_t \mathbb{1}_\Omega$ for all $t \in I$, then $\sum_{t \in I} |Z_t| \gamma_t$ simplifies to the harmonic series.

For the next example we need an elementary proposition.

Proposition 3.2.6. *Let $(a_n)_{n \in \mathbb{N}}$ be a sequence in $[0, \infty)$ converging to zero and such that $\sum_{n \in \mathbb{N}} a_n = \infty$. Then there exists a decreasing sequence $(b_n)_{n \in \mathbb{N}}$ in $(0, 1]$ converging to zero such that $\sum_{n \in \mathbb{N}} a_n b_n = \infty$ and $\sum_{n \in \mathbb{N}} a_n b_n^q < \infty$ for every $q > 1$.*

Proof. Define $s_n = a_1 + \cdots + a_n$ for $n \in \mathbb{N}_0$. Furthermore, there exists an upper bound $c \geq 1$ for $(a_n)_{n \in \mathbb{N}}$, because there is a $N \in \mathbb{N}$ such that $a_n \leq 1$ for all $n \geq N$. Starting with $n_0 := 0$, we can find iteratively for every $l \in \mathbb{N}$ an $n_l > n_{l-1}$ such that $l \leq s_{n_l} - s_{n_{l-1}} \leq 2cl$. Define $b_n = 1/l^2$ when $n \in \mathbb{N}$ satisfies $n_{l-1} < n \leq n_l$. Then

$$\sum_{n \in \mathbb{N}} a_n b_n = \sum_{l \in \mathbb{N}} \frac{s_{n_l} - s_{n_{l-1}}}{l^2} \geq \sum_{l \in \mathbb{N}} \frac{1}{l} = \infty$$

and, for every $q > 1$,

$$\sum_{n \in \mathbb{N}} a_n b_n^q = \sum_{l \in \mathbb{N}} \frac{s_{n_l} - s_{n_{l-1}}}{l^{2q}} \leq 2c \sum_{l \in \mathbb{N}} \frac{1}{l^{2q-1}} < \infty.$$

❑

Example 3.2.7. Given $p, q \in (1, \infty)$ and a probability measure ν not satisfying the decay condition (3.3.35), we claim (as in Example 3.2.5) that there is always a process $Z \in \mathcal{X}_{\nu,p,q}$ and a $\gamma \in \mathcal{M}_I^\nu$ such that $\sum_{t \in I} |Z_t| \gamma_t = \infty$ on Ω. Indeed, if (3.3.35) is not satisfied, then the support $\operatorname{supp}(\nu)$ of ν is countably infinite. Let $p' = (p' - 1)p$, $(q' - 1)q = q'$ and $(t_n)_{n \in \mathbb{N}}$ be an enumeration of $\operatorname{supp}(\nu)$. Apply Proposition 3.2.6 with $a_n := \nu_{t_n}^{q'/p'}$ for $n \in \mathbb{N}$. Use the corresponding sequence $(b_n)_{n \in \mathbb{N}}$ to define the deterministic process

$$Z_t = \begin{cases} b_n \nu_{t_n}^{q'/p' - 1} & \text{if } t = t_n \text{ for an } n \in \mathbb{N}, \\ 0 & \text{if } t \in I \setminus \operatorname{supp}(\nu). \end{cases}$$

Using Definition 3.1.1(e), it follows for every $\gamma \in \mathcal{M}_I^\nu$ and $n \in \mathbb{N}$ that

$$\begin{aligned} \left(\mathbb{E}[|Z_{t_n}|^p \gamma_{t_n}]\right)^{q/p} &= \left(Z_{t_n} (\mathbb{E}[\gamma_{t_n}])^{1/p}\right)^q = \left(Z_{t_n} \nu_{t_n}^{1/p}\right)^q \\ &= b_n^q \left(\nu_{t_n}^{q'/p' - 1 + 1/p}\right)^q = b_n^q \left(\nu_{t_n}^{(q'-1)/p'}\right)^q = b_n^q \nu_{t_n}^{q'/p'}, \end{aligned}$$

hence $\|Z\|_{\nu,p,q}^q = \sum_{n\in\mathbb{N}} a_n b_n^q < \infty$ by Proposition 3.2.6. If, as in Example 3.1.4, we take $\gamma_t := \nu_t \mathbb{1}_\Omega$ for all $t \in I$, then $\sum_{t\in I} |Z_t|\gamma_t$ simplifies to $\sum_{n\in\mathbb{N}} a_n b_n$, which diverges by Proposition 3.2.6.

The distribution-constrained optimization problem OptStop^γ can also be connected to a distribution-constrained optimal stopping problem OptStop^τ given an enlarged filtration $(\tilde{\mathcal{F}}_t)_{t\in I}$. By enlarging the filtration in an eligible way we can also embed the original set $\mathcal{M}_I^\nu$ into a set $\tilde{\mathcal{T}}_I^\nu$ corresponding to an enlarged filtration. Remember that the sets $\mathcal{M}_I^\nu$ and $\mathcal{T}_I^\nu$ depend on the underlying filtration and their defintions are more general to the ones in [33]. Now, the following theorem describes the construction of an appropriate stopping time, cf. [33, Theorem 2.41].

Theorem 3.2.8. *Let be $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in I}, \mathbb{P})$ the filtered probability space. We may assume w.l.o.g. that there exists a random variable U, uniformly distributed on $(0,1]$ and independent of $\mathcal{F}_T$, see Remark 3.2.10.*
Consider an adapted random probability measure $\gamma \in \mathcal{M}_I$ as in Definition 3.1.1 w.r.t. the filtration $\mathbb{F} := (\mathcal{F}_t)_{t\in I}$. Define the random time $\tau : \Omega \to \bar{I}$ by

$$\{\tau = t\} = \begin{cases} \{\gamma_{<t} < U \le \gamma_{\le t}\} & \text{for } t \in I \setminus \{T\}, \\ \{\gamma_{<T} < U\} & \text{for } t = T, \end{cases} \tag{3.2.9}$$

where $\gamma_{<t}$, $\gamma_{\le t}$ and $\gamma_{<T}$ are to be understood as in Notation 2.0.1(d), if $T \in I$. For $T \notin I$, it holds that $\gamma_{<T} = \sum_{t\in I} \gamma_t$. Then the following holds:

(a) Define the enlarged filtration $(\tilde{\mathcal{F}}_t)_{t\in I}$ by

$$\tilde{\mathcal{F}}_t = \sigma\Big(\mathcal{F}_t \cup \bigcup_{s\in I_{\le t}} \{\tau \le s\}\Big) \subseteq \sigma(\mathcal{F}_t \cup \sigma(U)) \quad \text{for } t \in I.$$

Then τ is a stopping time w.r.t. $(\tilde{\mathcal{F}}_t)_{t\in I}$ satisfying $\mathbb{P}(\tau = t|\mathcal{F}_T) \overset{a.s.}{=} \gamma_t$ for all $t \in I$.

(b) If $\gamma \in \mathcal{M}_I^\nu$, then $\mathcal{L}(\tau) = \nu$.

(c) Let $Z = (Z_t)_{t\in I}$ be an $\mathcal{F}_T$-measurable process such that $Z_\tau \in L^1$. Then

$$\mathbb{E}[Z_\tau|\mathcal{F}_T] \overset{a.s.}{=} Z_\gamma \qquad \text{and} \qquad \mathbb{E}[Z_\tau] = \mathbb{E}[Z_\gamma].$$

Proof. (a) We see that τ defined as in (3.2.9) is really a stopping time w.r.t. $(\tilde{\mathcal{F}}_t)_{t\in I}$. If $T \notin I$, then $\mathbb{P}(\tau \notin I) = \mathbb{P}(\gamma_{<T} < 1) = 0$ by Definition 3.1.1(c). As U is independent of $\mathcal{F}_T$ we also have $\mathbb{P}(\tau = t|\mathcal{F}_T) \overset{\text{a.s.}}{=} \gamma_t$ for $t \in I$.

(b) It follows immediately by part (a) and Definition 3.1.1(e).

(c) Since I is countable, there exists an increasing sequence $(I_k)_{k\in\mathbb{N}}$ of finite index sets with $\bigcup_{k\in\mathbb{N}} I_k = I$. Note that for every finite set I_k, $k \in \mathbb{N}$, it holds that $I_k = \bar{I}_k$. Since $Z_\tau = \lim_{k\to\infty} Z_\tau \mathbb{1}_{I_k}(\tau)$ pointwise on Ω and $|Z_\tau \mathbb{1}_{I_k}(\tau)| \le |Z_\tau|$ for every $k \in \mathbb{N}$, we can apply the dominated convergence theorem for conditional expectations, and have

$$\begin{aligned} \mathbb{E}[Z_\tau|\mathcal{F}_T] &= \mathbb{E}\Big[\sum_{t\in I} Z_t \mathbb{1}_{\{\tau=t\}}\Big|\mathcal{F}_T\Big] = \sum_{t\in I} \mathbb{E}[Z_t \mathbb{1}_{\{\tau=t\}}|\mathcal{F}_T] \\ &= \sum_{t\in I} Z_t\, \mathbb{P}(\tau = t|\mathcal{F}_T) = \sum_{t\in I} Z_t \gamma_t = Z_\gamma, \quad \text{a.s.} \end{aligned}$$

❑

Remark 3.2.10. If it is necessary to enlarge the probability space, this can be done by setting $\hat{\Omega} := (0,1] \times \Omega$, $\hat{\mathcal{F}} := \mathcal{B}_{(0,1]} \otimes \mathcal{F}$ and $\hat{\mathbb{P}} := \lambda \otimes \mathbb{P}$, where λ denotes the Lebesgue–Borel measure. Let $U\colon \hat{\Omega} \to (0,1]$ be the projection onto the first component. For a filtration on the extended probability space it would be sufficient to consider a filtration given by $\hat{\mathcal{F}}_t := \{\emptyset, (0,1]\} \otimes \mathcal{F}_t$ for $t \in I$. Let $\pi\colon (0,1] \times \Omega \to \Omega$ be the projection on the second component. We then consider on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ the process $\hat{Z}_t := Z_t \circ \pi$ for all $t \in I$. Similarly we consider $\hat{\gamma}_t := \gamma_t \circ \pi$ for all $t \in I$. Note that $\hat{Z}$ and $\hat{\gamma}$ are adapted to $(\hat{\mathcal{F}}_t)_{t\in I}$.

Remark 3.2.11. Due to the construction of the stopping time τ in Theorem 3.2.8 by using the given adapted random probability measure γ we have $\{\tau = t\} \subseteq \{\gamma_t > 0\}$ for all $t \in I \setminus \{T\}$.

Finally, we conclude that there is a solution to these induced problems of the form OptStop^τ, thus there also exists an optimal strategy. By Theorem 3.2.8 we have shown that the problem OptStop^γ can be connected to a distribution-constrained optimal stopping problem OptStop^τ given an enlarged filtration $(\tilde{\mathcal{F}}_t)_{t\in I}$. Therefore, there exists an optimal strategy $\tau \in \mathcal{T}_I^\nu$ for all processes Z which satisfies Theorem 3.3.5 or Theorem 3.3.34 and for such a problem OptStop^τ which can be traced back to a problem of form OptStop^γ.

Corollary 3.2.12. *Consider a real-valued process Z which satisfies Theorem 3.3.5 or Theorem 3.3.34. Then there always exists an optimal strategy $\tau^* \in \mathcal{T}_I^\nu$ for all problems OptStop^τ, which emerges from a problem OptStop^γ by means of the Theorem 3.2.8, such that τ^* solves*

$$\sup_{\tau \in \mathcal{T}_I^\nu} \mathbb{E}[Z_\tau] = \mathbb{E}[Z_{\tau^*}].$$

Proof. It follows by the combination of Theorem 3.3.5 or Theorem 3.3.34 and Theorem 3.2.8. ❑

3.3. Existence of an Optimal Strategy

After the introduction of our considered problem in the last section, an important question is whether an optimal strategy exists that yields the supremum we want to calculate. Again we take a look at a discrete time interval I. As we will see in this section an optimal $\gamma \in \mathcal{M}_I^\nu$ in Definition 3.1.1 always exists for a process Z satisfying the conditions in Theorem 3.3.5 or in Theorem 3.3.34. These are a generalization of [33, Theorem 3.10]. In [33, Chapter 3] the existence of an optimal strategy $\gamma \in \mathcal{M}_I^\nu$ is shown for all processes Z satisfying $\mathbb{E}[\sup_{t\in I} |Z_t|] < \infty$. This assumption for the process Z does not depend on the given probability measure ν on I. Thus known information does not be used. We have considered a more refined version, which uses the structure and information given by ν. This gives us the opportunity to look at many more processes, as the following example shows.

Example 3.3.1 (Motivation of the generalization)**.** Suppose that $(X_n)_{n\in\mathbb{N}}$ is a sequence of i.i.d. random variables satisfying

$$\mathbb{P}(X_n = 1) = \mathbb{P}(X_n = -1) = \frac{1}{2} \quad n \in \mathbb{N}.$$

Define the simple symmetric random walk as $Z_n := 1 + \sum_{i=1}^{n} X_i$ for all $n \in \mathbb{N}_0$. Let $\mathcal{F}_n = \sigma(Z_0, \dots, Z_n)$. Define the first hitting time of the level zero by $\tau_0 = \inf\{n \in \mathbb{N} : Z_n = 0\}$. Then the process $Z = (Z_n)_{n\in\mathbb{N}}$ is adapted and τ_0 is a stopping time.
We will show that $\mathbb{P}(\tau_0 < \infty) = 1$. Given $k \geq 2$, define the first exit time of the open interval $(0,k)$ by $\tau_{0,k} = \tau_0 \wedge \tau_k$ which is also a stopping time. By Borel-Cantelli lemma we get that $\mathbb{P}(\tau_{0,k} < \infty) = 1$, i.e., the random walk Z leaves a.s. every bounded Borel set of $\mathbb{Z}$. Since $\{Z_{\tau_{0,k}\wedge u}\}_{u\geq 0}$ is bounded by k, it is uniformly integrable. Applying of Doob's optional sampling theorem, we see that $Z_0 \overset{\text{a.s.}}{=} \mathbb{E}[Z_{\tau_{0,k}}|\mathcal{F}_0]$, hence

$$1 = \mathbb{E}[Z_0] = \mathbb{E}[Z_{\tau_{0,k}}] = k \cdot \mathbb{P}(Z_{\tau_{0,k}} = k) + 0 \cdot \mathbb{P}(Z_{\tau_{0,b}} = 0),$$

which implies that $\mathbb{P}(Z_{\tau_{0,k}} = k) = 1/k$ for $k \geq 2$. Hence, we get

$$\mathbb{P}(Z_{\tau_{0,k}} = 0) = 1 - \frac{1}{k} \nearrow 1 \quad \text{as } k \to \infty. \tag{3.3.2}$$

If the random walk leaves $(0,k)$ at 0, then it reaches 0, meaning that $\{\tau_{0,k} < \infty, Z_{\tau_{0,k}} = 0\} \subseteq \{\tau_0 < \infty\}$. Hence (3.3.2) proves that $\mathbb{P}(\tau_0 < \infty) = 1$. From now on, we consider the stopped process Z^{τ_0}, which is again a martingale and we know that

$$\mathbb{E}\Big[\sup_{n\in\mathbb{N}_0} Z_n^{\tau_0}\Big] = \sum_{k\in\mathbb{N}} \mathbb{P}\Big(\sup_{n\in\mathbb{N}_0} Z_n^{\tau_0} \geq k\Big) = \infty.$$

Let ν be an arbitrary measure on I with finite first moment, i.e.,

$$\sum_{k\in\mathbb{N}} k\,\nu_k < \infty,$$

then

$$\mathbb{E}[|Z_\gamma^\tau|] = \mathbb{E}\Big[\sum_{k\in\mathbb{N}} \underbrace{|Z_k^\tau|}_{\leq k+1}\,\gamma_k\Big] \leq \sum_{k\in\mathbb{N}} (k+1)\underbrace{\mathbb{E}[\gamma_k]}_{=\nu_k} < \infty.$$

This inequality holds for all martingales satisfying $|Z_k| \leq k+1$ for all $k \in \mathbb{N}_0$, such that we can consider our problem OptStop^γ for these processes.

The main results are given in Theorem 3.3.5 and Theorem 3.3.34. In this context you get Theorem 3.3.5 by using Theorem 3.3.34 with $p = q$, such that the second one is a generalization. But both have their validity. In some cases we get our distribution ν from empirical data such that we have no further information about the higher moments of this distribution. For example in actuarial mathematics we determine the distribution ν of the stopping time τ or the adapted random probability measure γ by using a life table, see Subsection 3.6. For this application the Subsection 3.3.1 is enough and there are a lot of additional results in this section. Otherwise, it is also possible that our given distribution ν belongs to a special family of distributions such that we know much more about ν. In this case we would consider Subsection 3.3.2. For better readability, some proofs have been moved to the Subsection 3.3.3. For the argumentation will be used partly analogous aids from functional analysis like in [33].
Remember that in general we can not expect an optimal strategy to be unique. As example,

consider for $\textsc{OptStop}^{\tau}$ a uniformly integrable martingale Z with index set $I = \mathbb{N}$. By Doob's optional stopping theorem, every stopping time $\tau \in \mathcal{T}_{\mathbb{N}}$ gives the same value for $\mathbb{E}[Z_\tau]$. Therefore, all strategies τ with $\mathcal{L}(\tau) = \nu$ are optimal for $\textsc{OptStop}^{\tau}$ and we do not have uniqueness.
Again, let I be a countable totally-ordered index set and $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a collection of σ-algebras $\mathbb{F} = (\mathcal{F}_t)_{t\in I}$.

3.3.1. One Fixed Given Distribution

We have introduced our problem in Section 3.1 and detected that the filtration is not a limiting factor to have a solution for $\textsc{OptStop}^{\gamma}$. In this section we want to answer the crucial question whether an optimal strategy exists that yields the supremum $V^{\nu}_{\mathcal{M}}(Z)$ we want to compute. Therefore, we have to concentrate on moment conditions for the adapted stochastic process $Z = (Z_t)_{t\in I}$. For this, we will define the vector spaces we are working with such that we can use results from functional analysis.
Let $q \in (1, \infty]$ be the conjugate Hölder exponent for $p \in [1, \infty)$, that means $(1 - q)p = q$. For $p \in [1, \infty)$ the main statement given in following theorem comprises that there exists an optimal strategy $\gamma \in \mathcal{M}^{\nu}_{I}$ of the problem $\textsc{OptStop}^{\gamma}$ for all processes Z for which the following should be assumed throughout:

Assumption 3.3.3. Let $Z = (Z_t)_{t\in I}$ be the real-valued and $\mathbb{F}$-adapted stochastic process of interest, which describes the underlying price process or a special payoff. Set $Z_T = 0$, if $T \notin I$. Furthermore, we define $\|\cdot\|_{\nu,p}$ as

$$\|Z\|_{\nu,p} := \sup_{\gamma\in\mathcal{M}^{\nu}_{I}} \sqrt[p]{\mathbb{E}\Big[\sum_{t\in I} |Z_t|^p \gamma_t\Big]} \quad \forall Z \in \prod_{t\in I} L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}), \tag{3.3.4}$$

which is a seminorm on $\{Z : \|Z\|_{\nu,p} < \infty\}$. The proof of the norm property is given by Theorem 3.3.13 and this set will be denoted by $X_{\nu,p}$, cf. (3.3.10).

For $p \in [1, \infty)$ and ν, we assume that

(a) $(Z_t)_{t\in I}$ is an element of $\prod_{t\in I} L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R})$ with $\|Z^+\|_{\nu,p} < \infty$.

In addition, for $p = 1$ we assume that

(b) $Z \in X_{\nu,p}$ and its positive part Z^+ can be approximated with respect to $\|\cdot\|_{\nu,p}$ by bounded processes unless otherwise stated. This set will be denoted by $\tilde{X}_{\nu,p}$, cf. (3.3.20).

This guarantees that all values we are interested in are well-posed and finite, see Lemma 3.3.25. The reader should be aware that some definitions also apply to a larger class of processes. However, we will not point this out every time.

Theorem 3.3.5 (Existence of an optimal strategy).
Let $Z \in \prod_{t\in I} L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R})$ with $Z^+ \in X_{\nu,p}$ for $p \in [1, \infty)$ and in addition, for $p = 1$ let $Z \in X_{\nu,p}$ with $Z^+ \in \tilde{X}_{\nu,p}$. Then there always exists an optimal adapted random probability measure $\gamma^* \in \mathcal{M}^{\nu}_{I}$ solving

$$\sup_{\gamma\in\mathcal{M}^{\nu}_{I}} \mathbb{E}[Z_\gamma] = \mathbb{E}[Z_{\gamma^*}].$$

For $p \in [1,\infty)$ the main statement given in Theorem 3.3.5 comprises that there exists an optimal strategy $\gamma \in \mathcal{M}_I^\nu$ of the problem OptStop^γ for all processes Z which satisfy Assumption 3.3.3. To prove this, we use different results from functional analysis. At first, we consider the underlying space, its properties and the subspace containing all Z with the necessary properties given above. It appears that we are working in Banach spaces. We go over to the corresponding dual space and specify a subset which is weak*-compact. With these preliminary considerations, we can finally prove the result.
At first, we want to introduce a convenient notation.

Definition 3.3.6. Let $p \in [1,\infty)$. For $\gamma \in \mathcal{M}_I$ we define

$$|Z|_{\gamma,p} = \Big(\sum_{t\in I} |Z_t|^p \gamma_t\Big)^{1/p}. \tag{3.3.7}$$

Remark 3.3.8. Note that in general $|Z|_{\gamma,p}$ does not agree with $|Z_\gamma| = \big|\sum_{t\in I} Z_t\gamma_t\big|$. However, by the convexity of $\mathbb{R} \ni x \mapsto |x|^p$ and Jensen's inequality,

$$|Z_\gamma| \le |Z|_{\gamma,p} \qquad \text{for all } \gamma \in \mathcal{M}_I. \tag{3.3.9}$$

For $p \in [1,\infty)$ we define the vector space $X_{\nu,p}$ of all $\mathbb{R}$-valued $\mathbb{F}$-adapted processes $Z = (Z_t)_{t\in I}$ with finite norm $\|Z\|_{\nu,p}$ by

$$X_{\nu,p} = \Big\{(Z_t)_{t\in I} \in \prod_{t\in I} L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}) : \ \|Z\|_{\nu,p} < \infty\Big\}, \tag{3.3.10}$$

where the seminorm given in (3.3.4) can be rewritten using Definition 3.3.6 as

$$\|Z\|_{\nu,p} = \sup_{\gamma\in\mathcal{M}_I^\nu} \big\| |Z|_{\gamma,p} \big\|_{L^p}. \tag{3.3.11}$$

The seminormed vector space obtained can be made into a normed vector space in a standard way; one simply takes the quotient space with respect to the kernel of $\|\cdot\|_{\nu,p}$. The resulting normed vector space is denoted by $(X_{\nu,p}, \|\cdot\|_{\nu,p})$.

Remark 3.3.12. For $p \in [1,\infty)$ the norm $\|\cdot\|_{\nu,p}$ depends on the given probability measure ν on I. Let $J = \operatorname{supp}(\nu) = \{t \in I \mid \nu_t > 0\}$ be the support of ν. Note that the restriction on the support of ν does not change the value of the norm of Z. This means that for every $\gamma \in \mathcal{M}_I^\nu$ and $t \in I \setminus J$ the correspondent summands in (3.3.7) would be equal to zero, a.s., because $\nu_t = 0$ implies that $\gamma_t = 0$, a.s.

Note that for every $t \in I$ the space $L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R})$ is a Banach space, where random variables are identified if they are equal, $\mathbb{P}$-a.s. This structure transfers to the previously defined space.

Theorem 3.3.13. *For every $p \in [1,\infty)$, the vector space $(X_{\nu,p}, \|\cdot\|_{\nu,p})$ is a Banach space.*

Proof. See Section 3.3.3. ❑

Lemma 3.3.14. *For $1 \le p < r \le \infty$ we have that $X_{\nu,r} \subseteq X_{\nu,p}$ and for $Z \in X_{\nu,r}$ it holds that $\|Z\|_{\nu,p} \le \|Z\|_{\nu,r}$.*

Proof. It follows immediately by equation (3.3.11) and the property of L^p-spaces that $L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{R}) \supseteq L^r(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{R})$ if $1 \leq p < r$ and for a real-valued measurable function f on Ω it holds that $\|f\|_{L^p} \leq \|f\|_{L^r}$. ❑

Remark 3.3.15. Note that we have for every $\gamma \in \mathcal{M}_I^\nu$ and $p \in [1,\infty)$ by (3.3.9) and (3.3.11)

$$\|Z_\gamma\|_{L^p} \leq \||Z|_{\gamma,p}\|_{L^p} \leq \|Z\|_{\nu,p}.$$

With

$$|Z|_{\gamma,p}^p = \sum_{t\in I} |Z_t|^p \gamma_t \leq \sup_{s\in I} |Z_s|^p \cdot \sum_{t\in I} \gamma_t \leq \sup_{s\in I} |Z_s|^p$$

we get that

$$\|Z\|_{\nu,p} = \sup_{\gamma\in\mathcal{M}_I^\nu} \Big(\mathbb{E}\Big[|Z|_{\gamma,p}^p\Big]\Big)^{1/p} \leq \Big(\mathbb{E}\Big[\sup_{s\in I} |Z_s|^p\Big]\Big)^{1/p}.$$

The last term is exactly the considered norm $\|\cdot\|_{X_{p,p'}}$ with $p \in [1,\infty)$ and $p' = \infty$ in [33, Chapter 3]. The introduced norm is thus an improved bound.

Lemma 3.3.16 (Convergence)**.**
Let $T \notin I$ and $Z \in X_{\nu,p}$, $p \in [1,\infty)$. We define for $u \in I$ and $\gamma \in \mathcal{M}_I^\nu$

$$Z_{\gamma,u} = \sum_{t\in I_{\leq u}} Z_t\gamma_t = \sum_{t\in I} Z_t \mathbb{1}_{I_{\leq u}}(t)\gamma_t = (Z\mathbb{1}_{I_{\leq u}})_\gamma. \tag{3.3.17}$$

Then it holds that

$$\lim_{u\nearrow T} \|Z_\gamma - Z_{\gamma,u}\|_{L^p} = 0.$$

Proof. Using (3.3.17) and Remark 3.3.15 we get for every $u \in I$ and $\gamma \in \mathcal{M}_I^\nu$ that

$$\|Z_\gamma - Z_{\gamma,u}\|_{L^p} \leq \|Z - Z\mathbb{1}_{I_{\leq u}}\|_{\nu,p} = \|Z\mathbb{1}_{I_{>u}}\|_{\nu,p} \leq \|Z\|_{\nu,p} < \infty. \tag{3.3.18}$$

Let $(u_n)_{n\in\mathbb{N}}$ be an ascending sequence such that $u_n \nearrow T$ for $n \to \infty$. Using dominated convergence with $\sum_{t>u_n} |Z_t|^p\gamma_t \leq |Z|_{\gamma,p}^p < \infty$, a.s., we get $\mathbb{E}[\sum_{t>u_n} |Z_t|^p\gamma_t] \to 0$ for $n \to \infty$. Combine this with (3.3.18) implies the statement. ❑

Remark 3.3.19. It is intuitively obvious that an optimal strategy should then be optimal even for a limited time horizon. The converse property is much less obvious (maybe it is possible to get something better by radically changing the strategy on a larger time horizon). The lemma above tell us that it is possible to approximate suitably every strategy and it converges then accordingly fast depending on the p.

There are two sets of processes of $X_{\nu,p}$ we are interested in, for which it turns out that they are closed, linear subspaces. For every $p \in [1,\infty)$ we define these sets of $\mathbb{R}$-valued $\mathbb{F}$-adapted processes $Z = (Z_t)_{t\in I}$ by

$$\tilde{X}_{\nu,p} = \Big\{(Z_t)_{t\in I} \in \prod_{t\in I} L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{R}) \Big| \limsup_{M\to\infty} \sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}\Big[\sum_{t\in I} (|Z_t|^p - M)_+\gamma_t\Big] = 0\Big\}, \tag{3.3.20}$$

where $x_+ = \max\{x, 0\}$ defines the positive part, and

$$\hat{X}_{\nu,p} = \left\{(Z_t)_{t\in I} \in \prod_{t\in I} L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}) \,\middle|\, \Big\{\sum_{t\in I} |Z_t|^p \gamma_t\Big\}_{\gamma\in\mathcal{M}_I^\nu} \text{ uniformly integrable}\right\}. \quad (3.3.21)$$

The set $\tilde{X}_{\nu,p}$ can be defined in two different ways. On the one hand it can be defined by the condition in (3.3.20) which can be considered as a uniform Fatou property. On the other hand it is the closure of the set of all bounded processes in $\|\cdot\|_{\nu,p}$ of $X_{\nu,p}$. In other words, this set includes all processes which can approximated by a bounded process with respect to $\|\cdot\|_{\nu,p}$. In this case, by a bounded process is meant that there exists a bounded representative in the corresponding equivalence class with respect to $\|\cdot\|_{\nu,p}$. This definition is very descriptive. Now we have to show that these two definitions are equivalent.

Lemma 3.3.22. *It holds for the set $\tilde{X}_{\nu,p}$ given in (3.3.20) that*

$$\tilde{X}_{\nu,p} = \overline{\{Z \in X_{\nu,p} \mid Z \text{ bounded}\}}^{\|\cdot\|_{\nu,p}}. \quad (3.3.23)$$

Proof. See Section 3.3.3. ❑

Both sets given in (3.3.20) and (3.3.21) are subsets of $X_{\nu,p}$ and equipped with $\|\cdot\|_{\nu,p}$ they are again Banach spaces as the following theorem shows:

Theorem 3.3.24. *Let $p \in [1, \infty)$.*

(a) The vector space $(\tilde{X}_{\nu,p}, \|\cdot\|_{\nu,p})$ is a Banach space.

(b) The vector space $(\hat{X}_{\nu,p}, \|\cdot\|_{\nu,p})$ is a Banach space.

(c) We have that

$$\tilde{X}_{\nu,p} \subseteq \hat{X}_{\nu,p} \subsetneq X_{\nu,p}.$$

The above Theorem 3.3.24 is only necessary for the proof of the existence in the case of $p = 1$ and therefore the reader is referred to the Subsection 3.3.3. Now, we consider the topological dual space of $(X_{\nu,p}, \|\cdot\|_{\nu,p})$ (respectively $(\tilde{X}_{\nu,p}, \|\cdot\|_{\nu,p})$) which is denoted by $X^*_{\nu,p}$ (respectively $\tilde{X}^*_{\nu,p}$) and is equipped with the operator norm

$$\|\phi\|_{X^*_{\nu,p}} := \sup\{|\phi(Z)| : Z \in X_{\nu,p}, \|Z\|_{\nu,p} \le 1\} \quad \text{for } \phi \in X^*_{\nu,p}$$

(analog $\|\phi\|_{\tilde{X}^*_{\nu,p}}$ for $\phi \in \tilde{X}^*_{\nu,p}$). In addition, $X^*_{\nu,p} \subseteq \tilde{X}^*_{\nu,p}$ and $\|\phi\|_{\tilde{X}^*_{\nu,p}} \le \|\phi\|_{X^*_{\nu,p}}$ for all $\phi \in X^*_{\nu,p}$. Due to [66, Theorem 4.1], $(\tilde{X}^*_{\nu,p}, \|\cdot\|_{\tilde{X}^*_{\nu,p}})$ and $(X^*_{\nu,p}, \|\cdot\|_{X^*_{\nu,p}})$ are again Banach spaces.
For $p \in [1, \infty)$ let $q \in (1, \infty]$ be the conjugate Hölder exponent, that means $1/p + 1/q = 1$. Note that for $p \in [1, \infty)$ the space $(L^q(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}), \|\cdot\|_{L^q}) = (L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}), \|\cdot\|_{L^p})^*$ is a Banach space for every $t \in I$, where random variables are identified if they are equal, $\mathbb{P}$-a.s. We want consider a subset of these linear and continuous functionals which satisfy the extremal equality of Hölder's inequality and is described in the following lemma.

Lemma 3.3.25. *For every $\gamma = (\gamma_t)_{t\in I} \in \mathcal{M}_I^\nu$ the map $\phi_\gamma : X_{\nu,p} \to \mathbb{R}$ defined by*

$$\phi_\gamma(Z) := \mathbb{E}\Big[\sum_{t\in I} Z_t \gamma_t\Big], \quad Z = (Z_t)_{t\in I} \in X_{\nu,p}, \tag{3.3.26}$$

is a well-defined element of $(X^*_{\nu,p}, \|\cdot\|_{X^*_{\nu,p}})$ and satisfies $\|\phi_\gamma\|_{X^*_{\nu,p}} \leq 1$.

Remark 3.3.27. The same considerations are possible for the restriction to the subspace $(\tilde{X}^*_{\nu,p}, \|\cdot\|_{\tilde{X}^*_{\nu,p}})$.

Proof. Using Hölder's inequality, we get for every $\gamma \in \mathcal{M}_I^\nu$ and for all $Z \in X_{\nu,p}$ with $\|Z\|_{\nu,p} \leq 1$ that

$$\begin{aligned} |\phi_\gamma(Z)| &\leq \mathbb{E}\Big[\sum_{t\in I} |Z_t|\gamma_t\Big] = \mathbb{E}\Big[\sum_{t\in I} |Z_t|\gamma_t^{1/p}\cdot\gamma_t^{1/q}\Big] \\ &\leq \Big(\mathbb{E}\Big[\sum_{t\in I} |Z_t|^p\gamma_t\Big]\Big)^{1/p}\Big(\underbrace{\mathbb{E}\Big[\sum_{t\in I}(\gamma_t^{1/q})^q\Big]}_{\leq 1}\Big)^{1/q} \leq \|Z\|_{\nu,p} \leq 1. \end{aligned} \tag{3.3.28}$$

Therefore, ϕ_γ as defined in (3.3.26) is a bounded linear functional, which implies that it is continuous. Thus it is a well-defined element of $(X^*_{\nu,p}, \|\cdot\|_{X^*_{\nu,p}})$ for every $\gamma \in \mathcal{M}_I^\nu$. By the definition of the operator norm we have that $\|\phi_\gamma\|_{X^*_{\nu,p}} \leq 1$. ❑

For $p \in [1,\infty)$, we define

$$B_{\nu,p} = \begin{cases} \{Z \in \tilde{X}_{\nu,1} : \|Z\|_{\nu,1} \leq 1\} & \text{for } p = 1, \\ \{Z \in X_{\nu,p} : \|Z\|_{\nu,p} \leq 1\} & \text{for } p \in (1,\infty). \end{cases}$$

Note that in the case of $p = 1$ we restrict to the smaller space $\tilde{X}_{\nu,1}$ instead of $X_{\nu,1}$. By the theorem of Banach–Alaoglu (see e.g. [66, Theorem 3.15]), we have that the polar set

$$K_{\nu,p} := \begin{cases} \{\phi \in \tilde{X}^*_{\nu,p} : \phi(Z)| \leq 1 \text{ for all } Z \in B_{\nu,p}\} & \text{for } p = 1, \\ \{\phi \in X^*_{\nu,p} : \phi(Z)| \leq 1 \text{ for all } Z \in B_{\nu,p}\} & \text{for } p \in (0,\infty), \end{cases}$$

is weak*-compact. For the proof of the main statement of this section we need the following key lemma:

Lemma 3.3.29. *For $p \in [1,\infty)$ considered in Lemma 3.3.25, the set $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is contained in $K_{\nu,p}$ and weak*-compact, where every ϕ_γ is of the form as in (3.3.26).*

Proof. See Section 3.3.3. ❑

With these preliminary considerations, we can finally prove the main result.

Proof of Theorem 3.3.5. For $p \in (1,\infty)$ let $Z \in X_{\nu,p}$. Then we can rewrite the considered problem by

$$\sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}[Z_\gamma] = \sup_{\gamma\in\mathcal{M}_I^\nu} \phi_\gamma(Z) = \sup_{\psi\in\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}} \psi(Z).$$

As $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is weak*-compact by Lemma 3.3.29 and $\mathcal{M}_I^\nu$ is not empty by Example 3.1.4, there exists a $\gamma^* \in \mathcal{M}_I^\nu$ such that

$$\sup_{\psi\in\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}} \psi(Z) = \phi_{\gamma^*}(Z),$$

as every continuous function on a non-empty compact set attains its supremum on this set (see e.g. [61, Chapter IV.3, p. 99] for compact sets that are Hausdorff). Now, let $Z \in \prod_{t\in I} L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{R})$ with $Z^+ \in X_{\nu,p}$ for $p \in (1,\infty)$. We define $Z_t^{(n)} = \max\{Z_t,-n\}$, $n \in \mathbb{N}$. It holds that $Z^{(n)} \in X_{\nu,p}$, $n \in \mathbb{N}$, and $Z^{(n)} \searrow Z$ as $n \to \infty$. Then we can prove that the functional

$$H : \{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu} \to \mathbb{R}, \quad \phi_\gamma \mapsto \phi_\gamma(Z),$$

is upper semicontinuous w.r.t. the weak* topology on $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$. For $Z \in X_{\nu,p}$ and $Z^+ \in \tilde{X}_{\nu,p}$, there exists a sequence $Z^{(n)} \in \tilde{X}_{\nu,p}$, $n \in \mathbb{N}$, such that $Z^{(n)} \searrow Z$ as $n \to \infty$. Furthermore, we can define the sequence of functionals

$$H_n : \{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu} \to \overline{\mathbb{R}}, \quad \phi_\gamma \mapsto \phi_\gamma(Z^{(n)}), \quad \text{where } \inf_n H_n(\phi_\gamma) = H(\phi_\gamma).$$

For every $n \in \mathbb{N}$, the functional H_n is continuous w.r.t. the weak* topology on $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$, because of Lemma 3.3.25. Then $\phi_\gamma(Z) = \inf_n \phi_\gamma(Z^{(n)})$ and using [3, Lemma 2.41] we get that H is upper semicontinuous. Furthermore, an upper semicontinuous function on a compact set attains a maximum value, and the non-empty set of maximizers is compact, see [3, Theorem 2.43].
For $p = 1$, let $Z \in X_{\nu,p}$ and $Z^+ \in \tilde{X}_{\nu,p}$. The proof follows as before only on the subspace $\tilde{X}_{\nu,p}$. ❑

Finally, we conclude that there is a solution for these induced problems of the form OptStop$^\tau$, thus there also exists an optimal strategy. By Theorem 3.2.8 we have shown that the problem OptStop$^\gamma$ can be connected to a distribution-constrained optimal stopping problem OptStop$^\tau$ given an enlarged filtration $(\tilde{\mathcal{F}}_t)_{t\in I}$. Therefore, there exists an optimal strategy $\tau \in \mathcal{T}_I^\nu$ for all processes Z which satisfy Assumption 3.3.3 and for such a problem OptStop$^\tau$ which can be traced back to a problem of form OptStop$^\gamma$.

Corollary 3.3.30. *For $p \in [1,\infty)$ let $Z \in \prod_{t\in I} L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{R})$ with $Z^+ \in X_{\nu,p}$ and in addition, for $p = 1$ let $Z \in X_{\nu,p}$ with $Z^+ \in \tilde{X}_{\nu,p}$. Then there always exists an optimal strategy $\tau^* \in \mathcal{T}_I^\nu$ for all problems OptStop$^\tau$, which emerges from a problem OptStop$^\gamma$ by means of the Theorem 3.2.8, such that τ^* solves*

$$\sup_{\tau\in\mathcal{T}_I^\nu} \mathbb{E}[Z_\tau] = \mathbb{E}[Z_{\tau^*}].$$

Proof. It follows by the combination of Theorem 3.3.5 and Theorem 3.2.8. ❑

3.3.2. Generalization

Let us introduce the corresponding notation in a slightly more general setting, where $\mathbb{K}$ denotes either $\mathbb{R}$ or $\mathbb{C}$. For exponents $p,q \in [1,\infty]$ define

$$\mathcal{X}_{\nu,p,q} = \Big\{ Z = (Z_t)_{t\in I} \in \prod_{t\in I} \mathcal{L}^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K}) \,\Big|\, \|Z\|_{\nu,p,q} < \infty \Big\}, \tag{3.3.31}$$

where $\mathcal{L}^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$ denotes the set of $\mathbb{K}$-valued $\mathcal{F}_t$-measurable random variables X with $\mathbb{E}[|X|^p] < \infty$, respectively $\mathbb{E}[\operatorname{ess\,sup}|X|] < \infty$ when $p = \infty$, (we pass to L^p when we identify $\mathbb{P}$-a.s. equal ones) and

$$\begin{aligned} \|Z\|_{\nu,p,q} &= \sup_{\gamma\in\mathcal{M}_I^\nu} \begin{cases} \sqrt[q]{\sum_{t\in I} \|Z_t\|^q_{\mathcal{L}^p(\gamma_t\mathbb{P})}} & \text{if } q\in[1,\infty), \\ \sup_{t\in I} \|Z_t\|_{\mathcal{L}^p(\gamma_t\mathbb{P})} & \text{if } q = \infty, \end{cases} \\ &= \sup_{\gamma\in\mathcal{M}_I^\nu} \Big\| \|Z_\cdot\|_{\mathcal{L}^p(\gamma_\cdot\mathbb{P})} \Big\|_{l^q(I)}. \end{aligned} \tag{3.3.32}$$

Here, for $t \in I$ and $\gamma \in \mathcal{M}_I^\nu$, we write $\gamma_t\mathbb{P}$ for the substochastic measure on $(\Omega,\mathcal{F}_t)$ which has density γ_t with respect to $\mathbb{P}$. Then $\|Z_t\|_{\mathcal{L}^p(\gamma_t\mathbb{P})}$ is a seminorm on $\mathcal{L}^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$ and $\prod_{s\in I}\mathcal{L}^p(\Omega,\mathcal{F}_s,\mathbb{P};\mathbb{K})$ for each $t \in I$. To calculate $\|Z\|_{\nu,p,q}$, we then take the l^q-norm of $I \ni t \mapsto \|Z_t\|_{\mathcal{L}^p(\gamma_t\mathbb{P})}$ with respect to the counting measure on I, and finally the l^∞-norm of $\mathcal{M}_I^\nu \ni \gamma \mapsto \big\|\|Z_\cdot\|_{\mathcal{L}^p(\gamma_\cdot\mathbb{P})}\big\|_{l^q(I)}$ with respect to the counting measure on $\mathcal{M}_I^\nu$. Hence (3.3.32) defines a seminorm on $\mathcal{X}_{\nu,p,q}$ and $\mathcal{X}_{\nu,p,q}$ is a vector space. Identifying $Z,Z' \in \mathcal{X}_{\nu,p,q}$ with $\|Z - Z'\|_{\nu,p,q} = 0$, we get a Banach space, see Theorem 3.3.47, which we denote by $(X_{\nu,p,q}, \|\cdot\|_{\nu,p,q})$.
We shall also use the subset $\tilde{\mathcal{X}}_{\nu,p,q}$ of all $Z \in \mathcal{X}_{\nu,p,q}$, for which there is a sequence $(Z^n)_{n\in\mathbb{N}}$ of bounded processes in $\mathcal{X}_{\nu,p,q}$ such that $\|Z - Z^n\|_{\nu,p,q} \to 0$ as $n \to \infty$. Of course, $\tilde{\mathcal{X}}_{\nu,p,q}$ and the corresponding set $\tilde{X}_{\nu,p,q}$ of equivalence classes are vector spaces with seminorm and norm $\|\cdot\|_{\nu,p,q}$, respectively. A standard argument shows that $\tilde{X}_{\nu,p,q}$ is a closed subset of $X_{\nu,p,q}$ and, therefore, itself a Banach space. For the case $p = q = 1$, Example 3.3.90 below shows that $\tilde{\mathcal{X}}_{\nu,1,1}$ is necessary because there is no solution of the problem OptStop$^\gamma$ for some process in $\mathcal{X}_{\nu,1,1}$.
For later reference consider $q \le \tilde{q}$ in $[1,\infty)$ and recall that for $x = (x_t)_{t\in I} \in l^q(I)$ (with respect to the counting measure on I) satisfying $\|x\|_{l^q} \le 1$, we have $|x_t| \le 1$ for each $t \in I$, hence $\|x\|^{\tilde{q}}_{l^{\tilde{q}}} = \sum_{t\in I}|x_t|^{\tilde{q}} \le \sum_{t\in I}|x_t|^q = \|x\|^q_{l^q} \le 1$. For general $x \in l^q(I)\setminus\{0\}$, apply this result to αx with $\alpha := 1/\|x\|_{l^q}$, implying that $\alpha^{\tilde{q}}\|x\|^{\tilde{q}}_{l^{\tilde{q}}} \le 1$. Hence

$$l^q(I) \subseteq l^{\tilde{q}}(I) \quad \text{and} \quad \|x\|_{l^{\tilde{q}}} \le \|x\|_{l^q}, \quad x \in l^q(I). \tag{3.3.33}$$

Since $|x_t| \le \|x\|_{l^q}$ for all $t \in I$ and $x \in l^q(I)$, (3.3.33) extends to all $q \le \tilde{q}$ in $[1,\infty]$.
The next result, which is the main one of this section, shows that under appropriate moment conditions on the process Z, which in turn depends on the decay of the probability distribution ν, problem OptStop$^\gamma$ has a solution.

Theorem 3.3.34 (Existence of an optimal strategy). *Consider a real-valued process $Z \in \mathcal{X}_{\nu,p,q}$, where $p \in [1,\infty)$ and $q \in [1,\infty]$, and assume that*

(a) *either* $p = q = 1$ *and* $Z^+ \in \tilde{\mathcal{X}}_{\nu,1,1}$,

(b) *or* $p > 1$ *and, with* $p' = \frac{p}{p-1}$ *and* q' *denoting the conjugate exponents of* p *and* q*, respectively,* ν *satisfies the decay condition*

$$\sum_{t\in I} \nu_t^{q'/p'} < \infty, \tag{3.3.35}$$

then the random variable Z_γ *is a well-defined element of* $L^1(\Omega, \mathcal{F}, \mathbb{P})$ *for every* $\gamma \in \mathcal{M}_I^\nu$ *and there exists an optimal adapted random probability measure* $\gamma^* \in \mathcal{M}_I^\nu$ *solving*

$$\sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}\big[Z_\gamma\big] = \mathbb{E}\big[Z_{\gamma^*}\big].$$

Remark 3.3.36. By Lemma 3.3.48 below, the set $\mathcal{X}_{\nu,p,q}$ of processes gets bigger when p decreases or q increases. The decay condition (3.3.35), which is trivially satisfied when $q'/p' \geq 1$, puts a threshold to this procedure.

Remark 3.3.37. Suppose we know for a $c \in (0,1)$ that $\sum_{t\in I} \nu_t^c < \infty$ and $p \in (1,\infty)$ is so small that a condition on q is required, i.e., $1/p' < c$, meaning that $p < 1/(1-c)$. Then we know that (3.3.35) is certainly satisfied if $q' \geq cp'$, meaning that

$$q \leq 1 + \frac{1}{cp'-1} = \frac{cp}{1-(1-c)p}.$$

Remark 3.3.38. For the missing pairs $(p,q) \in [1,\infty) \times [1,\infty]$ not be treated in the Theorem 3.3.34, namely $p = 1$ and $q \in (1,\infty]$ as well as $p > 1$ and $q \in (1,\infty]$ not satisfying (3.3.35), there always exist $Z \in \mathcal{X}_{\nu,p,q}$ and $\gamma \in \mathcal{M}_I^\nu$ such that $\sum_{t\in I} Z_t\gamma_t = \infty$ on Ω, see Examples 3.2.5 and 3.2.7 below.

Remark 3.3.39. To motivate the use of the spaces $\mathcal{X}_{\nu,p,q}$ and the decay condition (3.3.35) in Theorem 3.3.34, consider a corresponding $Z \in \mathcal{X}_{\nu,p,q}$.

(a) If $p = q = 1$, then (3.3.32) implies

$$\sum_{t\in I} \mathbb{E}[|Z_t|\gamma_t] \leq \|Z\|_{\nu,1,1}, \quad \gamma \in \mathcal{M}_I^\nu.$$

(b) Consider $p \in (1,\infty)$ and $q \in [1,\infty]$ satisfying (3.3.35). Then by Hölder's inequality for exponents p and p',

$$\begin{aligned}\mathbb{E}[|Z_t|\gamma_t] &= \mathbb{E}[|Z_t|\gamma_t^{1/p}\gamma_t^{1/p'}]\\ &\leq (\mathbb{E}[|Z_t|^p\gamma_t])^{1/p}(\mathbb{E}[\gamma_t])^{1/p'}, \quad t\in I, \gamma \in \mathcal{M}_I^\nu.\end{aligned}$$

Using that $\mathbb{E}[\gamma_t] = \nu_t \leq 1$ by Definition 3.1.1(e), we get for every $\gamma \in \mathcal{M}_I^\nu$ and $I' \subseteq I$ in the case $q = 1$

$$\sum_{t\in I'} \mathbb{E}[|Z_t|\gamma_t] \leq \sum_{t\in I'} (\mathbb{E}[|Z_t|^p\gamma_t])^{1/p} \cdot \sup_{t\in I'} \nu_t^{1/p'} \leq \|Z\|_{\nu,p,1} \sup_{t\in I'} \nu_t^{1/p'}, \tag{3.3.40}$$

in the case $q = \infty$ (using $q' = 1$ and (3.3.35))

$$\sum_{t\in I'} \mathbb{E}[|Z_t|\gamma_t] \leq \sup_{s\in I'} (\mathbb{E}[|Z_s|^p\gamma_s])^{1/p} \sum_{t\in I'} \nu_t^{q'/p'} \leq \|Z\|_{\nu,p,\infty} \sum_{t\in I'} \nu_t^{q'/p'}, \tag{3.3.41}$$

and in the remaining case $q \in (1,\infty)$ another use of Hölder's inequality for the exponents q and q' as well as (3.3.32) yield

$$\begin{aligned}\sum_{t\in I'} \mathbb{E}[|Z_t|\gamma_t] &\le \biggl(\sum_{t\in I'} (\mathbb{E}[|Z_t|^p\gamma_t])^{q/p}\biggr)^{1/q} \biggl(\sum_{t\in I'} \nu_t^{q'/p'}\biggr)^{1/q'} \\ &\le \|Z\|_{\nu,p,q} \biggl(\sum_{t\in I'} \nu_t^{q'/p'}\biggr)^{1/q'}.\end{aligned} \tag{3.3.42}$$

Hence in both cases (a) and (b), in the latter one using (3.3.40), (3.3.41) or (3.3.42), respectively, with $I' = I$, the series $\sum_{t\in I} Z_t\gamma_t$ defining Z_γ converges absolutely $\mathbb{P}$-a.s., $Z_\gamma \in L^1(\Omega,\mathcal{F},\mathbb{P})$ and $\{Z_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is L^1-bounded. The above calculation also shows that each $\gamma \in \mathcal{M}_I^\nu$ induces a bounded linear (hence continuous) functional

$$\mathcal{X}_{\nu,p,q} \ni Z \mapsto \phi_\gamma(Z) := \mathbb{E}[Z_\gamma].$$

For $(p,q) \in [1,\infty)\times[1,\infty]$ the main statement given in Theorem 3.3.34 comprises that there exists an optimal strategy $\gamma \in \mathcal{M}_I^\nu$ of the problem OptStop^γ for all processes Z which satisfy Theorem 3.3.34. To prove this, it suffices to find a topology on $\mathcal{M}_I^\nu$ which turns it into a compact set, and show that for each $Z \in \mathcal{X}_{\nu,p,q}$ the map $\mathcal{M}_I^\nu \ni \gamma \mapsto \mathbb{E}[Z_\gamma]$ is continuous with respect to this topology. We shall apply methods from functional analysis for this purpose. At first, we consider the underlying space, its properties and the subspace containing all Z with the necessary properties given above. It appears that we are working in Banach spaces. We go over to the corresponding dual space and specify a subset which is weak*-compact. With these preliminary considerations, we can finally prove the result.

Remark 3.3.43. The reader interested in the time-continuous case is referred to Part II. There we use a different approach and view of the problem, and the existence of an optimal strategy is proven through ideas and concepts from the theory of optimal transport, see Chapter 7.

Remark 3.3.44 (Restriction to support of ν). For $p,q \in [1,\infty]$ the seminorm $\|\cdot\|_{\nu,p,q}$ on $\mathcal{X}_{\nu,p,q}$ depends on the given probability measure ν on I. Let $J = \operatorname{supp}(\nu) = \{t \in I \mid \nu_t > 0\}$ be the support of ν. Note that the restriction on the support of ν does not change the value of the seminorm of Z defined in (3.3.32). This means that for every $\gamma \in \mathcal{M}_I^\nu$ and $t \in I\setminus J$ the corresponding term $\|Z_t\|_{\mathcal{L}^p(\gamma_t\mathbb{P})}$ in (3.3.32) is zero, because $\nu_t = 0$ implies that $\gamma_t = 0$ $\mathbb{P}$-a.s.

Remark 3.3.45 (Basic $\mathcal{L}^p$-inequalities). For future reference, let us show here that, for every $\gamma \in \mathcal{M}_I^\nu$, $p \in [1,\infty]$, $t \in I$, and $Z_t \in \mathcal{L}^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$,

$$\|Z_t\gamma_t\|_{\mathcal{L}^p(\mathbb{P})} \le \|Z_t\|_{\mathcal{L}^p(\gamma_t\mathbb{P})} \le \|Z_t\|_{\mathcal{L}^p(\mathbb{P})}. \tag{3.3.46}$$

Recall that $0 \le \gamma_t \le 1$ due to Definition 3.1.1(a) and (b). If $p \in [1,\infty)$, then $\gamma_t^p \le \gamma_t \le 1$, hence $\mathbb{E}[|Z_t\gamma_t|^p] \le \mathbb{E}[|Z_t|^p\gamma_t] \le \mathbb{E}[|Z_t|^p]$, which implies (3.3.46). If $p = \infty$, define $c = \|Z_t\|_{\mathcal{L}^\infty(\gamma_t\mathbb{P})}$. This means that $\{|Z_t| > c,\ \gamma_t > 0\}$ is a $\mathbb{P}$-null set. Since it contains $\{|Z_t|\gamma_t > c\}$, it follows that $\|Z_t\gamma_t\|_{\mathcal{L}^\infty(\mathbb{P})} \le c$, proving the first inequality in (3.3.46). The second one follows similarly, using $\{|Z_t|\gamma_t > c\} \subseteq \{|Z_t| > c\}$ for $c := \|Z_t\|_{\mathcal{L}^\infty(\mathbb{P})}$.

For every $p \in [1,\infty]$ and $t \in I$ the space $L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$ with norm $\|\cdot\|_{L^p(\mathbb{P})}$ is a Banach space (with random variables identified if they are $\mathbb{P}$-a.s. equal). This structure transfers to the quotient spaces $X_{\nu,p,q}$ derived from $\mathcal{X}_{\nu,p,q}$ given in (3.3.31).

Theorem 3.3.47. *For every probability measure ν on I and every choice of exponents $p,q \in [1,\infty]$, the normed vector space $(X_{\nu,p,q},\|\cdot\|_{\nu,p,q})$ is a Banach space.*

Proof. See Section 3.3.3. ❑

Lemma 3.3.48. *Consider $p,q \in [1,\infty)$.*

(a) If $\tilde{p} \in [1,p]$, then $\mathcal{X}_{\nu,p,q} \subseteq \mathcal{X}_{\nu,\tilde{p},q}$ and

$$\|Z\|_{\nu,\tilde{p},q} \le \|Z\|_{\nu,p,q}, \quad Z \in \mathcal{X}_{\nu,p,q}. \tag{3.3.49}$$

(b) If $\tilde{q} \in [q,\infty)$, then $\mathcal{X}_{\nu,p,q} \subseteq \mathcal{X}_{\nu,p,\tilde{q}}$ and

$$\|Z\|_{\nu,p,\tilde{q}} \le \|Z\|_{\nu,p,q}, \quad Z \in \mathcal{X}_{\nu,p,q}. \tag{3.3.50}$$

(c) Given $r \in (1,p]$ and $s \in (1,q]$, define $C_{p,q,r,s} = \sum_{t\in I} \nu_t^{qrs'/psr'}$, where r' and s' denote the conjugate exponents of r and s, respectively. If $C_{p,q,r,s} < \infty$, then $\mathcal{X}_{\nu,p,q} \subseteq \mathcal{X}_{\nu,p/r,q/s}$ and

$$\|Z\|_{\nu,p/r,q/s} \le C_{p,q,r,s}^{s/qs'} \|Z\|_{\nu,p,q}, \quad Z \in \mathcal{X}_{\nu,p,q}. \tag{3.3.51}$$

Proof. Note that $\mathcal{L}^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K}) \subseteq \mathcal{L}^{\tilde{p}}(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$ for $1 \le \tilde{p} \le p$ and each $t \in I$, because $\mathbb{P}$ is a finite measure.
(a) Let $\gamma \in \mathcal{M}_I^\nu$. Then Jensen's inequality applied to the substochastic probability measure $\gamma_t\mathbb{P}$ yields $(\mathbb{E}[|Z_t|^{\tilde{p}}\gamma_t])^{1/\tilde{p}} \le (\mathbb{E}[|Z_t|^p\gamma_t])^{1/p}$ for each $t \in I$, hence (3.3.49) follows via (3.3.32).
(b) For every $\gamma \in \mathcal{M}_I^\nu$, apply inequality (3.3.33) to $x_\gamma = (x_{\gamma,t})_{t\in I}$ with $x_{\gamma,t} := (\mathbb{E}[|Z_t|^p\gamma_t])^{1/p}$. Taking the supremum over $\gamma \in \mathcal{M}_I^\nu$ yields (3.3.50).
(c) By Hölder's inequality with exponents r and r',

$$\mathbb{E}\Big[|Z_t|^{p/r}\gamma_t\Big] = \mathbb{E}\Big[|Z_t|^{p/r}\gamma_t^{1/r}\gamma_t^{1/r'}\Big] \le \Big(\mathbb{E}[|Z_t|^p\gamma_t]\Big)^{1/r}\Big(\mathbb{E}[\gamma_t]\Big)^{1/r'}, \quad \gamma \in \mathcal{M}_I^\nu,\ t \in I. \tag{3.3.52}$$

Using this result and $\mathbb{E}[\gamma_t] = \nu_t$ from Definition 3.1.1(e), another application of Hölder's inequality with exponents s and s' yields, for every $\gamma \in \mathcal{M}_I^\nu$,

$$\begin{aligned}\sum_{t\in I}\Big(\mathbb{E}\Big[|Z_t|^{p/r}\gamma_t\Big]\Big)^{qr/ps} &\le \sum_{t\in I}\Big(\mathbb{E}\Big[|Z_t|^p\gamma_t\Big]\Big)^{q/ps}\nu_t^{qr/psr'} \\ &\le \bigg(\sum_{t\in I}\Big(\mathbb{E}\Big[|Z_t|^p\gamma_t\Big]\Big)^{q/p}\bigg)^{1/s} C_{p,q,r,s}^{1/s'}.\end{aligned}$$

Raising this inequality to the power s/q and using (3.3.32), estimate (3.3.51) follows. ❑

Lemma 3.3.53 (Convergence)**.**
Let $p \in [1,\infty]$ and $Z \in \mathcal{X}_{\nu,p,1}$. Then

$$\|Z_\gamma\|_{L^p(\mathbb{P})} \le \|Z\|_{\nu,p,1}, \quad \gamma \in \mathcal{M}_I^\nu. \tag{3.3.54}$$

For $\gamma \in \mathcal{M}_I^\nu$ and a sequence $(J_n)_{n\in\mathbb{N}}$ of increasing subsets of I with $\bigcup_{n\in\mathbb{N}} J_n = I$ define the approximations

$$Z_{\gamma,J_n} = \sum_{t\in J_n} Z_t\gamma_t = \sum_{t\in I} Z_t \mathbb{1}_{J_n}(t)\gamma_t = (Z\mathbb{1}_{J_n})_\gamma, \quad n \in \mathbb{N}. \tag{3.3.55}$$

Then

$$\lim_{n\to\infty} \|Z_\gamma - Z_{\gamma,J_n}\|_{L^p(\mathbb{P})} = 0. \tag{3.3.56}$$

Proof. To prove (3.3.54), we use Minkowski's integral inequality (for $\mathbb{P}$ and the counting measure on I) and then (3.3.46) and (3.3.32) to get that, for every $\gamma \in \mathcal{M}_I^\nu$,

$$\|Z_\gamma\|_{L^p(\mathbb{P})} = \Big\|\sum_{t\in I} Z_t\gamma_t\Big\|_{L^p(\mathbb{P})} \le \sum_{t\in I} \|Z_t\gamma_t\|_{L^p(\mathbb{P})} \le \sum_{t\in I} \|Z_t\|_{\mathcal{L}^p(\gamma_t,\mathbb{P})} \le \|Z\|_{\nu,p,1}. \tag{3.3.57}$$

To prove (3.3.56), use (3.3.55) to write $Z_\gamma - Z_{\gamma,J_n} = (Z\mathbb{1}_{I\setminus J_n})_\gamma$ and the first two estimates of (3.3.57) to get that

$$\|Z_\gamma - Z_{\gamma,J_n}\|_{L^p(\mathbb{P})} \le \sum_{t\in I\setminus J_n} \|Z_t\|_{\mathcal{L}^p(\gamma_t,\mathbb{P})}. \tag{3.3.58}$$

Since the last series in (3.3.57) converges in $\mathbb{R}$, the claim (3.3.56) follows from (3.3.58). ❑

Remark 3.3.59. It is intuitively obvious that an optimal strategy should then be optimal even for a limited time horizon. The converse property is much less obvious (maybe it is possible to get something better by radically changing the strategy on a larger time horizon). The lemma above tell us that it is possible to suitably approximate every strategy and then it converges accordingly fast depending on the p.
For example, let $T \notin I$ and $(u_n)_{n\in\mathbb{N}} \subseteq I$ be an increasing sequence in I with $u_n \le u_{n+1}$ for all $n \in \mathbb{N}$. Then $u_n \nearrow T$ for $n \to \infty$ and we can choose $J_n = I_{\le u_n} \cap J$ for all $n \in \mathbb{N}$.

Lemma 3.3.60. *Consider $p, q \in (1,\infty)$ and let q' denote the conjugate exponent of q. Given $r \in (1,p]$, define $K_{p,q,r} = \sum_{t\in I} \nu_t^{q'(r-1)/p}$. If $K_{p,q,r} < \infty$, then*

$$\|Z_\gamma\|_{L^{p/r}(\mathbb{P})} \le K_{p,q,r}^{1/q'} \|Z\|_{\nu,p,q}, \quad \gamma \in \mathcal{M}_I^\nu, Z \in \mathcal{X}_{\nu,p,q}.$$

Remark 3.3.61. Suppose we know for a $c \in (0,1)$ that $\sum_{t\in I} \nu_t^c < \infty$ and $p \in (1,\infty)$ is so small that a condition on q is required, i.e., $(r-1)/p < c$. Then we know that (3.3.35) is certainly satisfied if $q' \ge cp'$, meaning that

$$q \le \frac{cp}{cp-(r-1)}.$$

Proof of Lemma 3.3.60. Using Minkowski's integral inequality and that $|\gamma_t|^{p/r} \le \gamma_t$ for all $t \in I$, we get that

$$\|Z_\gamma\|_{L^{p/r}(\mathbb{P})} = \left(\mathbb{E}\left[\Big|\sum_{t\in I} Z_t\gamma_t\Big|^{p/r}\right]\right)^{r/p} \le \sum_{t\in I}\left(\mathbb{E}[|Z_t|^{p/r}\gamma_t]\right)^{r/p}. \tag{3.3.62}$$

Using Hölder's inequality with conjugate exponents r and $r/(r-1)$ like in (3.3.52) and that $\mathbb{E}[\gamma_t] = \nu_t$ we have

$$\mathbb{E}\left[|Z_t|^{p/r}\gamma_t\right] \le \left(\mathbb{E}\left[|Z_t|^p\gamma_t\right]\right)^{1/r}\nu_t^{(r-1)/r}, \quad t \in I.$$

Substitution into (3.3.62) and applying Hölder's inequality to the series with conjugate exponents q and q' shows that

$$\begin{aligned}\|Z_\gamma\|_{L^{p/r}(\mathbb{P})} &\le \sum_{t\in I}\left(\mathbb{E}\left[|Z_t|^p\gamma_t\right]\right)^{1/p}\nu_t^{(r-1)/p}\\ &\le \left(\sum_{t\in I}\left(\mathbb{E}\left[|Z_t|^p\gamma_t\right]\right)^{q/p}\right)^{1/q}\left(\sum_{t\in I}\nu_t^{q'(r-1)/p}\right)^{1/q'}.\end{aligned}$$

Using (3.3.32), the claimed inequality follows. ❑

There is a subset of processes of $X_{\nu,p,q}$ we are interested in, especially in the case $p = q = 1$, for which it turns out that it is a closed linear subspace.
Recall that the subset $\tilde{\mathcal{X}}_{\nu,p,q}$ is the set of all $Z \in \mathcal{X}_{\nu,p,q}$, for which there is a sequence $(Z^n)_{n\in\mathbb{N}}$ of bounded processes in $\mathcal{X}_{\nu,p,q}$ such that $\|Z - Z^n\|_{\nu,p,q} \to 0$ as $n \to \infty$. For every $p \in [1,\infty)$ and $q \in [1,\infty]$ we define these sets of $\mathbb{K}$-valued $\mathbb{F}$-adapted processes $Z = (Z_t)_{t\in I}$ by

$$\tilde{\mathcal{X}}_{\nu,p,q} = \overline{\{Z \in \mathcal{X}_{\nu,p,q} \mid Z \text{ bounded}\}}^{\|\cdot\|_{\nu,p,q}}. \tag{3.3.63}$$

Of course, $\tilde{\mathcal{X}}_{\nu,p,q}$ and the corresponding set $\tilde{X}_{\nu,p,q}$ of equivalence classes are vector spaces with seminorm and norm $\|\cdot\|_{\nu,p,q}$, respectively. A standard argument shows that $\tilde{X}_{\nu,p,q}$ is a closed subset of $X_{\nu,p,q}$ and, therefore, itself a Banach space.

Theorem 3.3.64. *Let $p \in [1,\infty)$ and $q \in [1,\infty]$.*

(a) The vector space $(\tilde{X}_{\nu,p,q}, \|\cdot\|_{\nu,p,q})$ is a Banach space.

(b) We have that

$$\tilde{X}_{\nu,p,q} \subsetneqq X_{\nu,p,q}.$$

The above Theorem 3.3.64 are only necessary for the proof of the existence in the case of $p = q = 1$ and therefore the reader is referred to the Section 3.3.3.
Now, we consider the topological dual spaces of $X_{\nu,p,q}$ and its subspace $\tilde{X}_{\nu,p,q}$ with respect to the norm $\|\cdot\|_{\nu,p,q}$. These are denoted by $X^*_{\nu,p,q}$ and $\tilde{X}^*_{\nu,p,q}$, respectively, and are equipped with the corresponding operator norms

$$\|\phi\|_{X^*_{\nu,p,q}} := \sup\{|\phi(Z)| : Z \in X_{\nu,p,q}, \|Z\|_{\nu,p,q} \le 1\}, \quad \phi \in X^*_{\nu,p,q},$$

and $\|\cdot\|_{\tilde{X}^*_{\nu,p,q}}$. In addition, $X^*_{\nu,p,q} \subseteq \tilde{X}^*_{\nu,p,q}$ and $\|\phi\|_{\tilde{X}^*_{\nu,p,q}} \le \|\phi\|_{X^*_{\nu,p,q}}$ for all $\phi \in X^*_{\nu,p,q}$. Due to [66, Theorem 4.1], $(X^*_{\nu,p,q}, \|\cdot\|_{X^*_{\nu,p,q}})$ and $(\tilde{X}^*_{\nu,p,q}, \|\cdot\|_{\tilde{X}^*_{\nu,p,q}})$ are again Banach spaces. The following lemma is a corresponding reformulation of Remark 3.3.39 taking equivalence classes into account.

Lemma 3.3.65. *For every $\gamma \in \mathcal{M}_I^\nu$ and $(p,q) \in [1,\infty) \times [1,\infty]$ with either $p = q = 1$ or $p > 1$ satisfying the decay condition (3.3.35), the map $X_{\nu,p,q} \ni Z \mapsto \phi_\gamma := \mathbb{E}[Z_\gamma]$ is a well-defined element of $X^*_{\nu,p,q}$ with*

$$\|\phi_\gamma\|_{X^*_{\nu,p,q}} \leq \begin{cases} 1 & \text{if } p = q = 1, \\ \left(\sum_{t\in I} \nu_t^{q'/p'}\right)^{1/q'} & \text{if } p > 1,\ q \in [1,\infty] \text{ satisfy (3.3.35).} \end{cases}$$

For $p,q = 1$, define $V_{\nu,p,q} = \{Z \in \tilde{X}_{\nu,p,q} : \|Z\|_{\nu,p,q} \leq 1\}$, and for $p > 1$ and $q \in [1,\infty]$ satisfying (3.3.35) define

$$V_{\nu,p,q} = \left\{Z \in X_{\nu,p,q} : \|Z\|_{\nu,p,q} \leq \left(\sum_{t\in I} \nu_t^{q'/p'}\right)^{-1/q'}\right\}.$$

In both cases $V_{\nu,p,q}$ is a neighborhood or the origin. By the Banach–Alaoglu theorem (see e.g. [66, Theorem 3.15]), the polar set

$$K_{\nu,p,q} := \begin{cases} \{\phi \in \tilde{X}^*_{\nu,p,q} : |\phi(Z)| \leq 1 \text{ for every } Z \in V_{\nu,p,q}\} & \text{if } p = q = 1, \\ \{\phi \in X^*_{\nu,p,q} : |\phi(Z)| \leq 1 \text{ for every } Z \in V_{\nu,p,q}\} & \text{if } p > 1,\ q \in [1,\infty] \text{ satisfy (3.3.35),} \end{cases}$$

is weak*-compact. For the proof of the main statement of this section we need the following key lemma:

Lemma 3.3.66. *For the pairs $(p,q) \in [1,\infty) \times [1,\infty]$ considered in Lemma 3.3.65, the set $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is contained in $K_{\nu,p,q}$ and weak*-compact.*

Proof. See Section 3.3.3. ❑

With these preliminary considerations, we can finally prove the main result.

Proof of Theorem 3.3.34. First we consider the case $(p,q) \in (1,\infty)\times[1,\infty]$ satisfying (3.3.35). We pass to the corresponding sets $X_{\nu,p,q}$ of equivalence classes. Thus, let $Z \in X_{\nu,p,q}$. Then we can rewrite the considered problem by using the notation of Lemma 3.3.65 as

$$\sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}[Z_\gamma] = \sup_{\gamma\in\mathcal{M}_I^\nu} \phi_\gamma(Z) = \sup_{\psi\in\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}} \psi(Z).$$

As $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is weak*-compact by Lemma 3.3.66 and $\mathcal{M}_I^\nu$ is not empty by Example 3.1.4, there exists a $\gamma^* \in \mathcal{M}_I^\nu$ such that

$$\sup_{\psi\in\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}} \psi(Z) = \phi_{\gamma^*}(Z),$$

as every continuous function on a non-empty compact set attains its supremum on this set (see e.g. [61, Chapter IV.3, p. 99] for compact sets that are Hausdorff).
For $p = q = 1$ let $Z \in X_{\nu,p,q}$ and $Z^+ \in \tilde{X}_{\nu,p,q}$. We define $Z_t^{(n)} = \max\{Z_t, -n\}$, $n \in \mathbb{N}$. It holds that $Z^{(n)} \in \tilde{X}_{\nu,p,q}$, $n \in \mathbb{N}$, and $Z^{(n)} \searrow Z$ as $n \to \infty$. Then we can prove that the functional

$$H : \{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu} \to \mathbb{R}, \quad \phi_\gamma \mapsto \phi_\gamma(Z),$$

is upper semicontinuous w.r.t. the weak*-topology on $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$. Furthermore, we can define the sequence of functionals

$$H_n : \{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu} \to \overline{\mathbb{R}}, \quad \phi_\gamma \mapsto \phi_\gamma(Z^{(n)}), \quad \text{where } \inf_n H_n(\phi_\gamma) = H(\phi_\gamma).$$

For every $n \in \mathbb{N}$, the functional H_n is continuous w.r.t. the weak*-topology on $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$, because of Lemma 3.3.65. Then $\phi_\gamma(Z) = \inf_n \phi_\gamma(Z^{(n)})$ and using [3, Lemma 2.41] we get that H is upper semicontinuous. Furthermore, an upper semicontinuous function on a compact set attains a maximum value, and the non-empty set of maximizers is compact, see [3, Theorem 2.43].

❑

3.3.3. Outstanding Proofs

Proofs of Section 3.3.1

In this section the remaining proofs of Section 3.3.1 are to be delivered.

Proof of Theorem 3.3.13.

(i) $\|\cdot\|_{\nu,p}$ is a norm: It is trivial that $|\cdot|_{\nu,p}$ given in Definition 3.3.6 is a norm. Using the norm properties of $|\cdot|_{\nu,p}$ and $\|\cdot\|_{L^p}$ we get that $\|\cdot\|_{\nu,p}$ is a norm for $p\in[1,\infty)$ by the representation (3.3.11). In this connection the absolute homogeneity follows by the iterative composition of the absolute homegeneity of $|\cdot|_{\nu,p}$ and $\|\cdot\|_{L^p}$. We get the triangle inequality by using the Minkowski inequality for $|\cdot|_{\nu,p}$ and the triangle inequality of $\|\cdot\|_{L^p}$.

(ii) Completeness: For a fixed $t\in I$ we have on the one hand for every $\gamma\in\mathcal{M}_I^\nu$ that

$$\Big(\mathbb{E}[|Z_t|^p\gamma_t]\Big)^{1/p}\le\Big(\mathbb{E}\Big[\sum_{s\in I}|Z_s|^p\gamma_s\Big]\Big)^{1/p}\le\sup_{\tilde\gamma\in\mathcal{M}_I^\nu}\Big(\mathbb{E}\Big[|Z|^p_{\tilde\gamma,p}\Big]\Big)^{1/p}=\|Z\|_{\nu,p}$$

and on the other one we get for the deterministic choice $\gamma_t=\nu_t\mathbb{1}_\Omega$ that $\mathbb{E}[|Z_t|^p\gamma_t]=\nu_t\mathbb{E}[|Z_t|^p]$. Due to Remark 3.3.12 we restrict our considerations on the support $J=\operatorname{supp}(\nu)$ of ν. Combining both thoughts we have for $t\in J$ that

$$\|Z_t\|_{L^p}=\Big(\mathbb{E}[|Z_t|^p]\Big)^{1/p}\le\nu_t^{-1/p}\|Z\|_{\nu,p},\quad Z\in X_{\nu,p}.\tag{3.3.67}$$

Now, let $(Z^n)_{n\in\mathbb{N}}$ be a $\|\cdot\|_{\nu,p}$-Cauchy sequence. Fix $t\in J$. Then inequality (3.3.67) implies that the corresponding t-components $(Z_t^n)_{n\in\mathbb{N}}$ form a $\|\cdot\|_{L^p}$-Cauchy sequence in $L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{R})$. By completeness of $L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{R})$ (see [65, Theorem 3.11]) there exists $Z_t\in L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{R})$ such that $\|Z_t-Z_t^n\|_{L^p}\to 0$ as $n\to\infty$. For $t\in I\setminus J$ we choose $Z_t=0$. Therefore, we have constructed an adapted process $Z=(Z_t)_{t\in I}\in\prod_{t\in I}L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{R})$, which is an eligible representative of the corresponding equivalence class with respect to $\|\cdot\|_{\nu,p}$.

Next, we will show that the sequence $(Z^n)_{n\in\mathbb{N}}$ converges to Z with respect to $\|\cdot\|_{\nu,p}$ as $n\to\infty$. Since J is countable, there exists an increasing sequence $(J_k)_{k\in\mathbb{N}}$ of finite index sets with $\bigcup_{k\in\mathbb{N}}J_k=J$. By monotone convergence (cf. [75, Theorem 5.3]) we have that

$$\mathbb{E}\Big[\sum_{t\in J}|Z_t-Z_t^n|^p\gamma_t\Big]=\lim_{k\to\infty}\mathbb{E}\Big[\sum_{t\in J_k}|Z_t-Z_t^n|^p\gamma_t\Big],\quad n\in\mathbb{N},\gamma\in\mathcal{M}_I^\nu,$$

and thus

$$\|Z-Z^n\|^p_{\nu,p}=\sup_{\gamma\in\mathcal{M}_I^\nu}\lim_{k\to\infty}\mathbb{E}\Big[\sum_{t\in J_k}|Z_t-Z_t^n|^p\gamma_t\Big],\quad n\in\mathbb{N}.\tag{3.3.68}$$

Fix $\varepsilon>0$. Since $(Z^n)_{n\in\mathbb{N}}$ is a $\|\cdot\|_{\nu,p}$-Cauchy sequence, there exists an $N_\varepsilon\in\mathbb{N}$ such that $\|Z^m-Z^n\|_{\nu,p}\le\varepsilon$ for all $m,n\in\mathbb{N}$ with $m,n\ge N_\varepsilon$. Fix $k,n\in\mathbb{N}$ with $n\ge N_\varepsilon$. Since

$\|Z_t - Z_t^m\|_{L^p} \to 0$ as $m \to \infty$ for every t in the finite set J_k, we may iteratively find a subsequence $(m_l)_{l\in\mathbb{N}}$ with $m_l \geq N_\varepsilon$ for all $l \in \mathbb{N}$ such that $(Z_t^{m_l})_{l\in\mathbb{N}}$ converges to Z_t a.s. for every $t \in J_k$. Using Fatou's lemma ([75, Section 5.4]) for every $n \geq N_\varepsilon$ and that $J_k \subseteq J$ we get therefore that for every $\gamma \in \mathcal{M}_I^\nu$

$$\mathbb{E}\Big[\sum_{t\in J_k} |Z_t - Z_t^n|^p \gamma_t\Big] \leq \liminf_{l\to\infty} \mathbb{E}\Big[\sum_{t\in J_k} |Z_t^{m_l} - Z_t^n|^p \gamma_t\Big] \tag{3.3.69}$$

$$\leq \liminf_{l\to\infty} \|Z^{m_l} - Z^n\|_{\nu,p}^p \leq \varepsilon^p. \tag{3.3.70}$$

We know that $Z - Z^n \in \prod_{t\in I} L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R})$ and we conclude from (3.3.68) and (3.3.70) that $\|Z - Z^n\|_{\nu,p} \leq \varepsilon$ for all $n \geq N_\varepsilon$ such that $Z - Z^n \in X_{\nu,p}$. We also have $Z \in X_{\nu,p}$, because $Z = (Z - Z^n) + Z^n$, where $Z - Z^n$ and Z^n are both elements of the vector space $X_{\nu,p}$. ❑

Proof of Lemma 3.3.22. At first, given a process $Z \in \tilde{X}_{\nu,p}$ we show that the uniform Fatou property implies that

$$\|Z - Z^M\|_{\nu,p} \to 0 \quad \text{as } M \to \infty,$$

where Z^M is defined for a constant $M > 0$ as

$$Z_t^M := (-M) \vee (Z_t \wedge M) = \max\{-M, \min\{Z_t, M\}\}, \quad t \in I.$$

It is obvious that Z^M is bounded by M. Therefore we know that $\|Z^M\|_{\nu,p}$ is finite and $Z^M \in \tilde{X}_{\nu,p}$. In addition, we have that

$$Z_t = Z_t^M + (Z_t - M)_+ - (Z_t + M)_-,$$

where $x_+ = \max\{x, 0\}$ defines the positive part and $x_- = -\min\{x, 0\}$ the negative part. Let $\varepsilon > 0$. By the uniform Fatou property given in (3.3.20), there exists $M > 0$ such that $\sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}[\sum_{t\in I}(|Z_t|^p - M)_+\gamma_t] \leq \varepsilon$. Furthermore, every convex function $f : \mathbb{R}_+ \to \mathbb{R}$ with $f(0) \geq 0$ satisfies

$$f(a) + f(b) \leq f(0) + f(a+b). \tag{3.3.71}$$

This follows by adding the inequality

$$f(a) = f\Big(\frac{a}{a+b}(a+b) + \frac{b}{a+b}0\Big) \leq \frac{a}{a+b} f(a+b) + \frac{b}{a+b} f(0)$$

to the corresponding one where a and b are changed. For $p \geq 1$ the function $\mathbb{R}_+ \ni x \mapsto x^p$ is convex and from (3.3.71) we get

$$(|Z_t| - M)^p \leq |Z_t|^p - M^p \quad \text{on } \{|Z_t| > M\}$$

and thus

$$((|Z_t| - M)_+)^p \leq (|Z_t|^p - M^p)_+ \quad \text{on } \Omega. \tag{3.3.72}$$

Using that $|Z_t - Z_t^M| = (|Z_t| - M)\mathbb{1}_{\{|Z_t|>M\}} = (|Z_t| - M)_+$ it holds that for all $M > 0$

$$\|Z_t - Z_t^M\|_{\nu,p}^p = \sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}\Big[\sum_{t\in I}(|Z_t| - M)_+^p\gamma_t\Big] \overset{(3.3.72)}{\leq} \sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}\Big[\sum_{t\in I}(|Z_t|^p - M^p)_+\gamma_t\Big].$$

Hence, for $M \to \infty$ it follows that $\|Z_t - Z_t^M\|_{\nu,p} \to 0$ by the uniform Fatou property. Furthermore, for a M that is big enough we have that $Z - Z^M \in X_{\nu,p}$. By $Z = Z - Z^M + Z^M$ it follows that also $Z \in X_{\nu,p}$ such that $\tilde{X}_{\nu,p}$ is a subset of $X_{\nu,p}$.
Finally, we want to show that the reverse inclusion holds. For the convex function $x \mapsto x^p$, $p \geq 1$ on $\mathbb{R}_+$ we have by Jensen's inequality that for $a, b \in \mathbb{R}_+$

$$(a+b)^p = 2^p\Big(\frac{1}{2}a + \frac{1}{2}b\Big)^p \leq 2^{p-1}(a^p + b^p). \tag{3.3.73}$$

Therefore, we have for every constant $M > 0$

$$|Z_t|^p - 2^{p-1}M^p \leq 2^{p-1}(|Z_t| - M)^p \quad \text{on } \{|Z_t| > M\}$$

and

$$(|Z_t|^p - 2^{p-1}M^p)_+ \leq 2^{p-1}(|Z_t| - M)_+^p \quad \text{on } \Omega, \tag{3.3.74}$$

because $\{|Z_t|^p > 2^{p-1}M^p\} \subseteq \{|Z_t| > M\}$. Using (3.3.74) we get with $\tilde{M} = \frac{(2M)^{1/p}}{2}$ that

$$\begin{aligned}\sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}\Big[\sum_{t\in I}(|Z_t|^p - M)_+\gamma_t\Big] &\leq 2^{p-1}\sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}\Big[\sum_{t\in I}(|Z_t| - \tilde{M})_+^p\gamma_t\Big] \\ &= 2^{p-1}\|Z_t - Z_t^{\tilde{M}}\|_{\nu,p} \to 0 \quad \text{as } M \to \infty.\end{aligned}$$

Thus, $\|Z - Z^M\|_{\nu,p} \to 0$ for $M \to \infty$ implies the uniform Fatou property. ❑

Proof of Theorem 3.3.24.

(a) We know by the first part of the proof of Theorem 3.3.13 that $\|\cdot\|_{\nu,p}$ is a norm. By the descriptive definition of $\tilde{X}_{\nu,p}$ given in (3.3.23) it is clear that $\tilde{X}_{\nu,p}$ is a closed linear subspace of $X_{\nu,p}$, cf. the proof of Lemma 3.3.22. A closed subset of a Banach space is again complete.

(b) We know that $\|\cdot\|_{\nu,p}$ is a norm, see the first part of the proof of Theorem 3.3.13. Moreover, we have to show that $\hat{X}_{\nu,p}$ is a closed linear subspace of $X_{\nu,p}$. A closed subspace of a Banach space is again complete.
(i) $\hat{X}_{\nu,p}$ is a vector space: Note that the condition, that $\{|Z|_{\gamma,p}^p\}_{\gamma\in\mathcal{M}_I^\nu}$ is uniformly integrable, keeps the vector space properties. This is easily verifiable. Therefore $\hat{X}_{\nu,p}$ is a vector space.
(ii) $\hat{X}_{\nu,p}$ is subset of $X_{\nu,p}$: Note that the condition, that $\mathcal{Z} := \{|Z|_{\gamma,p}^p\}_{\gamma\in\mathcal{M}_I^\nu}$ is uniformly integrable, implies that $\|Z\|_{\nu,p} < \infty$. Let $\mathcal{Z}$ be the uniformly integrable set of $\mathbb{R}$-valued functions on $(\Omega, \mathcal{F})$. Then we have that $\mathcal{Z} \subseteq L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ and that $\mathcal{Z}$ is bounded in L^p, that means

$$\|Z\|_{\nu,p} = \sup_{\gamma\in\mathcal{M}_I^\nu} \||Z|_{\gamma,p}\|_{L^p} = \sup_{f\in\Psi}\|f\|_{L^p} < \infty.$$

We get this in the following way, in the case $p = 1$ it is given by the equivalence of (i) and (ii) in [70, Theorem 16.8]. Given $\varepsilon > 0$, there exists $n \in \mathbb{N}$ such that $\mathbb{E}[|Z|^p_{\gamma,p}\mathbb{1}_{\{|Z|^p_{\gamma,p}>n\}}] \leq \varepsilon$ for all $\gamma \in \mathcal{M}^\nu_I$, because of the uniform integrability of the considered set. Therefore

$$\begin{aligned}
\|Z\|_{\nu,p} &= \sup_{\gamma\in\mathcal{M}^\nu_I}\left(\mathbb{E}\Big[\sum_{t\in I}|Z_t|^p\gamma_t\Big]\right)^{1/p} \\
&= \sup_{\gamma\in\mathcal{M}^\nu_I}\left(\mathbb{E}\Big[|Z|^p_{\gamma,p}(\mathbb{1}_{\{|Z|^p_{\gamma,p}>n\}}+\mathbb{1}_{\{|Z|^p_{\gamma,p}\leq n\}})^p\Big]\right)^{1/p} \\
&\leq \sup_{\gamma\in\mathcal{M}^\nu_I}\left(\mathbb{E}\Big[|Z|^p_{\gamma,p}\mathbb{1}_{\{|Z|^p_{\gamma,p}>n\}}\Big]\right)^{1/p} + \sup_{\gamma\in\mathcal{M}^\nu_I}\left(\mathbb{E}\Big[n\mathbb{1}_{\{|Z|^p_{\gamma,p}\leq n\}}\Big]\right)^{1/p} \\
&\quad \text{(Minkowski inequality)} \\
&\leq \varepsilon^{1/p} + n^{1/p} < \infty.
\end{aligned}$$

Therefore we have that $\hat{X}_{\nu,p} \subseteq X_{\nu,p}$.
(iii) $\hat{X}_{\nu,p}$ is closed: Let $(Z^n)_{n\in\mathbb{N}}$ be a sequence in $\hat{X}_{\nu,p}$ which converges with respect to $\|\cdot\|_{\nu,p}$ to $Z \in X_{\nu,p}$. Thus, we have to show that $Z \in \hat{X}_{\nu,p}$, i.e., the set $\{|Z|^p_{\gamma,p}\}_{\gamma\in\mathcal{M}^\nu_I}$ is uniformly integrable. Let $\varepsilon > 0$. We know that there exists $N_\varepsilon \in \mathbb{N}$ with $\|Z - Z^n\|^p_{\nu,p} \leq \varepsilon$ for all $n \geq N_\varepsilon$, in particular for $n = N_\varepsilon$. Since $\{\sum_{t\in I}|Z^{N_\varepsilon}_t|^p\gamma_t\}_{\gamma\in\mathcal{M}^\nu_I}$ is uniformly integrable, there exists $\tilde{M} \in \mathbb{R}$ such that

$$\mathbb{E}\Big[|Z^{N_\varepsilon}|^p_{\gamma,p}\mathbb{1}_{\{|Z^{N_\varepsilon}|^p_{\gamma,p}>\tilde{M}\}}\Big] \leq \varepsilon \quad \text{for all } \gamma \in \mathcal{M}^\nu_I.$$

Since $\|Z\|_{\nu,p} < \infty$, there exits $M > 0$ such that $\tilde{M}\|Z\|^p_{\nu,p} \leq \varepsilon M$. Hence we get with $Z = (Z - Z^{N_\varepsilon}) + Z^{N_\varepsilon}$ and (3.3.73) that

$$\begin{aligned}
\mathbb{E}\Big[|Z|^p_{\gamma,p}\mathbb{1}_{\{|Z|^p_{\gamma,p}>M\}}\Big] &\leq 2^{p-1}\,\mathbb{E}\Big[|Z - Z^{N_\varepsilon}|^p_{\gamma,p}\Big] + 2^{p-1}\,\mathbb{E}\Big[|Z^{N_\varepsilon}|^p_{\gamma,p}\mathbb{1}_{\{|Z^{N_\varepsilon}|^p_{\gamma,p}>\tilde{M}\}}\Big] \\
&\quad + 2^{p-1}\,\mathbb{E}\Big[|Z^{N_\varepsilon}|^p_{\gamma,p}\mathbb{1}_{\{|Z|^p_{\gamma,p}>M,\,|Z^{N_\varepsilon}|^p_{\gamma,p}\leq\tilde{M}\}}\Big] \\
&\leq 2^{p-1}\|Z - Z^{N_\varepsilon}\|^p_{\nu,p} + 2^{p-1}\varepsilon + 2^{p-1}\tilde{M}\mathbb{P}(|Z|^p_{\gamma,p} > M) \\
&\leq 2^p\varepsilon + 2^{p-1}\tilde{M}\frac{\mathbb{E}[|Z|^p_{\gamma,p}]}{M} \leq 2^p\varepsilon + 2^{p-1}\tilde{M}\frac{\|Z\|^p_{\nu,p}}{M} \\
&\leq (2^p + 1)\varepsilon
\end{aligned}$$

for all $\gamma \in \mathcal{M}^\nu_I$ such that $Z \in \hat{X}_{\nu,p}$.

(c) We get in the parts (a) and (b) of the proof of Theorem 3.3.24 that $\tilde{X}_{\nu,p} \subseteq X_{\nu,p}$ and $\hat{X}_{\nu,p} \subseteq X_{\nu,p}$. Now, we have to show that $\tilde{X}_{\nu,p}$ is a subspace of $\hat{X}_{\nu,p}$. In other words, the uniform Fatou property implies the uniform integrability of $\{|Z|^p_{\gamma,p}\}_{\gamma\in\mathcal{M}^\nu_I}$. We have to show that for every $\varepsilon > 0$ there exists $M > 0$ such that

$$\mathbb{E}\Big[|Z|^p_{\gamma,p}\mathbb{1}_{\{|Z|^p_{\gamma,p}>M\}}\Big] \leq \varepsilon \quad \text{for all } \gamma \in \mathcal{M}^\nu_I.$$

Since Z satisfies $\limsup_{M\to\infty}\sup_{\gamma\in\mathcal{M}_I^\nu}\mathbb{E}\Big[\sum_{t\in I}(|Z_t|^p - M)_+\gamma_t\Big] = 0$, we get that $\|Z\|_{\nu,p} < \infty$ and there exists $\tilde{M} > 0$ such that

$$\mathbb{E}\Big[\sum_{t\in I}(|Z_t|^p - \tilde{M})_+\gamma_t\Big] \le \varepsilon \quad \text{for all } \gamma\in\mathcal{M}_I^\nu. \tag{3.3.75}$$

Since $\|Z\|_{\nu,p} < \infty$, there exists $M > 0$ such that $\tilde{M}\|Z\|_{\nu,p} \le \varepsilon M$.
Furthermore,

$$|Z|_{\gamma,p}^p - \tilde{M} = \sum_{t\in I}(|Z_t|^p - \tilde{M})\gamma_t \le \sum_{t\in I}(|Z_t|^p - \tilde{M})_+\gamma_t$$

and therefore

$$(|Z|_{\gamma,p}^p - \tilde{M})_+ \le \sum_{t\in I}(|Z_t|^p - \tilde{M})_+\gamma_t. \tag{3.3.76}$$

Using (3.3.75), (3.3.76) and the transformation of $|Z|_{\gamma,p}^p\mathbb{1}_{\{|Z|_{\gamma,p}^p>M\}}$ as

$$\begin{aligned}|Z|_{\gamma,p}^p\mathbb{1}_{\{|Z|_{\gamma,p}^p>M\}} &= (|Z|_{\gamma,p}^p - \tilde{M})\mathbb{1}_{\{|Z|_{\gamma,p}^p>M\}} + \tilde{M}\mathbb{1}_{\{|Z|_{\gamma,p}^p>M\}}\\ &\le (|Z|_{\gamma,p}^p - \tilde{M})_+ + \tilde{M}\frac{\|Z\|_{\nu,p}}{M},\end{aligned}$$

yield that

$$\begin{aligned}\mathbb{E}\Big[|Z|_{\gamma,p}^p\mathbb{1}_{\{|Z|_{\gamma}^p>M\}}\Big] &= \mathbb{E}\Big[(|Z|_{\gamma,p}^p - \tilde{M})_+\Big] + \mathbb{E}[\tilde{M}\mathbb{1}_{\{|Z|_{\gamma,p}^p>M\}}]\\ &= \mathbb{E}\Big[(|Z|_{\gamma,p}^p - \tilde{M})_+\Big] + \tilde{M}\mathbb{P}(|Z|_{\gamma,p}^p > M)\\ &\le \mathbb{E}\Big[(|Z|_{\gamma,p}^p - \tilde{M})_+\Big] + \tilde{M}\frac{\mathbb{E}[|Z|_{\gamma,p}^p]}{M}\\ &\le \mathbb{E}\Big[\sum_{t\in I}(|Z_t|^p - \tilde{M})_+\gamma_t\Big] + \tilde{M}\frac{\|Z\|_{\nu,p}}{M} \le 2\varepsilon.\end{aligned}$$

Thus, we get that $\{|Z|_{\gamma,p}^p\}_{\gamma\in\mathcal{M}_I^\nu}$ is uniformly integrable. ❑

Remark 3.3.77. It is unknown to us whether $\tilde{X}_{\nu,p}$ is really a proper subset of $\hat{X}_{\nu,p}$.

The next example shows us that in general $\hat{X}_{\nu,p}$ is really a proper subset of $X_{\nu,p}$. For that we specify a process $(Z_t)_{t\in I}\in X_{\nu,p}$ such that $\{\sum_{t\in I}|Z_t|^p\gamma_t\}_{\gamma\in\mathcal{M}_I^\nu}$ is not uniformly integrable. Therefore we choose $p = 1$.

Example 3.3.78. Let $I = \mathbb{N}$, $((0,1],\mathcal{B}_{(0,1]},\mathbb{F},\lambda)$ be the considered probability space with λ denoting the Lebesgue-Borel measure and the filtration $\mathbb{F}$ given as

$$\mathcal{F}_n := \sigma(\{((k-1)2^{-n}, k2^{-n}] \mid k\in\{1,\dots,2^n\}\})$$

for all $n \in \mathbb{N}$. The distributional restriction is given by $\nu_n = 2^{-n}$ for all $n \in \mathbb{N}$ and our underlying process $(Z_n)_{n\in\mathbb{N}}$ of random variables is defined by $Z_n = 2^n \mathbb{1}_{(0,2^{-n}]}$ for $n \in \mathbb{N}$. Note that the sequence $(Z_n)_{n\in\mathbb{N}}$ is bounded in $L^1(\lambda)$ but not uniformly integrable, because $\mathbb{E}[Z_n \mathbb{1}_{\{Z_n>c\}}] = \mathbb{E}[Z_n] = 1$ for all $c > 0$ and $n \in \mathbb{N}$ with $n > c$. It converges pointwise to zero, hence there can be no integrable random variable dominating the sequence to apply the dominated convergence theorem.
(a) $\|Z\|_{\nu,1} < \infty$: Using Theorem 3.2.8 we find an enlarged filtration and a random stopping time τ such that $\mathbb{E}[Z_\gamma] = \mathbb{E}[Z_\tau]$. Furthermore, we know that for the non-negative martingale Z we get $\mathbb{E}[Z_\tau] = \mathbb{E}[Z_1] = 1$. Therefore $\mathbb{E}[Z_\gamma]$ is bounded and consequently $\|Z\|_\nu < \infty$.

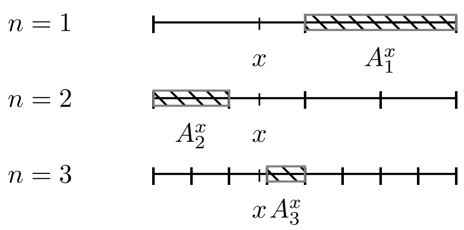


Figure 3.1.: Iteration steps of A_n^x for a given x

(b) $\{\sum_{t\in\mathbb{N}} |Z_t|\gamma_t\}_{\gamma\in\mathcal{M}_{\mathbb{N}}^\nu}$ is not uniformly integrable: For a fixed $x \in (0,1]$, we define γ^x by $\gamma_n^x := \mathbb{1}_{A_n^x}$ for all $n \in \mathbb{N}$, where

$$A_n^x := \left(\frac{\lceil 2^{n-1}x\rceil - 1}{2^{n-1}}, \frac{\lceil 2^{n-1}x\rceil}{2^{n-1}}\right] \Big\backslash \left(\frac{\lceil 2^n x\rceil - 1}{2^n}, \frac{\lceil 2^n x\rceil}{2^n}\right], \tag{3.3.79}$$

see Figure 3.1. Note that $x \notin A_n^x$, because $x \in \left(\frac{\lceil 2^n x\rceil - 1}{2^n}, \frac{\lceil 2^n x\rceil}{2^n}\right]$ and $\mathbb{E}[\gamma_n^x] = \lambda(A_n^x) = 2^{-n} = \nu_n$ for all $n \in \mathbb{N}$. Furthermore, we have that $|Z|_\gamma = Z_\gamma = \sum_{j\in\mathbb{N}} Z_j \gamma_j^x = Z_n$ if $A_n^x = (0, 2^{-n}]$. For $n = \min\{j \in \mathbb{N} \mid 2^j x > 1\}$ we get that $A_n^x = (0, 2^{-n}]$. Thus, $\{|Z|_\gamma\}_{\gamma\in\mathcal{M}_{\mathbb{N}}^\nu}$ is not uniformly integrable, because $\{Z_n\}_{n\in\mathbb{N}} \subseteq \{|Z|_\gamma\}_{\gamma\in\mathcal{M}_{\mathbb{N}}^\nu}$ and we know that $(Z_n)_{n\in\mathbb{N}}$ is not uniformly integrable.

Proof of Lemma 3.3.29. Using the definition of $B_{\nu,p}$ and the estimates for Lemma 3.3.25, it follows that $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is indeed a subset of $K_{\nu,p}$. Furthermore, by Lemma 3.3.25 we know that $\|\phi_\gamma\|_{X_{\nu,p}^*} \leq 1$.
It remains to show that $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is weak*-compact. As a closed subset of a compact space it is compact in the relative topology, see e.g. Proposition in [61, Chapter IV.3, p. 99], it is enough to show that $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is weak*-closed.
For this let $\psi \in X_{\nu,p}^* \subseteq \tilde{X}_{\nu,p}^*$ (or $\psi \in \tilde{X}_{\nu,1}^*$, respectively) be in the weak*-closure of $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$. We have to show that $\psi \in \{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ by proving that there exists a $\tilde{\gamma} \in \mathcal{M}_I^\nu$ such that $\psi = \phi_{\tilde{\gamma}}$ on $X_{\nu,p}$.

(a) **Existence of $\tilde{\gamma}$:** Fix $t \in I$ and define $\psi_t : L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}) \to \mathbb{R}$ by $\psi_t(Z_t) = \psi(Z^{(t)})$ for all $Z_t \in L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R})$, where $Z^{(t)} \in X_{\nu,p}$ is given by

$$Z_s^{(t)} = Z_t \mathbb{1}_{\{t\}}(s) \quad \forall s \in I. \tag{3.3.80}$$

Note that

$$\|Z^{(t)}\|_{\nu,p} = \sup_{\gamma \in \mathcal{M}_I^\nu} \left(\mathbb{E}\Big[\sum_{s \in I} |Z_t \mathbb{1}_{\{t\}}(s)|^p \gamma_s \Big] \right)^{1/p} = \sup_{\gamma \in \mathcal{M}_I^\nu} (\mathbb{E}[|Z_t|^p \gamma_t])^{1/p}$$
$$\leq (\mathbb{E}[|Z_t|^p])^{1/p} = \|Z_t\|_{L^p},$$

because $\gamma_t \leq 1$ a.s. The functional ψ_t is linear and

$$|\psi_t(Z_t)| = |\psi(Z^{(t)})| \leq \|\psi\|_{X^*_{\nu,p}} \|Z^{(t)}\|_{\nu,p} \leq \|\psi\|_{X^*_{\nu,p}} \|Z_t\|_{L^p},$$

hence $\psi_t \in (L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}))^* = L^q(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R})$. By [65, Theorem 6.16] there exists a $\tilde{\gamma}_t \in L^q(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R})$ with $\psi_t(Z_t) = \mathbb{E}[\tilde{\gamma}_t Z_t]$ for all $Z_t \in L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R})$. Hence we have $\tilde{\gamma} = (\tilde{\gamma}_t)_{t \in I} \in \prod_{t \in I} L^q(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R})$ such that $\phi_{\tilde{\gamma}}(Z^{(t)})$ is well-defined and $\psi(Z^{(t)}) = \phi_{\tilde{\gamma}}(Z^{(t)})$ for all $t \in I$.

(b) **The candidate $\tilde{\gamma}$ is in $\mathcal{M}_I^\nu$:**

(i) Note that $\tilde{\gamma}_t$ is $\mathcal{F}_t$-measurable for all $t \in I$ by construction.

(ii) Distribution of $\tilde{\gamma}$: For $t \in I$ define $Z_t = \mathbb{1}_\Omega$ and $Z^{(t)}$ via (3.3.80). Then

$$\phi_\gamma(Z^{(t)}) = \mathbb{E}[Z_t \gamma_t] = \mathbb{E}[\mathbb{1}_\Omega \gamma_t] = \mathbb{E}[\gamma_t] = \nu_t$$

for every $\gamma \in \mathcal{M}_I^\nu$, hence $\mathbb{E}[\tilde{\gamma}_t] = \psi(Z^{(t)}) = \nu_t$. For this, consider for $Z \in X_{\nu,p}$ the functional $\Theta_Z : X^*_{\nu,p} \to \mathbb{R}$ defined as $\Theta_Z(\phi) := \phi(Z)$. Note that Θ_Z is an element of the second dual $X^{**}_{\nu,p}$. Using that ψ is in the weak*-closure of $\{\phi_\gamma\}_{\gamma \in \mathcal{M}_I^\nu}$ and the continuity of every element of $X^{**}_{\nu,p}$, we get that for all $\varepsilon > 0$ and for for every $Z \in X_{\nu,p}$ there exists a $\gamma \in \mathcal{M}_I^\nu$ such that $|\Theta_Z(\psi) - \Theta_Z(\Phi_\gamma)| < \varepsilon$. Particularly, we get with $\Theta_{Z^{(t)}}(\phi_\gamma) = \phi_\gamma(Z^{(t)}) = \nu_t$ that $\psi(Z^{(t)}) = \Theta_{Z^{(t)}}(\psi) = \nu_t$.

(iii) $\tilde{\gamma}$ is non-negative: Fix $t \in I$. Define $A = \{\tilde{\gamma}_t < 0\}$ and note that $Z_t := \mathbb{1}_A \in L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R})$. $Z^{(t)}$ is defined as in (3.3.80). For every $\gamma \in \mathcal{M}_I^\nu$ we have $\phi_\gamma(Z^{(t)}) = \mathbb{E}[\gamma_t \mathbb{1}_A] \geq 0$, hence $0 \leq \psi(Z^{(t)}) = \phi_{\tilde{\gamma}}(Z^{(t)}) = \mathbb{E}[\tilde{\gamma}_t \mathbb{1}_A]$. This implies $\mathbb{P}(A) = 0$.

(iv) $\tilde{\gamma}$ is a probability measure satisfying conditions (b) and (c) of Definition 3.1.1: Consider an increasing sequence $(I_n)_{n \in \mathbb{N}}$ of finite index sets with $\bigcup_{n \in \mathbb{N}} I_n = I$. Define $Z^{\Omega,n} \in X_{\nu,p}$ by $Z_s^{\Omega,n} = \mathbb{1}_\Omega$ for all $s \in I_n$ and $Z_s^{\Omega,n} = 0$ for all $s \in I \setminus I_n$. Since for all $\gamma \in \mathcal{M}_I^\nu$

$$\phi_\gamma(Z^{\Omega,n}) = \mathbb{E}\Big[\sum_{t \in I_n} \gamma_t \Big] = \sum_{t \in I_n} \nu_t \leq 1,$$

we have $\psi(Z^{\Omega,n}) = \sum_{t \in I_n} \nu_t$, hence

$$\mathbb{E}\Big[\sum_{t \in I_n} \tilde{\gamma}_t \Big] = \phi_{\tilde{\gamma}}(Z^{\Omega,n}) = \psi(Z^{\Omega,n}) = \sum_{t \in I_n} \nu_t \nearrow \sum_{t \in I} \nu_t = 1 \text{ as } n \to \infty.$$

Further, by monotone convergence,

$$\mathbb{E}\Big[\sum_{t\in I_n}\tilde{\gamma}_t\Big] \nearrow \mathbb{E}\Big[\sum_{t\in I}\tilde{\gamma}_t\Big] \text{ as } n\to\infty.$$

Therefore $\mathbb{E}[\sum_{t\in I}\tilde{\gamma}_t]=1$.
Define $A=\{\sum_{t\in I}\tilde{\gamma}_t>1\}\in\sigma(\bigcup_{t\in I}\mathcal{F}_t)$. Furthermore, set $A_n=\{\sum_{t\in I_n}\tilde{\gamma}_t>1\}\in\sigma(\bigcup_{t\in I}\mathcal{F}_t)$. Then $A=\bigcup_{n\in\mathbb{N}}A_n$. For $n\in\mathbb{N}$ consider the processes Z^{A_n} with $Z_t^{A_n}=\mathbb{E}[\mathbb{1}_{A_n}|\mathcal{F}_t]$ for all $t\in I_n$ and $Z_t^{A_n}=0$ for all $t\in I\setminus I_n$. For every $\gamma\in\mathcal{M}_I^\nu$,

$$\phi_\gamma(Z^{A_n})=\mathbb{E}\Big[\sum_{t\in I_n}\mathbb{E}[\mathbb{1}_{A_n}|\mathcal{F}_t]\gamma_t\Big]=\sum_{t\in I_n}\mathbb{E}[\mathbb{1}_{A_n}\gamma_t]\le\mathbb{E}\Big[\mathbb{1}_{A_n}\sum_{t\in I}\gamma_t\Big]=\mathbb{P}(A_n).$$

Therefore, $\psi(Z^{A_n})\le\mathbb{P}(A_n)$. By the same calculation

$$\mathbb{E}\Big[\mathbb{1}_{A_n}\sum_{t\in I_n}\tilde{\gamma}_t\Big]=\phi_{\tilde{\gamma}}(Z^{A_n})=\psi(Z^{A_n})\le\mathbb{P}(A_n),$$

which implies that $\mathbb{P}(A_n)=0$. Since A is a countable union of null sets, we have also that $\mathbb{P}(A)=0$. Hence

$$\sum_{t\in I}\tilde{\gamma}_t=1,\ \text{a.s.}$$

Finally, $\tilde{\gamma}$ satisfies all properties of Definition 3.1.1, thus $\tilde{\gamma}\in\mathcal{M}_I^\nu$.

(c) **$\psi(Z)=\phi_{\tilde{\gamma}}(Z)$ for all $Z\in B_{\nu,p}$:** Let $Z\in X_{\nu,p}$ be fixed and $\varepsilon>0$ arbitrary. By Lemma 3.3.25 and $\tilde{\gamma}\in\mathcal{M}_I^\nu$ we have that $\phi_{\tilde{\gamma}}$ is well-defined. We will now make a case-by-case analysis. In fact, the case $p=1$ forms an exception and the corresponding proof is different.
Let $p\in(1,\infty)$. As $Z\in X_{\nu,p}$ we have that $\|Z\|_{\nu,p}<\infty$. Set $M>0$ such that $\|Z\|_{\nu,p}\le M$. Further there exists a finite set $\tilde{J}\subseteq I$ such that $(\sum_{t\in I\setminus\tilde{J}}\nu_t)^{1/q}\le\frac{\varepsilon}{M}$, where q is the Hölder conjugate of p. Define the process $Z^{\tilde{J}}\in X_{\nu,p}$ by

$$Z_t^{\tilde{J}}=\begin{cases}0 & \text{for } t\in\tilde{J},\\ Z_t & \text{for } t\in I\setminus\tilde{J}.\end{cases} \tag{3.3.81}$$

Using the Hölder's inequality like in (3.3.28), we get for every $\gamma\in\mathcal{M}_I^\nu$ that

$$\begin{aligned}|\phi_\gamma(Z^{\tilde{J}})|&\le\mathbb{E}\Big[\sum_{t\in I\setminus\tilde{J}}|Z_t|\gamma_t\Big]=\mathbb{E}\Big[\sum_{t\in I\setminus\tilde{J}}|Z_t|\gamma_t^{1/p}\cdot\gamma_t^{1/q}\Big]\\&\le\Big(\mathbb{E}\Big[\sum_{t\in I\setminus\tilde{J}}|Z_t|^p\gamma_t\Big]\Big)^{1/p}\Big(\mathbb{E}\Big[\sum_{t\in I\setminus\tilde{J}}(\gamma_t^{1/q})^q\Big]\Big)^{1/q}\\&\le\|Z\|_{\nu,p}\Big(\mathbb{E}\Big[\sum_{t\in I\setminus\tilde{J}}\gamma_t\Big]\Big)^{1/q}\le\|Z\|_{\nu,p}\Big(\sum_{t\in I\setminus\tilde{J}}\nu_t\Big)^{1/q}\le\varepsilon.\end{aligned}$$

Then also $|\psi(Z^{\tilde{J}})| \leq \varepsilon$. By the construction of $\tilde{\gamma}$ and linearity of ψ we have

$$\psi(Z - Z^{\tilde{J}}) = \phi_{\tilde{\gamma}}(Z - Z^{\tilde{J}}).$$

Finally

$$|\psi(Z) - \phi_{\tilde{\gamma}}(Z)| \leq |\psi(Z - Z^{\tilde{J}}) - \phi_{\tilde{\gamma}}(Z - Z^{\tilde{J}})| + |\psi(Z^{\tilde{J}})| + |\phi_{\tilde{\gamma}}(Z^{\tilde{J}})| \leq 2\varepsilon.$$

It follows that $\psi(Z) = \phi_{\tilde{\gamma}}(Z)$ for all $Z \in B_{\nu,p}$, $p \in (1,\infty)$.
Now, we consider the case $p = 1$. As $Z \in \tilde{X}_{\nu,p}$ we have that $\{\sum_{t\in I} |Z_t|^p \gamma_t\}_{\gamma\in\mathcal{M}_I^\nu}$ is uniformly integrable, because $\tilde{X}_{\nu,1}$ is a subset of $\hat{X}_{\nu,1}$.

(1) Let Z be additionally (uniformly) bounded. Then there exists $M > 0$ such that $|Z_t| \leq M$ a.s. for all $t \in I$. Further there exists a finite set $\tilde{J} \subseteq I$ such that $\sum_{t\in I\setminus\tilde{J}} \nu_t \leq \frac{\varepsilon}{M}$. Let the process $Z^{\tilde{J}} \in \tilde{X}_{\nu,p}$ be defined like in (3.3.81). Then for every $\gamma \in \mathcal{M}_I^\nu$

$$|\phi_\gamma(Z^{\tilde{J}})| \leq \mathbb{E}\Big[\sum_{t\in I\setminus\tilde{J}} |Z_t|\gamma_t\Big] \leq M\ \mathbb{E}\Big[\sum_{t\in I\setminus\tilde{J}} \gamma_t\Big] \leq \varepsilon.$$

Then also $|\psi(Z^{\tilde{J}})| \leq \varepsilon$. By the construction of $\tilde{\gamma}$ and linearity of ψ we have

$$\psi(Z - Z^{\tilde{J}}) = \phi_{\tilde{\gamma}}(Z - Z^{\tilde{J}}).$$

Finally

$$|\psi(Z) - \phi_{\tilde{\gamma}}(Z)| \leq |\psi(Z - Z^{\tilde{J}}) - \phi_{\tilde{\gamma}}(Z - Z^{\tilde{J}})| + |\psi(Z^{\tilde{J}})| + |\phi_{\tilde{\gamma}}(Z^{\tilde{J}})| \leq 2\varepsilon.$$

Thus, $\psi(Z) = \phi_{\tilde{\gamma}}(Z)$ for all uniformly bounded $Z \in X_{\nu,p}$.

(2) At first, recall that we define for the constant $M > 0$

$$Z_t^M := (-M) \vee (Z_t \wedge M) = \max\{-M, \min\{Z_t, M\}\}, \quad t \in I,$$

such that Z^M is bounded by M. Using that $Z_t = Z_t^M + (Z_t - M)_+ - (Z_t + M)_-$ and the linearity of ϕ_γ for all $\gamma \in \mathcal{M}_I^\nu$, we have that

$$\phi_\gamma(Z) = \phi_\gamma(Z^M) + \phi_\gamma(((Z_t - M)_+)_{t\in I}) - \phi_\gamma(((Z_t + M)_-)_{t\in I}).$$

For the bounded process Z^M we get with item (1) that $\phi_\gamma(Z^M) = \psi(Z^M)$ for all $\gamma \in \mathcal{M}_I^\nu$. Now, we have to consider $\phi_\gamma(((Z_t - M)_+)_{t\in I})$ and $\phi_\gamma(((Z_t + M)_-)_{t\in I})$. For this, let $\varepsilon > 0$. As $\limsup_{M\to\infty} \sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}[\sum_{t\in I}(|Z_t|^p - M)_+\gamma_t] = 0$, there exists $M > 0$ such that $\sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}[\sum_{t\in I}(|Z_t|^p - M)_+\gamma_t] \leq \varepsilon$. Using that $0 \leq (Z_t + M)_- \leq (|Z_t| - M)_+ \leq (|Z_t|^p - M)_+$ and $0 \leq (Z_t - M)_+ \leq (|Z_t|^p - M)_+$ we have besides that

$$\sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}\Big[\sum_{t\in I}(Z_t + M)_-\gamma_t\Big] \leq \varepsilon \quad \text{and} \tag{3.3.82}$$

$$\sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}\Big[\sum_{t\in I}(Z_t - M)_+\gamma_t\Big] \leq \varepsilon. \tag{3.3.83}$$

Furthermore, we get that $|\phi_\gamma((Z_t+M)_-)_{t\in I})| \le \varepsilon$ and $|\phi_\gamma((Z_t-M)_+)_{t\in I})| \le \varepsilon$ for all $\gamma \in \mathcal{M}_I^\nu$ because of (3.3.82) and (3.3.83). In particular, it applies to $\tilde{\gamma}$ such that we have $|\psi((Z_t+M)_-)_{t\in I})| \le \varepsilon$ and $|\psi((Z_t-M)_+)_{t\in I})| \le \varepsilon$. Finally we get

$$\begin{aligned}|\phi_{\tilde{\gamma}}(Z)-\psi(Z)| \le &\underbrace{|\phi_{\tilde{\gamma}}(Z^M)-\psi(Z^M)|}_{=0}+|\phi_{\tilde{\gamma}}((Z_t+M)_-)_{t\in I})|\\ &+|\psi((Z_t+M)_-)_{t\in I})|+|\phi_{\tilde{\gamma}}((Z_t-M)_+)_{t\in I})|+|\psi((Z_t-M)_+)_{t\in I}|\\ &\le 4\varepsilon.\end{aligned}$$

It follows that $\psi(Z)=\phi_{\tilde{\gamma}}(Z)$ for all $Z \in B_{\nu,1}$. ❑

Note that the uniform Fatou property is essential for the case $p=1$ and the last step of the proof of Lemma 3.3.29. It will be clear by the next example. Remember that we consider for $p=1$ the Banach space $(X_{\nu,1}, \|\cdot\|_{\nu,1})$, $B_{\nu,1} = \{Z \in \tilde{X}_{\nu,1} : \|Z\|_{\nu,1} \le 1\}$ and the polar set $K_{\nu,1} = \{\phi \in \tilde{X}^*_{\nu,1} : |\phi(Z)| \le 1 \text{ for all } Z \in B_{\nu,1}\}$. K is weak*-compact. The following example shows that $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is contained in K but not weak*-closed. For this we will choose a $\psi \in \tilde{X}^*_{\nu,1}$ in the weak*-closure of $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ for which there exists no $\tilde{\gamma} \in \mathcal{M}_I^\nu$ such that $\psi(Z)=\phi_{\tilde{\gamma}}(Z)$ for all $Z \in X_{\nu,1}$ with $\{|Z|_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is uniformly integrable, which implies the uniform Fatou property.

Example 3.3.84 (Extension of Example 3.3.78)**.** First of all, define $\tilde{\gamma}_n = \mathbb{1}_{(2^{-n},2^{-n+1}]}$ for all $n \in \mathbb{N}$. Then we have that $\sum_{n=1}^\infty \tilde{\gamma}_n = 1$ on $\Omega=(0,1]$ and for the given process $(Z_n)_{n\in\mathbb{N}}$ in Example 3.3.78 it holds that $\sum_{n=1}^\infty Z_n\tilde{\gamma}_n = 0$, because for $Z_n = 2^n\mathbb{1}_{(0,2^{-n}]}$ every summand is zero. Let $Y=(Y_n)_{n\in\mathbb{N}}$ be an $\mathbb{F}$-adapted process with $Y^* := \sup_{n\in\mathbb{N}}|Y_n| \in L^1$. Then $Y \in X_{\nu,1}$, because $\{|Y|_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is uniformly integrable. Remember the definition of γ^x for $x \in (0,1]$, which is given by $\gamma_n^x := \mathbb{1}_{A_n^x}$ for all $n \in \mathbb{N}$, where A_n^x is given as in (3.3.79). Futhermore we have for $n = \min\{j \in \mathbb{N} \mid 2^j x > 1\}$ that $A_n^x = (0, 2^{-n}]$. Thus, for every $m \in \mathbb{N}$ with $m \le n-1$ we get that $\gamma_m^{2^{-n}} = \tilde{\gamma}_m$.

$$\begin{aligned}\mathbb{E}\Big[\Big|\sum_{m=1}^\infty Y_m\tilde{\gamma}_m - \sum_{m=1}^\infty Y_m\gamma_m^{2^{-n}}\Big|\Big] &\le \mathbb{E}\Big[Y^*\sum_{m=1}^\infty |\tilde{\gamma}_m-\gamma_m^{2^{-n}}|\Big]\\ &= \mathbb{E}\Big[Y^*\sum_{m=n}^\infty |\tilde{\gamma}_m-\gamma_m^{2^{-n}}|\Big] \xrightarrow{n\to\infty} 0.\end{aligned}$$

The term converges to zero because of dominated convergence. Thus, we get that $\phi_{\gamma^{2^{-n}}} \to \phi_{\tilde{\gamma}}$ for $n\to\infty$. From the proof of Lemma 3.3.29 we know that $\phi_{\tilde{\gamma}}$ is the limit of sequence $(\phi_{\gamma^{2^{-n}}})_{n\in\mathbb{N}}$ for all $Y \in \tilde{X}_{\nu,1}$. Finally, we have an element $\psi \in \tilde{X}^*_{\nu,1}$ in the weak*-closure of $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ for which there exists $\tilde{\gamma} \in \mathcal{M}_I^\nu$ such that $\psi(Y)=\phi_{\tilde{\gamma}}(Y)$ for all $Y \in \tilde{X}_{\nu,1}$. Now, we want to show that this equation does not hold for an element of $X_{\nu,1}\setminus\tilde{X}_{\nu,1}$. We consider the sequence $Z:=(Z_m)_{m\in\mathbb{N}}$ defined in Example 3.3.78. Then Z is in $X_{\nu,1}\setminus\tilde{X}_{\nu,1}$ and for every $n \in \mathbb{N}$ we have that $\sum_{m=1}^\infty Z_m\gamma_m^{2^{-n}} = Z_n$ and $\mathbb{E}[Z_n]=1$. But $\sum_{m=1}^\infty Z_m\tilde{\gamma}_m = 0$ such that $\psi(Z) \ne \phi_{\tilde{\gamma}}(Z)$.

Proofs of Section 3.3.2

In this section the remaining proofs of Section 3.3.2 are to be delivered.

Proof of Theorem 3.3.47. As explained just after its definition in (3.3.32), $\|\cdot\|_{\nu,p,q}$ is a semi-norm on $\mathcal{X}_{\nu,p,q}$, which is a vector space. Therefore, it is a norm on the quotient space $X_{\nu,p,q}$.
To prove completeness, fix a $t \in I$ and a process $Z \in X_{\nu,p,q}$ first. Then by (3.3.32),

$$\|Z_t\|_{L^p(\gamma_t\mathbb{P})} \le \Big\| \|Z_.\|_{L^p(\gamma_.\mathbb{P})} \Big\|_{l^q(I)} \le \|Z\|_{\nu,p,q}, \quad \gamma \in \mathcal{M}_I^\nu. \tag{3.3.85}$$

For the deterministic choice $\gamma_t = \nu_t \mathbb{1}_\Omega$ from Example 3.1.4 we have that

$$\|Z_t\|_{L^p(\gamma_t\mathbb{P})} = c_{p,t}\|Z_t\|_{L^p(\mathbb{P})} \quad \text{with} \quad c_{p,t} = \begin{cases} \nu_t^{1/p} & \text{if } p \in [1,\infty), \\ \nu_t & \text{if } p = \infty. \end{cases} \tag{3.3.86}$$

Due to Remark 3.3.44 we restrict our considerations on the support $J = \operatorname{supp}(\nu)$ of ν. Combining (3.3.85) and (3.3.86),

$$\|Z_t\|_{L^p(\mathbb{P})} \le c_{p,t}^{-1}\|Z\|_{\nu,p,q}, \quad t \in J,\, Z \in X_{\nu,p,q}. \tag{3.3.87}$$

Now, let $(Z^n)_{n\in\mathbb{N}}$ be a $\|\cdot\|_{\nu,p,q}$-Cauchy sequence in $X_{\nu,p,q}$. Fix $t \in J$. Then inequality (3.3.87) implies that the corresponding t-components $(Z_t^n)_{n\in\mathbb{N}}$ form a $\|\cdot\|_{L^p(\mathbb{P})}$-Cauchy sequence in $L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$. By completeness of $L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$ (see [65, Theorem 3.11]) there exists $Z_t \in L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$ such that $\|Z_t - Z_t^n\|_{L^p(\mathbb{P})} \to 0$ as $n \to \infty$. Due to (3.3.46), it follows for each $\gamma \in \mathcal{M}_I^\nu$ that $\|Z_t - Z_t^n\|_{L^p(\gamma_t\mathbb{P})} \to 0$ as $n \to \infty$. For $t \in I \setminus J$ we choose $Z_t = 0$. Therefore, we have constructed an adapted process $Z = (Z_t)_{t\in I}$ in $\prod_{t\in I} L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$.
Next we will show that the sequence $(Z^n)_{n\in\mathbb{N}}$ converges to Z with respect to $\|\cdot\|_{\nu,p,q}$ and that $\|Z\|_{\nu,p,q} < \infty$. Since J is countable, there exists an increasing sequence $(J_k)_{k\in\mathbb{N}}$ of finite index sets with $\bigcup_{k\in\mathbb{N}} J_k = J$. By monotone convergence (cf. [75, Theorem 5.3]) for the counting measure on J in the case $q \in [1,\infty)$, and the property of the supremum of an increasing sequence in the case $q = \infty$, respectively, we have that

$$\Big\| \|Z_. - Z_.^n\|_{L^p(\gamma_.\mathbb{P})} \Big\|_{l^q(J)} = \lim_{k\to\infty} \Big\| \|Z_. - Z_.^n\|_{L^p(\gamma_.\mathbb{P})} \Big\|_{l^q(J_k)}, \quad n \in \mathbb{N},\, \gamma \in \mathcal{M}_I^\nu$$

and thus by (3.3.32) and Remark 3.3.44,

$$\|Z - Z^n\|_{\nu,p,q} = \sup_{\gamma\in\mathcal{M}_I^\nu} \lim_{k\to\infty} \Big\| \|Z_. - Z_.^n\|_{L^p(\gamma_.\mathbb{P})} \Big\|_{l^q(J_k)}, \quad n \in \mathbb{N}. \tag{3.3.88}$$

Fix $\varepsilon > 0$. Since $(Z^n)_{n\in\mathbb{N}}$ is a $\|\cdot\|_{\nu,p,q}$-Cauchy sequence, there exists an $N_\varepsilon \in \mathbb{N}$ such that $\|Z^m - Z^n\|_{\nu,p,q} \le \varepsilon$ for all $m,n \in \mathbb{N}$ with $m,n \ge N_\varepsilon$. Fix $k,n \in \mathbb{N}$ with $n \ge N_\varepsilon$. Since $\|Z_t - Z_t^m\|_{L^p(\mathbb{P})} \to 0$ as $m \to \infty$ for every t in the finite set J_k, we may iteratively find a subsequence $(m_l)_{l\in\mathbb{N}}$ with $m_l \ge N_\varepsilon$ for all $l \in \mathbb{N}$ such that $(Z_t^{m_l})_{l\in\mathbb{N}}$ converges to Z_t a.s. for every $t \in J_k$ (see [65, Theorem 3.12]). Using Fatou's lemma ([75, Section 5.4]) for every $n \ge N_\varepsilon$ and that $J_k \subseteq J$ we get therefore that, for every $\gamma \in \mathcal{M}_I^\nu$,

$$\begin{aligned} \Big\| \|Z_. - Z_.^n\|_{L^p(\gamma_.\mathbb{P})} \Big\|_{l^q(J_k)} &\le \liminf_{l\to\infty} \Big\| \|Z_.^{m_l} - Z_.^n\|_{L^p(\gamma_.\mathbb{P})} \Big\|_{l^q(J_k)} \\ &\le \liminf_{l\to\infty} \|Z^{m_l} - Z^n\|_{\nu,p,q} \le \varepsilon. \end{aligned} \tag{3.3.89}$$

We know that $Z-Z^n \in \prod_{t\in I} L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$ and we conclude from (3.3.88) and (3.3.89) that $\|Z-Z^n\|_{\nu,p,q} \le \varepsilon$ for all $n \ge N_\varepsilon$ such that $Z-Z^n \in X_{\nu,p,q}$. We also have $Z \in X_{\nu,p,q}$, because $Z=(Z-Z^n)+Z^n$ and $Z-Z^n$ and Z^n are both elements of the vector space $X_{\nu,p,q}$. ❑

Proof of Theorem 3.3.64. We know by the first part of the proof of Theorem 3.3.47 that $\|\cdot\|_{\nu,p,q}$ is a norm. By the descriptive definition of $\tilde{X}_{\nu,p,q}$ given in (3.3.63) it is clear that $\tilde{X}_{\nu,p,q}$ is a closed linear subspace of $X_{\nu,p,q}$. A closed subset of a Banach space is complete again. ❑

The next example shows that there are processes in $\mathcal{X}_{\nu,1,1}$ for which problem OptStop^{γ} has no solution although the supremum in (3.1.7) has a well-defined real value.

Example 3.3.90. Let $I=\mathbb{N}$, consider the filtered probability space $((0,1],\mathcal{B}_{(0,1]},\mathbb{F},\mathbb{P})$ with $\mathbb{P}$ denoting the Lebesgue–Borel measure and the filtration $\mathbb{F}=(\mathcal{F}_n)_{n\in\mathbb{N}}$ given by $\mathcal{F}_n=\sigma(\{((k-1)2^{-n},k2^{-n}]\mid k\in\{1,\dots,2^n\}\})$ for all $n\in\mathbb{N}$. The distributional restriction ν is given by $\nu_n=2^{-n}$ and the non-negative adapted process Z by $Z_n=2^n(1-\frac{1}{n})\mathbb{1}_{(0,2^{-n}]}$ for $n\in\mathbb{N}$. Given $\gamma\in\mathcal{M}_I^\nu$, the random variable γ_n, due to $\mathcal{F}_n$-measurability, has to be constant on $(0,2^{-n}]$, say α_n, for every $n\in\mathbb{N}$, and it follows from Definition 3.1.1(b) that $\sum_{n\in\mathbb{N}}\alpha_n\le 1$. Hence

$$\mathbb{E}[Z_\gamma]=\sum_{n\in\mathbb{N}}\mathbb{E}[\gamma_n Z_n]=\sum_{n\in\mathbb{N}}\alpha_n\,\mathbb{E}[Z_n]=\sum_{n\in\mathbb{N}}\alpha_n\Big(1-\frac{1}{n}\Big)<1$$

for every $\gamma\in\mathcal{M}_I^\nu$. By (3.3.32), $\|Z\|_{\nu,1,1}\le 1$. For every $n\in\mathbb{N}$ there exists a stopping time $\tau\in\mathcal{T}_I^\nu$ with $\{\tau=n\}=(0,2^{-n}]$ (and by Example 3.1.2 a corresponding $\gamma\in\mathcal{M}_I^\nu$), hence $\mathbb{E}[Z_\tau]\ge\mathbb{E}[Z_n]=(1-\frac{1}{n})$. Therefore, the supremum in (3.1.7) equals 1 but there is no $\gamma^*\in\mathcal{M}_I^\nu$ with $\mathbb{E}[Z_\gamma]=1$, hence OptStop^{γ} has no solution.

The next example shows us that in general $\tilde{X}_{\nu,p,q}$ is really a proper subset of $X_{\nu,p,q}$. We choose the case $p=q=1$. To show this, we specify a process $(Z_t)_{t\in I}\in X_{\nu,1,1}$ such that $\{\sum_{t\in I}|Z_t|\gamma_t\}_{\gamma\in\mathcal{M}_I^\nu}$ is not uniformly integrable and particularly not bounded.

Example 3.3.91. Let $I=\mathbb{N}$, consider the filtered probability space $((0,1],\mathcal{B}_{(0,1]},\mathbb{F},\mathbb{P})$ with $\mathbb{P}$ denoting the Lebesgue–Borel measure and the filtration $\mathbb{F}=(\mathcal{F}_n)_{n\in\mathbb{N}}$ given by $\mathcal{F}_n=\sigma(\{((k-1)2^{-n},k2^{-n}]\mid k\in\{1,\dots,2^n\}\})$ for all $n\in\mathbb{N}$. The distributional restriction ν is given by $\nu_n=2^{-n}$ for all $n\in\mathbb{N}$ and our underlying process $(Z_n)_{n\in\mathbb{N}}$ of random variables is defined by $Z_n=2^n\mathbb{1}_{(0,2^{-n}]}$ for $n\in\mathbb{N}$. Note that the sequence $(Z_n)_{n\in\mathbb{N}}$ is bounded in $L^1(\lambda)$ but not uniformly integrable, because $\mathbb{E}[Z_n\mathbb{1}_{\{Z_n>c\}}]=\mathbb{E}[Z_n]=1$ for all $c>0$ and $n\in\mathbb{N}$ with $n>c$. It converges pointwise to zero, hence there can be no integrable random variable dominating the sequence to apply the dominated convergence theorem.

(a) $\|Z\|_{\nu,1,1}<\infty$: Using Theorem 3.2.8 we find an enlarged filtration and a random stopping time τ such that $\mathbb{E}[Z_\gamma]=\mathbb{E}[Z_\tau]$. Furthermore we know that for the non-negative martingale Z we get $\mathbb{E}[Z_\tau]=\mathbb{E}[Z_1]=1$. Therefore $\mathbb{E}[Z_\gamma]$ is bounded and consequently $\|Z\|_{\nu,1,1}<\infty$.

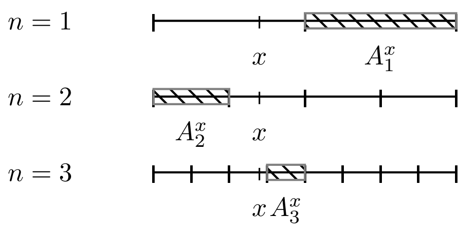


Figure 3.2.: Iteration steps of A_n^x for a given x

(b) $\{\sum_{t\in\mathbb{N}}|Z_t|\gamma_t\}_{\gamma\in\mathcal{M}_{\mathbb{N}}^{\nu}}$ is not uniformly integrable: For a fixed $x\in(0,1]$ we define γ^x by $\gamma_n^x := \mathbb{1}_{A_n^x}$ for all $n\in\mathbb{N}$, where

$$A_n^x := \left(\frac{\lceil 2^{n-1}x\rceil - 1}{2^{n-1}}, \frac{\lceil 2^{n-1}x\rceil}{2^{n-1}}\right] \Big\backslash \left(\frac{\lceil 2^{n}x\rceil - 1}{2^{n}}, \frac{\lceil 2^{n}x\rceil}{2^{n}}\right], \tag{3.3.92}$$

see Figure 3.2. Note that $x\notin A_n^x$, because $x\in\left(\frac{\lceil 2^n x\rceil-1}{2^n},\frac{\lceil 2^n x\rceil}{2^n}\right]$ and $\mathbb{E}[\gamma_n^x]=\lambda(A_n^x)=2^{-n}=\nu_n$ for all $n\in\mathbb{N}$. Furthermore we have that $|Z|_\gamma = Z_\gamma = \sum_{j\in\mathbb{N}} Z_j\gamma_j^x = Z_n$ if $A_n^x=(0,2^{-n}]$. For $n=\min\{j\in\mathbb{N}\mid 2^j x>1\}$ we get that $A_n^x=(0,2^{-n}]$. Thus $\{|Z|_\gamma\}_{\gamma\in\mathcal{M}_{\mathbb{N}}^{\nu}}$ is not uniformly integrable, because $\{Z_n\}_{n\in\mathbb{N}}\subseteq\{|Z|_\gamma\}_{\gamma\in\mathcal{M}_{\mathbb{N}}^{\nu}}$ and we know that $(Z_n)_{n\in\mathbb{N}}$ is not uniformly integrable.

Proof of Lemma 3.3.66. Using the definition of $V_{\nu,p,q}$ and the estimates in Remark 3.3.39, it follows that $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is indeed a subset of $K_{\nu,p,q}$.
It remains to show that $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is weak*-compact. As a closed subset of a compact space it is compact in the relative topology, see e.g. Proposition in [61, Chapter IV.3, p. 99], it is enough to show that $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ is weak*-closed.
For this let $\psi\in X_{\nu,p,q}^*\subseteq\tilde{X}_{\nu,p,q}^*$ (or $\psi\in\tilde{X}_{\nu,1,1}^*$, respectively) be in the weak*-closure of $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$. We have to show that $\psi\in\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ by proving that there exists a $\tilde{\gamma}\in\mathcal{M}_I^\nu$ such that $\psi=\phi_{\tilde{\gamma}}$ on $X_{\nu,p,q}$.
Step 1 (Existence of $\tilde{\gamma}$). Fix $t\in I$ and define $\psi_t\colon L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})\to\mathbb{K}$ by $\psi_t(Z_t)=\psi(Z^{(t)})$ for all $Z_t\in L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$, where $Z^{(t)}$ given by

$$Z_s^{(t)} = Z_t\mathbb{1}_{\{t\}}(s), \quad s\in I, \tag{3.3.93}$$

is always in the smaller space $\tilde{X}_{\nu,p,q}$. Note that by (3.3.32) and Definition 3.1.1(a) and (b),

$$\|Z^{(t)}\|_{\nu,p,q} = \sup_{\gamma\in\mathcal{M}_I^\nu}(\mathbb{E}[|Z_t|^p\gamma_t])^{1/p} \le (\mathbb{E}[|Z_t|^p])^{1/p} = \|Z_t\|_{L^p}.$$

The functional ψ_t is linear and

$$|\psi_t(Z_t)| = |\psi(Z^{(t)})| \le \|\psi\|_{\tilde{X}_{\nu,p,q}^*}\|Z^{(t)}\|_{\nu,p,q} \le \|\psi\|_{\tilde{X}_{\nu,p,q}^*}\|Z_t\|_{L^p},$$

hence $\psi_t\in(L^p(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K}))^*=L^{p'}(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$, where $p'\in(1,\infty]$ is the conjugate exponent. By [65, Theorem 6.16] there exists a $\tilde{\gamma}_t\in L^{p'}(\Omega,\mathcal{F}_t,\mathbb{P};\mathbb{K})$ with $\psi_t(Z_t)=\mathbb{E}[\tilde{\gamma}_t Z_t]$ for all

$Z_t \in L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{K})$.
Putting all $\tilde{\gamma}_t$ for $t \in I$ together, we have $\tilde{\gamma} := (\tilde{\gamma}_t)_{t\in I} \in \prod_{t\in I} L^{p'}(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{K})$ such that $\phi_{\tilde{\gamma}}(Z^{(t)}) = \mathbb{E}[Z_t \tilde{\gamma}_t]$ is well-defined and $\psi(Z^{(t)}) = \phi_{\tilde{\gamma}}(Z^{(t)})$ for every $t \in I$ and $Z_t \in L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{K})$. By linearity, it follows that $\psi(Z) = \phi_{\tilde{\gamma}}(Z) = \sum_{t\in I} \mathbb{E}[Z_t \tilde{\gamma}_t]$ for all $Z = (Z_t)_{t\in I} \in \prod_{t\in I} L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{K})$ with $Z_t \overset{\text{a.s.}}{=} 0$ for all but finitely many $t \in I$.

Step 2 (A version of the candidate $\tilde{\gamma}$ is in $\mathcal{M}_I^\nu$). Let $(\tilde{\gamma})_{t\in I} \in \prod_{t\in I} \mathcal{L}^{p'}(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{K})$ be any version from Step 1. We construct a version satisfying Definition 3.1.1

(a) *$\tilde{\gamma}$ is adapted:* By construction, $\tilde{\gamma}_t$ is $\mathcal{F}_t$-measurable for every $t \in I$.

(b) *Distribution of $\tilde{\gamma}$:* For $t \in I$ define $Z_t = \mathbb{1}_\Omega$ and $Z^{(t)}$ via (3.3.93). Then

$$\phi_\gamma(Z^{(t)}) = \mathbb{E}[Z_t \gamma_t] = \mathbb{E}[\gamma_t] = \nu_t$$

for every $\gamma \in \mathcal{M}_I^\nu$. We will prove next that $\mathbb{E}[\tilde{\gamma}_t] = \psi(Z^{(t)}) = \nu_t$. For this, consider for $Z \in X_{\nu,p,q}$ the functional $\Theta_Z \colon X^*_{\nu,p,q} \to \mathbb{K}$ defined as $\Theta_Z(\phi) = \phi(Z)$. Note that Θ_Z is an element of the second dual $X^{**}_{\nu,p,q}$. Using that ψ is in the weak*-closure of $\{\phi_\gamma\}_{\gamma\in\mathcal{M}_I^\nu}$ and the continuity of every element of $X^{**}_{\nu,p,q}$, we get that for every $\varepsilon > 0$ and for every $Z \in X_{\nu,p,q}$, there exists a $\gamma \in \mathcal{M}_I^\nu$ such that $|\Theta_Z(\psi) - \Theta_Z(\phi_\gamma)| < \varepsilon$. In particular, we get with $\Theta_{Z^{(t)}}(\phi_\gamma) = \phi_\gamma(Z^{(t)}) = \nu_t$ that $\psi(Z^{(t)}) = \Theta_{Z^{(t)}}(\psi) = \nu_t$.

(c) *$\tilde{\gamma}$ can be chosen real and non-negative:* Fix $t \in I$. Given $A \in \mathcal{F}_t$, note that $Z_t := \mathbb{1}_A$ is in $L^p(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{K})$. Let $Z^{(t)}$ by defined by (3.3.93). For every $\gamma \in \mathcal{M}_I^\nu$ we have $\phi_\gamma(Z^{(t)}) = \mathbb{E}[\gamma_t \mathbb{1}_A] \geq 0$, hence $\psi(Z^{(t)}) = \phi_{\tilde{\gamma}}(Z^{(t)}) = \mathbb{E}[\tilde{\gamma}_t \mathbb{1}_A]$ has to be real-valued and non-negative. This implies that the three $\mathcal{F}_t$-measurable sets $\{\mathfrak{Re}(\tilde{\gamma}_t) < 0\}$, $\{\mathfrak{Im}(\tilde{\gamma}_t) > 0\}$ and $\{\mathfrak{Im}(\tilde{\gamma}_t) < 0\}$ have $\mathbb{P}$-measure zero, where $\mathfrak{Re}(\cdot)$ denotes the real part and $\mathfrak{Im}(\cdot)$ the imaginary part. Hence, from now on we may take a non-negative $\tilde{\gamma}_t \in \mathcal{L}^{p'}(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R})$.

(d) *$\tilde{\gamma}$ satisfies Definition 3.1.1(c) and can be chosen to satisfy (b):* Consider an increasing sequence $(I_n)_{n\in\mathbb{N}}$ of finite index sets with $\bigcup_{n\in\mathbb{N}} I_n = I$. Define $Z^{\Omega,n} \in X_{\nu,p,q}$ by $Z_t^{\Omega,n} = \mathbb{1}_\Omega$ for all $t \in I_n$ and $Z_t^{\Omega,n} = 0$ for all $t \in I \setminus I_n$. Since

$$\phi_\gamma(Z^{\Omega,n}) = \mathbb{E}\Big[\sum_{t\in I_n} \gamma_t\Big] = \sum_{t\in I_n} \nu_t \leq 1, \quad \gamma \in \mathcal{M}_I^\nu,$$

by Definition 3.1.1(e), we have $\psi(Z^{\Omega,n}) = \sum_{t\in I_n} \nu_t$, hence

$$\mathbb{E}\Big[\sum_{t\in I_n} \tilde{\gamma}_t\Big] = \phi_{\tilde{\gamma}}(Z^{\Omega,n}) = \psi(Z^{\Omega,n}) = \sum_{t\in I_n} \nu_t \overset{n\to\infty}{\longrightarrow} \sum_{t\in I} \nu_t = 1.$$

Further, by monotone convergence,

$$\mathbb{E}\Big[\sum_{t\in I_n} \tilde{\gamma}_t\Big] \overset{n\to\infty}{\longrightarrow} \mathbb{E}\Big[\sum_{t\in I} \tilde{\gamma}_t\Big].$$

Therefore $\mathbb{E}[\sum_{t\in I} \tilde{\gamma}_t] = 1$.

For $n \in \mathbb{N}$ define $A_n = \{\sum_{t\in I_n} \tilde{\gamma}_t > 1\}$ and the processes Z^{A_n} by

$$Z_t^{A_n} = \begin{cases} \mathbb{E}\big[\mathbb{1}_{A_n} | \mathcal{F}_t\big] & \text{for } t \in I_n, \\ 0 & \text{for } t \in I \setminus I_n. \end{cases}$$

For every $\gamma \in \mathcal{M}_I^\nu$, using Definition 3.1.1(b),

$$\phi_\gamma(Z^{A_n}) = \mathbb{E}\Big[\sum_{t\in I_n} \mathbb{E}[\mathbb{1}_{A_n}|\mathcal{F}_t]\gamma_t\Big] = \sum_{t\in I_n} \mathbb{E}\big[\mathbb{1}_{A_n}\gamma_t\big] \le \mathbb{E}\Big[\mathbb{1}_{A_n}\sum_{t\in I}\gamma_t\Big] \le \mathbb{P}(A_n).$$

Therefore, $\psi(Z^{A_n}) \le \mathbb{P}(A_n)$. By the same calculation

$$\mathbb{E}\Big[\mathbb{1}_{A_n}\sum_{t\in I_n}\tilde{\gamma}_t\Big] = \phi_{\tilde{\gamma}}(Z^{A_n}) = \psi(Z^{A_n}) \le \mathbb{P}(A_n),$$

which implies that $\mathbb{P}(A_n) = 0$. Hence $A := \{\sum_{t\in I}\tilde{\gamma}_t > 1\} = \bigcup_{n\in\mathbb{N}} A_n$ is also a $\mathbb{P}$-null set, which implies that $\sum_{t\in I}\tilde{\gamma}_t = 1$ a.s., hence $\tilde{\gamma}$ satisfies Definition 3.1.1(c). Using Notation 2.0.1(d) we replace $\tilde{\gamma}_t$ by $\min\{\tilde{\gamma}_t, (1-\tilde{\gamma}_{<t})^+\}$ for each $t \in I$, which is a change only on the $\mathbb{P}$-null set A.[1] Then $(\tilde{\gamma}_t)_{t\in I}$ also satisfies Definition 3.1.1(b).

Step 3 ($\psi(Z) = \phi_{\tilde{\gamma}}(Z)$ for all $Z \in X_{\nu,p,q}$ or $Z \in \tilde{X}_{\nu,1,1}$, respectively). Fix $Z \in X_{\nu,p,q}$. Since $\tilde{\gamma} \in \mathcal{M}_I^\nu$ by Step 2, it follows from Remark 3.3.39 for the exponents p,q considered in Theorem 3.3.34 that $Z_{\tilde{\gamma}}$ is a well-defined element of $L^1(\mathbb{P})$, hence $\phi_{\tilde{\gamma}}(Z) = \mathbb{E}[Z_{\tilde{\gamma}}]$ is a well-defined extension of $\phi_{\tilde{\gamma}}$ given at the end of Step 1. Note that this extension is linear.

(a) First, we consider the case $p \in (1,\infty)$. Take any $\varepsilon > 0$. If $q = 1$, then there exists a finite set $I_\varepsilon \subseteq I$ such that

$$\|Z\|_{\nu,p,1} \cdot \sup_{t\in I\setminus I_\varepsilon} \nu_t^{1/p'} \le \varepsilon. \tag{3.3.94}$$

If $q \in (1,\infty]$, then due to (3.3.35) there exists a finite set $I_\varepsilon \subseteq I$ such that

$$\|Z\|_{\nu,p,q} \cdot \Big(\sum_{t\in I\setminus I_\varepsilon} \nu_t^{q'/p'}\Big)^{1/q'} \le \varepsilon, \tag{3.3.95}$$

where $p' \in (1,\infty)$ and $q' \in [1,\infty)$ are the conjugate exponents of p and q, respectively. In both cases define the process $Z^{I_\varepsilon} \in X_{\nu,p,q}$ by

$$Z_t^{I_\varepsilon} = \begin{cases} 0 & \text{for } t \in I_\varepsilon, \\ Z_t & \text{for } t \in I \setminus I_\varepsilon. \end{cases} \tag{3.3.96}$$

Using the estimates (3.3.40), (3.3.41) or (3.3.42), respectively, for $I' := I \setminus I_\varepsilon$, it follows in combination with (3.3.94) or (3.3.95), respectively, that $|\phi_\gamma(Z^{I_\varepsilon})| \le \varepsilon$ for every $\gamma \in \mathcal{M}_I^\nu$, hence also $|\psi(Z^{I_\varepsilon})| \le \varepsilon$. By the construction of $\phi_{\tilde{\gamma}}$ at the end of Step 1, $\psi(Z - Z^{I_\varepsilon}) = \phi_{\tilde{\gamma}}(Z - Z^{I_\varepsilon})$. By the linearity of ψ and $\phi_{\tilde{\gamma}}$ on $X_{\nu,p,q}$ as well as the triangle inequality,

$$|\psi(Z) - \phi_{\tilde{\gamma}}(Z)| \le |\psi(Z - Z^{I_\varepsilon}) - \phi_{\tilde{\gamma}}(Z - Z^{I_\varepsilon})| + |\psi(Z^{I_\varepsilon})| + |\phi_{\tilde{\gamma}}(Z^{I_\varepsilon})| \le 2\varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $\psi(Z) = \phi_{\tilde{\gamma}}(Z)$.

[1] If $I = \mathbb{Q}$ and $\tilde{\gamma}_t(\omega) = 1$ for all $t \in I$, then this construction produces the zero path for ω, which is (besides Example 3.2.4) another reason to have the exceptional null set in Definition 3.1.1(c).

(b) Now, we consider the case $p = q = 1$. First we consider a $Z \in \tilde{X}_{\nu,1,1}$ such that there exists a real number $M > 0$ satisfying $|Z_t| \leq M$ for all $t \in I$ (i.e., Z is bounded). Take any $\varepsilon > 0$. There exists a finite set $I_\varepsilon \subseteq I$ such that $\sum_{t \in I \setminus I_\varepsilon} \nu_t \leq \varepsilon / M$. Let the process Z^{I_ε} be defined by (3.3.96). Then, using Definition 3.1.1(e),

$$|\phi_\gamma(Z^{I_\varepsilon})| \leq \mathbb{E}\Big[\sum_{t \in I \setminus I_\varepsilon} |Z_t| \gamma_t\Big] \leq M\, \mathbb{E}\Big[\sum_{t \in I \setminus I_\varepsilon} \gamma_t\Big] \leq \varepsilon, \quad \gamma \in \mathcal{M}_I^\nu,$$

hence also $|\psi(Z^{I_\varepsilon})| \leq \varepsilon$. As in the previous case, it follows that $\psi(Z) = \phi_{\tilde{\gamma}}(Z)$.
For a general $Z \in \tilde{X}_{\nu,1,1}$ there exists by definition a sequence $(Z^n)_{n \in \mathbb{N}}$ of bounded processes in $\tilde{X}_{\nu,1,1}$ with $\|Z - Z^n\|_{\nu,1,1} \to 0$ as $n \to \infty$. Then, using linearity and the above result,

$$\begin{aligned} |\psi(Z) - \phi_{\tilde{\gamma}}(Z)| &\leq |\psi(Z - Z^n)| + |\psi(Z^n) - \phi_{\tilde{\gamma}}(Z^n)| + |\phi_{\tilde{\gamma}}(Z^n - Z)| \\ &\leq \Big(\|\psi\|_{\tilde{X}^*_{\nu,p,q}} + \|\phi_{\tilde{\gamma}}\|_{\tilde{X}^*_{\nu,p,q}}\Big)\|Z - Z^n\|_{\nu,1,1} \to 0 \end{aligned}$$

as $n \to \infty$, hence $\psi(Z) = \phi_{\tilde{\gamma}}(Z)$. ❑

Note that in the case $p = q = 1$ special consideration is required what the Example 3.3.84 from the previous section shows. Remember that we consider there for $p = 1$ the Banach space $(\tilde{X}_{\nu,1}, \|\cdot\|_{\nu,1})$, $B_{\nu,1} = \{Z \in X_{\nu,1} : \|Z\|_{\nu,1} \leq 1\}$ and the polar set $K = \{\phi \in \tilde{X}^*_{\nu,1} : |\phi(Z)| \leq 1 \text{ for all } Z \in B_{\nu,1}\}$. K is weak*-compact. The example shows that $\{\phi_\gamma\}_{\gamma \in \mathcal{M}_I^\nu}$ is contained in K but not weak*-closed. For this we choose a $\psi \in X^*_{\nu,1}$ in the weak*-closure of $\{\phi_\gamma\}_{\gamma \in \mathcal{M}_I^\nu}$ for which there exists no $\tilde{\gamma} \in \mathcal{M}_I^\nu$ such that $\psi(Z) = \phi_{\tilde{\gamma}}(Z)$ for all $Z \in \tilde{X}_{\nu,1}$ with $\{|Z|_\gamma\}_{\gamma \in \mathcal{M}_I^\nu}$ is uniformly integrable, which implies the uniform Fatou property.

3.4. General Results and Bounds

Some general bounds have already been specified in [33, Section 4.2] and can be easily transferred. We can find a lower bound by assuming that the process Z and the stopping time or adapted random probability measure are independent. If we assume that the process Z and the stopping time τ are independent, we have that

$$V^{\nu}_{\text{ind}}(Z) = \sum_{t\in I} \mathbb{E}[Z_t \mathbb{1}_{\{\tau=t\}}] = \sum_{t\in I} \mathbb{E}[Z_t]\nu_t.$$

We get the same result, if Z and the adapted random probabilty measure are independent, see Lemma 3.5.1 below, such that it holds:

$$V^{\nu}_{\text{ind}}(Z) \leq V^{\nu}_{\mathcal{T}}(Z) \leq V^{\nu}_{\mathcal{M}}(Z),$$

with $V^{\nu}_{\mathcal{T}}(Z)$ as in (2.0.5) and $V^{\nu}_{\mathcal{M}}(Z)$ as in (3.1.10). We have seen in Section 3.2 that there are examples for $V^{\nu}_{\mathcal{T}}(Z) < V^{\nu}_{\mathcal{M}}(Z)$, cf. Example 3.2.3. Furthermore by (2.0.5) and (3.2.1) we have $V^{\nu}_{\mathcal{T}}(Z) \leq V_{\mathcal{T}}(Z)$ and $V^{\nu}_{\mathcal{M}}(Z) \leq V_{\mathcal{M}}(Z)$, because $\mathcal{T}^{\nu}_I \subseteq \mathcal{T}_I$ and $\mathcal{M}^{\nu}_I \subseteq \mathcal{M}_I$. However, it is very questionable if $V^{\nu}_{\mathcal{M}}(Z) \leq V_{\mathcal{T}}(Z)$, as in [33] discussed.
Furthermore we modify [33, Theorem 2.49] in the following way:

Theorem 3.4.1.

(a) Let $I \subseteq \mathbb{N}_0$ with $0 \in I$. Let Z be a uniformly integrable supermartingale. Then, for every $\gamma \in \mathcal{M}_I$, the random variable Z_γ is well-defined, integrable and satisfies $\mathbb{E}[Z_\gamma] \leq \mathbb{E}[Z_0]$. If, further, Z is a martingale, then, for every $\gamma \in \mathcal{M}_I$, $\mathbb{E}[Z_\gamma] = \mathbb{E}[Z_0]$.

(b) Let be I a totally ordered countable set. Let Z be a closable martingale. Then, for every $\gamma \in \mathcal{M}_I$, the random variable Z_γ is well-defined, integrable and satisfies $\mathbb{E}[Z_\gamma] = \mathbb{E}[Z_t]$ for all $t \in I$.

(c) Let $I \subseteq \mathbb{N}_0$ with $0 \in I$, ν be a probability distribution on I. Let Z be a supermartingale. Then, for every $\gamma \in \mathcal{M}^{\nu}_I$, the random variable Z_γ is well-defined, integrable and satisfies $\mathbb{E}[Z_\gamma] \leq \mathbb{E}[Z_0]$ in each of the following situations:

1. *There exists a $t \in I$ with $\nu_{\leq t} = 1$,*
2. *Z is bounded a.s.,*
3. *$\sum_{t\in I} |t|\nu_t < \infty$ and for some $K > 0$ we have $|Z_t(\omega) - Z_s(\omega)| \leq K|t-s|$ a.s. for all $s, t \in I \setminus \{0\}$ and for all $\omega \in \Omega$.*

(d) If $I \subseteq \mathbb{N}_0$ with $0 \in I$, any of the conditions (c1), (c2) or (c3) holds and Z is a martingale, then, for every $\gamma \in \mathcal{M}^{\nu}_I$, Z_γ is well-defined, integrable and $\mathbb{E}[Z_\gamma] = \mathbb{E}[Z_0]$.

(e) If $I \subseteq \mathbb{Z}$ is a countably infinite index set, ν is a probability distribution on I, Z is a martingale and there exists $t \in I$ with $\nu_{\leq t} = 1$, then Z_γ is integrable and $\mathbb{E}[Z_\gamma] = \mathbb{E}[Z_t]$ for all $t \in I$.

(f) Let I be a totally ordered countable set, ν be a probability distribution on I. Let Z be a martingale with $Z^* = \sup_{t\in I} |Z_t| \in L^1$. Then, for every $\gamma \in \mathcal{M}^{\nu}_I$, the random variable is well-defined, integrable and satisfies $\mathbb{E}[Z_\gamma] = \mathbb{E}[Z_t]$ for every $t \in I$.

By [33, Remark 2.50] it should hold that $\sum_{t\in I} t\nu_t = \sum_{t\in I}(1-\nu_{\leq t})$, when we use Condition (c3) of [33, Theorem 2.49], i.e., $I \subseteq \mathbb{N}_0$ with $0 \in I$, ν is a probability distribution on I and $\sum_{t\in I} |t|\nu_t < \infty$, which is similar to Condition (c3) of Theorem 3.4.1. The next example is a counterexample for that. The corrected version is given in Remark 3.4.3.

Example 3.4.2. Counterexample for [33, Remark 2.50]:
Let $I = \{0,2,4\}$ and choose $\nu_t = \frac{1}{3}$ for all $t \in I$. We get for the right side of the equation

$$\begin{aligned}\sum_{t\in I}(1-\nu_{\leq t}) &= (1-\nu_{\leq 0}) + (1-\nu_{\leq 2}) + (1-\nu_{\leq 4})\\ &= 3 - \nu_0 - \nu_0 - \nu_2 - \nu_0 - \nu_2 - \nu_4\\ &= 3 - 3\nu_0 - 2\nu_2 - \nu_4 = 3 - 6\cdot\frac{1}{3} = 1\end{aligned}$$

and for the left side

$$\begin{aligned}\sum_{t\in I} t\nu_t &= 0\cdot\nu_0 + 2\cdot\nu_2 + 4\cdot\nu_4\\ &= 2\cdot\nu_2 + \nu_4 + 3\cdot\nu_4 = 2\cdot\nu_2 + \nu_4 + 3\cdot(1-\nu_0-\nu_2)\\ &= 3 - 3\nu_0 - \nu_2 + \nu_4 = 3 - 3\cdot\frac{1}{3} = 2.\end{aligned}$$

Remark 3.4.3. Given $I \subseteq \mathbb{N}_0$ with $0 \in I$, let ν be a probability distribution on I with $\sum_{t\in I} t\nu_t < \infty$. Let $|I|$ be the number of elements of the index set I. Then

$$\sum_{t\in I} |I_{<t}|\nu_t = \sum_{t\in I}(1-\nu_{\leq t})$$

or rather

$$\sum_{t\in I} |I_{\leq t}|\nu_t = \sum_{t\in I}(1-\nu_{<t}).$$

This is obtained as follows

$$\sum_{t\in I}(1-\nu_{\leq t}) = \sum_{t\in I}\sum_{\substack{s\in I\\ s>t}} \nu_s = \sum_{t\in I}(|I| - \sum_{\substack{s\in I\\ s\geq t}} 1)\nu_t = \sum_{t\in I}|I_{<t}|\nu_t \quad \text{and}$$

$$\begin{aligned}\sum_{t\in I}|I_{\leq t}|\nu_t &= \sum_{t\in I}(|I_{<t}|+1)\nu_t = \sum_{t\in I}|I_{<t}|\nu_t + \sum_{t\in I}\nu_t \text{ (rearrangement of the first sum)}\\ &= \sum_{t\in I}(1-\nu_{\leq t}) + \sum_{t\in I}\nu_t = \sum_{t\in I}(1-\nu_{<t}-\nu_t) + \sum_{t\in I}\nu_t\\ &= \sum_{t\in I}(1-\nu_{<t}).\end{aligned}$$

Example 3.4.4. We consider again Example 3.4.2, where $I = \{0,2,4\}$ and $\nu_t = \frac{1}{3}$ for all $t \in I$. It holds that

$$\sum_{t\in I}(1-\nu_{\leq t}) = 3 - 3\nu_0 - 2\nu_2 - \nu_4 = 3 - 6\cdot\frac{1}{3} = 1$$

and

$$\sum_{t\in I} |I_{<t}|\nu_t = 0\cdot\nu_0 + 1\cdot\nu_2 + 2\cdot\nu_4 (= 3\cdot\frac{1}{3} = 1)$$
$$= 1\cdot\nu_2 + 3\cdot\nu_4 - \nu_4 = 1\cdot\nu_2 + 3\cdot(1-\nu_0-\nu_2) - \nu_4$$
$$= 3 - 3\nu_0 - 2\nu_2 - \nu_4.$$

Proof of Theorem 3.4.1. At first, we prove (b): Using monotone convergence and Jensen's inequality, we get

$$\mathbb{E}[|Z_\gamma|] = \mathbb{E}\Big[\Big|\sum_{t\in I} Z_t\gamma_t\Big|\Big] \le \mathbb{E}\Big[\sum_{t\in I}|Z_t|\underbrace{|\gamma_t|}_{\ge 0 \text{ a.s.}}\Big] \overset{Z \text{ closable}}{=} \mathbb{E}\Big[\sum_{t\in I}\underbrace{|\mathbb{E}[Z_\infty|\mathcal{F}_t]|}_{\le \mathbb{E}[|Z_\infty||\mathcal{F}_t] \text{ (Jensen)}}\gamma_t\Big]$$
$$\overset{\gamma_t\, \mathcal{F}_t\text{-measurable}}{\le} \sum_{t\in I}\mathbb{E}[\mathbb{E}[|Z_\infty|\gamma_t|\mathcal{F}_t]] = \sum_{t\in I}\mathbb{E}[|Z_\infty|\gamma_t] = \mathbb{E}\Big[|Z_\infty|\underbrace{\sum_{t\in I}\gamma_t}_{=1}\Big] = \mathbb{E}[|Z_\infty|] < \infty.$$

This implies that Z_γ is well-defined and integrable. Repeating the calculation without absolute values, which is allowed due to the almost surely absolute convergence of Z_γ, we get

$$\mathbb{E}[Z_\gamma] = \mathbb{E}\Big[\sum_{t\in I} Z_t\gamma_t\Big] = \sum_{t\in I}\mathbb{E}[\mathbb{E}[Z_\infty\gamma_t|\mathcal{F}_t]] = \sum_{t\in I}\mathbb{E}[Z_\infty\gamma_t] = \mathbb{E}\Big[Z_\infty\sum_{t\in I}\gamma_t\Big]$$
$$= \mathbb{E}[Z_\infty] = \mathbb{E}[Z_t] \quad \forall t\in I,$$

because Z is a closable martingale.
The rest of the results of this theorem follows by Theorem 2.0.6 and Theorem 3.2.8. The conditions imposed on Z and ν in this theorem are equivalent to the conditions on the process and the stopping time in Theorem 2.0.6, where the stopping time is now found using Theorem 3.2.8. ❑

As in [33] discussed, we can find an upper bound by the value of an optimal stopping problem with the same underlying process. We get that $V_T^\nu(Z) \le \mathbb{E}[Z_T]$ by [33, Lemma 4.37], if Z is a submartingale. At first we will take a closer look at this inequality given by [33, Lemma 4.7] in the following way:

Lemma 3.4.5. *Assume $I = \mathbb{N}_0$ and that the process Z is a submartingale.*

(a) If $\mathbb{E}[\sup_{t\in I}|Z_t|] < \infty$, then for every $\gamma \in \mathcal{M}_I$

$$\mathbb{E}[Z_0] \le \mathbb{E}[Z_\gamma] \le \mathbb{E}[Z_\infty].$$

(b) Using the Doob decomposition $M + A$ of the submartingale Z and assuming that M (or in case $\gamma \in \mathcal{M}_I^\nu$ we consider M and ν) satisfies one of the conditions of [33, Theorem 2.49], we find

$$\mathbb{E}[M_0] \le \mathbb{E}[Z_\gamma] \le \mathbb{E}[M_0] + \mathbb{E}[A_\infty].$$

Remark 3.4.6. The results are shown for $I = \mathbb{N}_0$. If $I = \{0, \dots, T\}$ we simply have to replace Z_∞ and A_∞ by Z_T and A_T, respectively.

If Z is a supermartingale, we have

$$\begin{aligned}\mathbb{E}[Z_T] &= \mathbb{E}[Z_T]\cdot 1 = \mathbb{E}[Z_T]\sum_{t\in I}\nu_t = \sum_{t\in I}\mathbb{E}[Z_T]\nu_t \\ &\le \sum_{t\in I}\mathbb{E}[Z_t]\nu_t = V^\nu_{\text{ind}}(Z),\end{aligned}$$

because it holds then that $\mathbb{E}[Z_t] \le \mathbb{E}[Z_{t-1}]$ for all $t \in I$, in particular $\mathbb{E}[Z_T] \le \mathbb{E}[Z_t]$ for all $t \in I$. The equality holds only for martingales. The results of Lemma 3.4.5 and the version for supermartingales, [33, Lemma 4.35], were not observed in [33, Section 5.5], such that the binomial model was taken up again in this work as an example. The results for the binomial model are additionally extended, see Section 3.5.3. As we already have discussed, the assumption $\mathbb{E}[\sup_{t\in I}|Z_t|] < \infty$ for the process Z is a strong condition and does not depend on the given probability measure ν on I. For finite time domains like $I = \{0, \dots, T\}$ and processes that do not explode, this assumption makes sense, but the given information remains unused. Furthermore we have considered a more refined version by introducing the norms $\|\cdot\|_{\nu,p}$ for $p \in [1,\infty)$ and $\|\cdot\|_{\nu,p,q}$ for $(p,q) \in [1,\infty)\times[1,\infty]$, which uses the structure and information given by ν. For example it is enough when the processes satisfy such condition on the support of ν. This can give us the opportunity to look at many more processes. Now, we want to remember the important bounds from the Section 3.3. If we only have given a distribution ν on I and we do not have more information about this distribution. Then we consider processes Z which satisfy Theorem 3.3.5 and the following bounds.

Lemma 3.4.7 (Bounds)**.** *Let $p \in [1,\infty)$ and Z be a process which statisfies Theorem 3.3.5. Then for every $\gamma \in \mathcal{M}^\nu_I$*

(a)

$$\|Z\|_{\nu,p} \le \sup_{\gamma\in\mathcal{M}^\nu_I}\left(\mathbb{E}\Big[\sup_{s\in I}|Z_s|^p\sum_{t\in I}\gamma_t\Big]\right)^{1/p} \le \left(\mathbb{E}\Big[\sup_{s\in I}|Z_s|^p\Big]\right)^{1/p}.$$

(b)

$$\mathbb{E}[Z_\gamma] \le |\mathbb{E}[Z_\gamma]| \le \mathbb{E}\Big[\sum_{t\in I}|Z_t|\gamma_t\Big] \le \left(\mathbb{E}\Big[\sum_{t\in I}|Z_t|^p\gamma_t\Big]\right)^{1/p}\left(\sum_{t\in I}\nu_t\right)^{1/q} \le \|Z\|_{\nu,p}.$$

(c)

$$\|Z_\gamma\|_{L^p} \le \left(\mathbb{E}\Big[\sum_{t\in I}|Z_t|^p\gamma_t\Big]\right)^{1/p} \le \|Z\|_{\nu,p}.$$

(d)

$$\|Z_t\|_{L^p} \le \nu_t^{-1/p}\|Z\|_{\nu,p}, \quad t \in J,$$

where J is the support of ν given by $\operatorname{supp}(\nu) = \{t \in I \mid \nu_t > 0\}$.

Furthermore, for $1 \le p < r \le \infty$ we have for $Z \in X_{\nu,r}$ that

(e)

$$\|Z\|_{\nu,p} \le \|Z\|_{\nu,r}.$$

The first item of the lemma above shows us that the introduced norm is an improved bound, see Remark 3.3.15.
If we have more information about the given distribution ν on I, we consider processes Z which satisfy Theorem 3.3.34. The most important inequalities are given in Remark 3.3.39 and Lemma 3.3.48.

3.5. Results for Special Cases

3.5.1. Independent Stochastic Components

As described in the introduction, there are many situations in financial and actuarial mathematics where independence is assumed for two stochastic components. It is questionable whether this assumption is always correct, but we will take a look at it in this section, because it is a special case. The following lemma is the generalized version of [33, Corollary 2.39].

Lemma 3.5.1. *Consider a totally-ordered countable set I, a process $Z = (Z_t)_{t\in I}$ that satisfies Theorem 3.3.5 (or Theorem 3.3.34) and a $\gamma \in \mathcal{M}_I^\nu$ as in Definition 3.1.1. Then*

$$\mathbb{E}[Z_\gamma] = \sum_{t\in I} \mathbb{E}[Z_t \gamma_t]. \tag{3.5.2}$$

If Z and γ are independent, then

$$\mathbb{E}[Z_\gamma] = \sum_{t\in I} \mathbb{E}[Z_t]\mathbb{E}[\gamma_t] = \sum_{t\in I} \mathbb{E}[Z_t]\nu_t.$$

Proof. First, we consider $Z_t^n := \max\{-n, \min\{Z_t, n\}\} \nearrow Z_t$ as $n \to \infty$ for every $t \in I$. Then for every $n \in \mathbb{N}$ the process $Z^{(n)} = (Z_t^{(n)})_{t\in I}$ is bounded. Using $\sum_{t\in I} \gamma_t \le 1$, a.s., as in Definition 3.1.1 we get that

$$Z_\gamma^n = \sum_{t\in I} Z_t^n \gamma_t \le \Big(\sup_{t\in I} Z_t^n\Big) \sum_{t\in I} \gamma_t \le n, \quad \text{a.s.}$$

By dominated convergence (see [75, Theorem 5.9]) we see that we can exchange the expected value and the series, which proves (3.5.2) for Z^n. If Z_γ is well-defined, the process Z satisfies (3.5.2) by monotone convergence (see [75, Theorem 5.3]).
Now, we want to show that Z_γ is well-defined for a general process Z satisfying Theorem 3.3.5 (or Theorem 3.3.34). Note that $\mathbb{E}[Z_\gamma^n] < \infty$ and Z_γ^n is well-defined for all $n \in \mathbb{N}$.

- For processes Z satisfying Theorem 3.3.5, we have on the one hand that $\mathbb{E}[|Z_\gamma^n|] \leq \mathbb{E}[|Z_\gamma|]$ and on the other hand we get with (3.3.28)

$$\mathbb{E}[|Z_\gamma|] \leq \mathbb{E}[|Z_\gamma - Z_\gamma^n|] + \mathbb{E}[|Z_\gamma^n|] \leq \|Z - Z^n\|_{\nu,p} + \mathbb{E}[|Z_\gamma^n|].$$

  We have that $\mathbb{E}[|Z_\gamma^n|] < \infty$ and $\|Z - Z^n\|_{\nu,p} \to 0$ as $n \to \infty$ by Theorem 3.3.5.

- For a process Z satisfying Theorem 3.3.34, we have on the one hand that $\mathbb{E}[|Z_\gamma^n|] \leq \mathbb{E}[|Z_\gamma|]$ and on the other hand by using Remark 3.3.39 we get

$$\mathbb{E}[|Z_\gamma|] \leq \mathbb{E}[|Z_\gamma - Z_\gamma^n|] + \mathbb{E}[|Z_\gamma^n|] \leq \mathbb{E}[|Z_\gamma^n|] + \|Z - Z^n\|_{\nu,p,q} \bigg(\sum_{t \in I'} \nu_t^{q'/p'} \bigg)^{1/q'}.$$

  We have that $\mathbb{E}[|Z_\gamma^n|] < \infty$ and $\|Z - Z^n\|_{\nu,p,q} \to 0$ as $n \to \infty$ by Theorem 3.3.34.

Thus $\mathbb{E}[|Z_\gamma|] < \infty$ and by dominated convergence (see [75, Theorem 5.9]) we have that Z_γ is well-defined. Due to the independence of Z and γ, we have $\mathbb{E}[Z_t \gamma_t] = \mathbb{E}[Z_t]\,\mathbb{E}[\gamma_t]$ for all $t \in I$. ❑

3.5.2. Product of a Martingale and a Deterministic Function

The results of [33, Subsection 5.6] can be corrected, generalized and extended in the following way. We resort to the preliminaries from Section A.1.
Let I be a totally ordered countable index set, M be a martingale, ν be a given distribution on I and the support of ν is defined as

$$J = \operatorname{supp}(\nu) := \{t \in I \mid \nu_t > 0\}. \tag{3.5.3}$$

The considered, adapted process Z is given in the form

$$Z_t = f(t) M_t, \quad t \in I, \tag{3.5.4}$$

where f is an element of the set

$$\mathcal{F}_\nu(M) := \Big\{ f : I \to \mathbb{R} \;\Big|\; f \text{ non-decreasing function}, \sum_{t \in I} |f(t) M_t| \gamma_t \text{ integrable for all } \gamma \in \mathcal{M}_I^\nu \Big\}.$$

Using the adapted random probability measure to describe our problem we are interested in

$$V_{\mathcal{M}}^\nu(Z) := \sup_{\gamma \in \mathcal{M}_I^\nu} \mathbb{E}\bigg[\sum_{t \in I} Z_t \gamma_t \bigg].$$

The main theorem of this section will give us a characterization of an optimal strategy for this problem. We generalize the Lemma [33, Lemma 5.55] and additionally extend it to an if-and-only-if condition. One part of the proof is leaned on the one given in [33], the other parts are proved by using the known structure of any $f \in \mathcal{F}_\nu(M)$.

The Lemma [33, Lemma 5.55] starts with considering processes Z which are given as a product of a martingale M bounded from below and a non-decreasing deterministic function f and are satisfied $\mathbb{E}[\sup_{t\in I} Z^+] < \infty$ or $\mathbb{E}[\sup_{t\in I} Z^-] < \infty$. We were starting our consideration with the assumption for the martingale that $\mathbb{E}[\sup_{t\in I}|M_t|] < \infty$. This assumption makes sure that all value we are interested in are well-defined and finite. Step by step it was possible to generalize this assumption and we stop with the assumption that $\mathcal{F}_\nu(M)$ includes the identically one function. Before we formulate the main theorem we want to take a look at $\mathcal{F}_\nu(M)$ and some specific properties.

Lemma 3.5.5 ($\mathcal{F}_\nu(M)$)**.**

(a) For every $f \in \mathcal{F}_\nu(M)$ and $\gamma \in \mathcal{M}_I^\nu$ the series defining Z_γ is absolutely convergent almost surely and Z_γ is integrable.

(b) If $\mathcal{F}_\nu(M)$ includes the identically one function, then

(i) M_γ is absolutely convergent almost surely and integrable for every $\gamma \in \mathcal{M}_I^\nu$,

(ii) all bounded non-decreasing functions $f : I \to \mathbb{R}$ are in $\mathcal{F}_\nu(M)$, especially all constant functions.

Proof. (a) For every $f \in \mathcal{F}_\nu(M)$ we have that $\sum_{t\in I}|f(t)M_t|\gamma_t$ is integrable for all $\gamma \in \mathcal{M}_I^\nu$ such that

$$\mathbb{E}[|Z_\gamma|] = \mathbb{E}\left[\left|\sum_{t\in I} Z_t\gamma_t\right|\right] \le \mathbb{E}\left[\sum_{t\in I}|Z_t|\gamma_t\right] = \mathbb{E}\left[\sum_{t\in I}|f(t)M_t|\gamma_t\right] < \infty.$$

This implies the almost surely absolute convergence of Z_γ and using dominated convergence we have that Z_γ is integrable. In addition, we get for every $\gamma \in \mathcal{M}_I^\nu$ that

$$\mathbb{E}[Z_\gamma] = \mathbb{E}\left[\sum_{t\in I} Z_t\gamma_t\right] = \mathbb{E}\left[\sum_{t\in I} f(t)M_t\gamma_t\right] = \sum_{t\in I} f(t)\mathbb{E}[M_t\gamma_t].$$

(b) (i) If $\mathcal{F}_\nu(M)$ includes the identically one function, we have that $\sum_{t\in I}|M_t|\gamma_t$ is integrable for all $\gamma \in \mathcal{M}_I^\nu$ such that M_γ is absolute convergent almost surely and integrable, cf. the thoughts above. Moreover $\mathbb{E}[M_\gamma]$ exists and is finite.

(ii) $\mathcal{F}_\nu(M)$ includes the identically one function which means that $\sum_{t\in I}|M_t|\gamma_t$ is integrable for all $\gamma \in \mathcal{M}_I^\nu$. Moreover, for every bounded non-decreasing functions $f : I \to \mathbb{R}$ it holds that $\sup_{t\in I}|f(t)| < \infty$. Using these statements we get that $\sum_{t\in I}|f(t)M_t|\gamma_t \le \sup_{t\in I}|f(t)|\sum_{t\in I}|M_t|\gamma_t$ is integrable for all $\gamma \in \mathcal{M}_I^\nu$. Thus $f \in \mathcal{F}_\nu(M)$. Especially for every constant function $f \equiv c$, $c \in \mathbb{R}$, we have the equality $\sum_{t\in I}|f(t)M_t|\gamma_t = |c|\sum_{t\in I}|M_t|\gamma_t$. ❑

Theorem 3.5.6. *Given a totally ordered countable index set $I \subseteq \mathbb{R} \cup \{-\infty, \infty\}$, a probability distribution ν on I and a martingale M. Assume that $\mathcal{F}_\nu(M)$ contains the identically one function. Then for an adapted random probability measure $\gamma^* \in \mathcal{M}_I^\nu$ the following properties are equivalent:*

(a) *γ^* is optimal for all processes $(Z_t)_{t\in I}$ given by*

$$Z_t = f(t)M_t, \quad t \in I, \tag{3.5.7}$$

with $f \in \mathcal{F}_\nu(M)$ and f is bounded.

(b) *γ^* satisfies $\mathbb{E}[M_{\gamma^*}] = \mathbb{E}[M_\gamma]$ and*

$$\mathbb{E}\Bigg[\sum_{t\in I_{>s}} M_t \gamma_t^*\Bigg] \geq \mathbb{E}\Bigg[\sum_{t\in I_{>s}} M_t \gamma_t\Bigg] \tag{3.5.8}$$

for all $s \in I$ and all $\gamma \in \mathcal{M}_I^\nu$.

(c) *γ^* is optimal for all processes $(Z_t)_{t\in I}$ given by*

$$Z_t = f(t)M_t, \quad t \in I,$$

with $f \in \mathcal{F}_\nu(M)$.

Remark 3.5.9. Note that the condition (3.5.8) is independent from the choice of the non-decreasing function $f \in \mathcal{F}_\nu(M)$, such that we get an optimal adapted random probability measure for the whole class of processes, which can be written as a product of a martingale and a deterministic function and satisfies certain conditions.

Proof of Theorem 3.5.6. We prove that $(a) \Leftrightarrow (b)$ and $(c) \Leftrightarrow (b)$. The proof of the return direction is the same in both cases. We will give two different proofs of the only-if-part. One proof uses the knowledge about the structure of our considered functions $f \in \mathcal{F}_\nu(M)$ and the other one uses the dominance in first order.

1. (a) implies (b):
Note that based on the claimed conditions we have that Z_γ and M_γ are well-defined and in L^1 for all $\gamma \in \mathcal{M}_I^\nu$, see Lemma 3.5.5. A random probability measure $\gamma^* \in \mathcal{M}_I^\nu$ is optimal for $(Z_t)_{t\in I}$, if $\mathbb{E}[Z_{\gamma^*}] \geq \mathbb{E}[Z_\gamma]$ for all $\gamma \in \mathcal{M}_I^\nu$. For every $(Z_t)_{t\in I}$ given in the form as in condition (a) of Theorem 3.5.6, the optimality of γ^* implies

$$\mathbb{E}[Z_{\gamma^*}] = \mathbb{E}\Bigg[\sum_{t\in I} f(t)M_t\gamma_t^*\Bigg] \geq \mathbb{E}[Z_\gamma] = \mathbb{E}\Bigg[\sum_{t\in I} f(t)M_t\gamma_t\Bigg], \quad \forall \gamma \in \mathcal{M}_I^\nu. \tag{3.5.10}$$

For a fixed $s \in I$ the function $f_s : I \to \mathbb{R}$ is defined as

$$f_s(t) := \mathbb{1}_{I_{>s}}(t), \quad t \in I, \tag{3.5.11}$$

is a special non-decreasing deterministic function and bounded by 1. We have that

$$f_s(t)M_t = \begin{cases} 0 & \text{for } t \in I_{\leq s}, \\ M_t & \text{for } t \in I_{>s}, \end{cases}$$

and $\sum_{t\in I}|f_s(t)M_t|\gamma_t = \sum_{t\in I_{>s}}|M_t|\gamma_t \le \sum_{t\in I}|M_t|\gamma_t$. Due to the identically one function being in $\mathcal{F}_\nu(M)$, we get that $f_s \in \mathcal{F}_\nu(M)$. The inequality (3.5.10) holds for every non-decreasing deterministic function $f \in \mathcal{F}_\nu(M)$, particularly for f_s, $s \in I$. For $s \in I$ and for every $\gamma \in \mathcal{M}_I^\nu$ we have

$$\mathbb{E}\Big[\sum_{t\in I} f_s(t)M_t\gamma_t\Big] = \mathbb{E}\Big[\sum_{t\in I_{>s}} M_t\gamma_t\Big], \tag{3.5.12}$$

so that we get for every $s \in I$ and every $\gamma \in \mathcal{M}_I^\nu$ with the special choice f_s that

$$\mathbb{E}\Big[\sum_{t\in I_{>s}} M_t\gamma_t^*\Big] \ge \mathbb{E}\Big[\sum_{t\in I_{>s}} M_t\gamma_t\Big].$$

Note that we need M_γ is well-defined and integrable for the existence of the terms above. Now, we want to show that $\mathbb{E}[M_\gamma]$ is the same real number for all $\gamma \in \mathcal{M}_I^\nu$. Applying the inequation (3.5.10) for the identically one function, we get immediately that

$$\mathbb{E}[M_{\gamma^*}] \ge \mathbb{E}[M_\gamma], \quad \forall \gamma \in \mathcal{M}_I^\nu. \tag{3.5.13}$$

If $\mathcal{F}_\nu(M)$ includes the identically one function, then we know that all constant functions are in $\mathcal{F}_\nu(M)$, see Lemma 3.5.5. Therefore we can also apply the inequation (3.5.10) for the function which is identically minus one, i.e., $g(t) = -1$, for all $t \in I$. Note that g is bounded by one. We get for every $\gamma \in \mathcal{M}_I^\nu$ that

$$\mathbb{E}[M_{\gamma^*}] \le \mathbb{E}[M_\gamma], \quad \forall \gamma \in \mathcal{M}_I^\nu. \tag{3.5.14}$$

Putting the inequations (3.5.13) and (3.5.14) together we get the assertion.

2. (b) implies (a) and (b) implies (c):
For every $t \in I$ and $\gamma \in \mathcal{M}_I^\nu$ we define

$$\mu_{\gamma,t} := \mathbb{E}[M_t\gamma_t]. \tag{3.5.15}$$

- Because $\mathcal{F}_\nu(M)$ contains the identically one function, we get that M_γ is well-defined and integrable, cf. Lemma 3.5.5. Thus we get that μ_γ is σ-additive. Using that $|\mathbb{E}[M_t\gamma_t]| \le \mathbb{E}[|M_t|\gamma_t]$ and monotone convergence we have additionally that

$$|\mu_\gamma|(I) = \sum_{t\in I}|\mathbb{E}[M_t\gamma_t]| \le \sum_{t\in I}\mathbb{E}[|M_t|\gamma_t] = \mathbb{E}\Big[\sum_{t\in I}|M_t|\gamma_t\Big] < \infty.$$

  Therefore μ_γ is a signed measure of finite total variation.

- Due to $\mathbb{E}[M_\gamma]$ is the same real number for all $\gamma \in \mathcal{M}_I^\nu$, we have for every $\gamma, \tilde{\gamma} \in \mathcal{M}_I^\nu$ that $\mu_\gamma(I) = \mathbb{E}[M_\gamma] = \mathbb{E}[M_{\tilde{\gamma}}] = \mu_{\tilde{\gamma}}(I)$.

Now, we assume that $\gamma^* \in \mathcal{M}_I^\nu$ is optimal and satisfies

$$\mathbb{E}\Big[\sum_{t\in I_{>s}} M_t\gamma_t^*\Big] \ge \mathbb{E}\Big[\sum_{t\in I_{>s}} M_t\gamma_t\Big]$$

for all $s \in I$ and $\gamma \in \mathcal{M}_I^\nu$. It follows with monotone convergence that for all $s \in I$

$$\sum_{t \in I_{>s}} \mathbb{E}[M_t \gamma_t^*] \geq \sum_{t \in I_{>s}} \mathbb{E}[M_t \gamma_t].$$

Using equation (3.5.15) we have

$$\sum_{t \in I_{>s}} \mu_{\gamma^*,t} \geq \sum_{t \in I_{>s}} \mu_{\gamma,t}, \quad \forall s \in I,$$

this means μ_{γ^*} dominates μ_γ in first order, cf. Definition A.1.11. Due to Lemma A.1.15 we have equivalently

$$\sum_{t \in I} f(t) \mu_{\gamma^*,t} \geq \sum_{t \in I} f(t) \mu_{\gamma,t}, \tag{3.5.16}$$

for all non-decreasing functions f for which the expectations exist. Because of the claimed conditions of Theorem 3.5.6 we know that the expectations in equation (3.5.16) exist for all $f \in \mathcal{F}_\nu(M)$, especially also for bounded functions, and Z_γ is well-defined and integrable. Therefore we have by monotone convergence that

$$\mathbb{E}[Z_\gamma] = \mathbb{E}\left[\sum_{t \in I} Z_t \gamma_t\right] = \mathbb{E}\left[\sum_{t \in I} f(t) M_t \gamma_t\right] = \sum_{t \in I} f(t) \mathbb{E}[M_t \gamma_t] = \sum_{t \in I} f(t) \mu_{\gamma,t}. \tag{3.5.17}$$

Finally, with equation (3.5.17) and inequation (3.5.16) it follows for every $(Z_t)_{t \in I}$ given in the form as in condition (a) of Lemma 3.5.6 that

$$\mathbb{E}[Z_{\gamma^*}] \geq \mathbb{E}[Z_\gamma], \quad \forall \gamma \in \mathcal{M}_I^\nu,$$

such that we get the assertion.

3. Alternative proof for (b) implies (a) and (b) implies (c):
Let $\gamma^* \in \mathcal{M}_I^\nu$ satisfying

$$\mathbb{E}\left[\sum_{t \in I_{>s}} M_t \gamma_t^*\right] \geq \mathbb{E}\left[\sum_{t \in I_{>s}} M_t \gamma_t\right]$$

for all $s \in I$ and $\gamma \in \mathcal{M}_I^\nu$. Then using equation (3.5.12) and $f_s \in \mathcal{F}_\nu(M)$ for all $s \in I$ which is given in (3.5.11) we get that $\gamma^* \in \mathcal{M}_I^\nu$ is optimal for $(Z_t)_{t \in I}$ with $Z_t = f_s(t) M_t = \mathbb{1}_{I_{>s}}(t) M_t$. Thus we have

$$\mathbb{E}\left[\sum_{t \in I} \mathbb{1}_{I_{>s}}(t) M_t \gamma_t^*\right] \geq \mathbb{E}\left[\sum_{t \in I} \mathbb{1}_{I_{>s}}(t) M_t \gamma_t\right], \quad \forall \gamma \in \mathcal{M}_I^\nu. \tag{3.5.18}$$

Now, we show this statement for any non-decreasing deterministic function in $\mathcal{F}_\nu(M)$. For simplification we consider the case that I has no accumulation points. If I has some accumulation points, we only need additionally some limit arguments.

(i) Let $f \in \mathcal{F}_\nu(M)$ be bounded and non-negative. Then f can be written as linear combination of the f_s given in (3.5.11), i.e.,

$$f(t) = a + \sum_{s \in I} a_s f_s(t) = a + \sum_{s \in I_{\leq t}} a_s f_s(t) = a + \sum_{s \in I_{\leq t}} a_s \mathbb{1}_{I_{>s}}(t), \quad t \in I,$$

where a is constant and a_s, $s \in I$, are suitably chosen coefficients which are given by

$$a_s := \inf_{t \in I_{>s}} (f(t) - f(s)) \geq 0.$$

These coefficients describe the size of the jumps. Due to f is non-negative, it holds that $a_s \geq 0$ for every $s \in I$ and $a \geq 0$. Using inequality (3.5.18) and that $\mathbb{E}[M_\gamma] = \mathbb{E}[M_{\gamma^*}]$ for $\gamma, \gamma^* \in \mathcal{M}_I^\nu$ we get

$$\begin{aligned}
\mathbb{E}[Z_{\gamma^*}] &= \mathbb{E}\Big[\sum_{t\in I} f(t) M_t \gamma_t^*\Big] = \mathbb{E}\Big[\sum_{t\in I}\Big(a + \sum_{s \in I_{\leq t}} a_s \mathbb{1}_{\{s<t\}}\Big) M_t \gamma_t^*\Big] \\
&= a\,\mathbb{E}[M_{\gamma^*}] + \mathbb{E}\Big[\sum_{t\in I}\sum_{s\in I} a_s \mathbb{1}_{\{s<t\}} M_t \gamma_t^*\Big] = a\,\mathbb{E}[M_{\gamma^*}] + \sum_{s\in I} a_s \mathbb{E}\Big[\sum_{t\in I} \mathbb{1}_{\{s<t\}} M_t \gamma_t^*\Big] \\
&\overset{(3.5.18)}{\geq} a\,\mathbb{E}[M_\gamma] + \sum_{s\in I} a_s \mathbb{E}\Big[\sum_{t\in I} \mathbb{1}_{\{s<t\}} M_t \gamma_t\Big] \\
&= \mathbb{E}\Big[\sum_{t\in I}\Big(a + \sum_{s\in I_{\leq t}} a_s \mathbb{1}_{\{s<t\}}\Big) M_t \gamma_t\Big] = \mathbb{E}[Z_\gamma].
\end{aligned}$$

The rearrangement of the sums is possible, because the sums are well-defined due to $\sum_{t\in I} |f(t) M_t| \gamma_t$ is integrable and $\mathbb{E}[M_\gamma] < \infty$ for all $\gamma \in \mathcal{M}_I^\nu$. So we have that $\gamma^* \in \mathcal{M}_I^\nu$ is optimal for $(Z_t)_{t\in I}$ with $f \in \mathcal{F}_\nu(M)$ is bounded and non-negative.

(ii) Let $f \in \mathcal{F}_\nu(M)$ be bounded. We know that every f has the representation $a + \sum_{s\in I_{\leq t}} a_s f_s(t)$. We choose $c := \|f\|_\infty$. Then $f + c$ is bounded and non-negative such we can use the thoughts above. It holds that

$$\begin{aligned}
\mathbb{E}[Z_{\gamma^*}] &= \mathbb{E}\Big[\sum_{t\in I} (f(t) + c - c) M_t \gamma_t^*\Big] = \mathbb{E}\Big[\sum_{t\in I} \underbrace{(f(t)+c)}_{\geq 0} M_t \gamma_t^*\Big] - c\,\underbrace{\mathbb{E}\Big[\sum_{t\in I} M_t \gamma_t^*\Big]}_{=\mathbb{E}[M_{\gamma^*}]<\infty} \\
&\geq \mathbb{E}\Big[\sum_{t\in I} (f(t)+c) M_t \gamma_t\Big] - c\,\mathbb{E}[M_{\gamma^*}] = \mathbb{E}\Big[\sum_{t\in I} f(t) M_t \gamma_t\Big] + c\,\underbrace{(\mathbb{E}[M_\gamma] - \mathbb{E}[M_{\gamma^*}])}_{=0} \\
&= \mathbb{E}[Z_\gamma].
\end{aligned}$$

Thereby we get also that $\gamma^* \in \mathcal{M}_I^\nu$ is optimal for $(Z_t)_{t\in I}$ with $f \in \mathcal{F}_\nu(M)$ is bounded.

(iii) Now, we consider the case that $f \in \mathcal{F}_\nu(M)$ and f is unbounded. Therefore let $\alpha \in \mathbb{R}$ and $\alpha \leq 0$. We define $g := \alpha \vee f$. Then g is in $\mathcal{F}_\nu(M)$, bounded from below and we know that $\gamma^* \in \mathcal{M}_I^\nu$ is optimal for $(Z_t)_{t\in I}$ with $Z_t = g(t) M_t$. Furthermore we have that $|g(t) M_t| = |(\alpha \vee f)(t)||M_t| \leq |Z_t|$ and that $g(t) = (\alpha \vee f)(t)$ converge to $f(t)$ for $\alpha \to -\infty$. Using dominated convergence we get our claim.

4. (c) implies (b):
It follows immediately using (a) implies (b). ❑

Remark 3.5.19. The conditions in Lemma 3.5.6 are chosen in such a way that the key properties are satisfied for the above proof. So, the key properties used and needed are:

(a) Z_γ is well-defined and integrable for every $\gamma \in \mathcal{M}_I^\nu$,

(b) M_γ is well-defined and integrable for every $\gamma \in \mathcal{M}_I^\nu$,

(c) $\mathbb{E}[M_\gamma] = \mathbb{E}[M_{\tilde{\gamma}}]$ for all $\gamma, \tilde{\gamma} \in \mathcal{M}_I^\nu$.

The next lemmas and remarks give us different assumptions for the processes Z and M such that the key properties given in Remark 3.5.19 are fulfilled. We will show them directly.

Lemma 3.5.20. *Given a totally ordered countable set I. Z_γ is well-defined and integrable for every $\gamma \in \mathcal{M}_I^\nu$ if Z satisfies one of the following conditions:*

(a) $\sum_{t\in I} |f(t)M_t|\gamma_t$ *integrable for all* $\gamma \in \mathcal{M}_I^\nu$,

(b) $\mathbb{E}[\sup_{t\in I} |Z_t|] < \infty$,

(c) $\mathbb{E}[\sup_{t\in \operatorname{supp}(\nu)} |Z_t|] < \infty$.

Proof. (a) The condition $\sum_{t\in I} |f(t)M_t|\gamma_t$ integrable for all $\gamma \in \mathcal{M}_I^\nu$ implies that $f \in \mathcal{F}_\nu(M)$. Due to Lemma 3.5.5 we know that Z_γ is well-defined and integrable.

(b) Let $\mathbb{E}[\sup_{t\in I} |Z_t|] < \infty$. Using that $|\sum_{t\in I} Z_t\gamma_t| \le \sum_{t\in I} |Z_t|\gamma_t$ and monotone convergence, we get that

$$\mathbb{E}[|Z_\gamma|] = \mathbb{E}\left[\left|\sum_{t\in I} Z_t\gamma_t\right|\right] \le \sum_{t\in I} \mathbb{E}[|Z_t|\gamma_t] \le \mathbb{E}\left[\sup_{t\in I} |Z_t| \underbrace{\sum_{t\in I} \gamma_t}_{=1 \text{ a.s.}}\right] = \mathbb{E}\left[\sup_{t\in I} |Z_t|\right] < \infty.$$

The absolute convergence almost surely of Z_γ implies immediately that Z_γ is well-defined and integrable.

(c) If $\mathbb{E}[\sup_{t\in \operatorname{supp}(\nu)} |Z_t|] < \infty$ hold we can prove the claim analogously to (2), because the restriction on the support of ν does not change anything. For every $t \in I \setminus \operatorname{supp}(\nu)$ the correspondent summands would be equal to zero, because $\nu_t = 0$ for $t \in I$ implies that $\gamma_t = 0$, a.s. Therefore we have that $\sum_{t\in I} \mathbb{E}[Z_t\gamma_t] = \sum_{t\in J} \mathbb{E}[Z_t\gamma_t]$, where $J = \{t \in I | \nu_t > 0\}$. Note that the restriction is a weaker condition on the process. ❑

The following lemma show us that the condition (2) and (3) of Lemma 3.5.20 are much stronger.

Lemma 3.5.21. *Let I' be a subset of a totally ordered countable index set I, M be a martingale and $f : I \to \mathbb{R}$ be a non-decreasing deterministic function. Then $\mathbb{E}[\sup_{t\in I'} |f(t)M_t|] < \infty$ implies that $f \in \mathcal{F}_\nu(M)$ for all $\nu \in \mathcal{D}_1(I)$ with $\operatorname{supp}(\nu) \subseteq I'$, where $\mathcal{D}_1(I)$ is the set of all probability distribution on I.*

Proof. For a fixed $\nu \in \mathcal{D}_1(I)$ with $\operatorname{supp}(\nu) \subseteq I'$ we have to show that $\sum_{t\in I} |f(t)M_t|\gamma_t$ is integrable for all $\gamma \in \mathcal{M}_I^\nu$. Note that for every $t \in I \setminus \operatorname{supp}(\nu)$ we know that $\nu_t = 0$. This implies $\gamma_t = 0$ a.s for these $t \in I \setminus \operatorname{supp}(\nu)$ and every correspondent summand would be equal to zero. Therefore we have that

$$\mathbb{E}\Big[\sum_{t\in I} |f(t)M_t|\gamma_t\Big] = \mathbb{E}\Big[\sum_{t\in \operatorname{supp}(\nu)} |f(t)M_t|\gamma_t\Big] = \mathbb{E}\Big[\sum_{t\in I'} |f(t)M_t|\gamma_t\Big].$$

Using the equality above we have that

$$\mathbb{E}\Big[\sum_{t\in I} |f(t)M_t|\gamma_t\Big] = \mathbb{E}\Big[\sum_{t\in I'} |f(t)M_t|\gamma_t\Big] \leq \mathbb{E}\Big[\sup_{t\in I'} |f(t)M_t| \underbrace{\sum_{t\in I} \gamma_t}_{\leq 1 \text{ a.s.}}\Big] \leq \mathbb{E}\Big[\sup_{t\in I'} |f(t)M_t|\Big] < \infty,$$

such that $\sum_{t\in I} |f(t)M_t|\gamma_t$ is integrable for all $\gamma \in \mathcal{M}_I^\nu$ and therefore $f \in \mathcal{F}_\nu(M)$. ❑

Lemma 3.5.22. *Given a totally ordered countable set I. For every $\gamma \in \mathcal{M}_I^\nu$, the random variable M_γ is well-defined and integrable, if the martingale M satisfies one of the following conditions:*

(a) $\mathbb{E}[\sup_{t\in I} |M_t|] < \infty$,

(b) $M = (M_t)_{t\in I}$ *is closable,*

(c) $\sum_{t\in I} |M_t|\gamma_t$ *integrable for all* $\gamma \in \mathcal{M}_I^\nu$,

(d) $\mathbb{E}[\sup_{t\in \operatorname{supp}(\nu)} |M_t|] < \infty$,

(e) $M = (M_t)_{t\in \operatorname{supp}(\nu)}$ *is closable.*

Remark 3.5.23. Note that the condition (a) and (b) are independent from ν such that the lemma would hold for the set $\mathcal{M}_I$ instead of $\mathcal{M}_I^\nu$. The condition (c), (d) and (e) depend on ν. The restriction on the support of ν does not change anything in the calculations. This means that for every $t \in I \setminus \operatorname{supp}(\nu)$ the correspondent summands would be equal to zero, because $\nu_t = 0$ for $t \in I$ implies that $\gamma_t = 0$, a.s. Therefore we have that $\sum_{t\in I} \mathbb{E}[Z_t\gamma_t] = \sum_{t\in J} \mathbb{E}[Z_t\gamma_t]$, where $J = \{t \in I | \nu_t > 0\}$. The restriction on the support of ν, however, places a weaker condition on the martingale.

Proof of Lemma 3.5.22. We want to prove that if M satisfies one of the conditions given in Lemma 3.5.22, the random variable M_γ is well-defined and integrable for every $\gamma \in \mathcal{M}_I^\nu$. For condition (a), (c) and (d) the proof is the same like the proof of Lemma 3.5.20. Now, let $M = (M_t)_{t\in I}$ be a closable martingale. The same consideration are possible for $M = (M_t)_{t\in \operatorname{supp}(\nu)}$, because the restriction on the support of ν does not change the calculations. Using the property that M is closable, i.e., $M_t = \mathbb{E}[M_\infty|\mathcal{F}_t]$ for all $t \in I$, we get with Jensen's inequality that

$$|M_t| = |\mathbb{E}[M_\infty|\mathcal{F}_t]| \leq \mathbb{E}[|M_\infty||\mathcal{F}_t]$$

such that using γ_t is $\mathcal{F}_t$-measurable, we have

$$\mathbb{E}\Big[\sum_{t\in I} |M_t|\gamma_t\Big] = \mathbb{E}\Big[\sum_{t\in I} |\mathbb{E}[M_\infty|\mathcal{F}_t]|\gamma_t\Big] \leq \mathbb{E}\Big[\sum_{t\in I} \mathbb{E}[|M_\infty|\gamma_t|\mathcal{F}_t]\Big].$$

Furthermore, using monotone convergence we get that

$$\mathbb{E}\Big[\sum_{t\in I}\mathbb{E}[|M_\infty|\gamma_t|\mathcal{F}_t]\Big]=\sum_{t\in I}\mathbb{E}[|M_\infty|\gamma_t]=\mathbb{E}\Big[|M_\infty|\underbrace{\sum_{t\in I}\gamma_t}_{=1\text{ a.s.}}\Big]=\mathbb{E}[|M_\infty|]<\infty.$$

Using dominated convergence we have that $\sum_{t\in I}M_t\gamma_t\in L^1$ and we can define $M_\gamma:=\sum_{t\in I}M_t\gamma_t$ which is well-defined and integrable. ❑

Remark 3.5.24. (a) Note that we get the same result as in Lemma 3.5.6, if we use respectively one of the condition in Remark 3.5.20 and Remark 3.5.22 for the definition of $\mathcal{F}_\nu(M)$ and the condition of M. Then we get that Z_γ and M_γ are well-defined and integrable. The rest of the proof of the modified Theroem 3.5.6 would be analog.

(b) With $\mathbb{E}[\sup_{t\in\operatorname{supp}(\nu)}|M_t|]<\infty$ and $\mathbb{E}[\sup_{t\in\operatorname{supp}(\nu)}|Z_t|]<\infty$ we restrict the conditions on the support of ν. This is a weaker condition on the martingale and process.

Remark 3.5.25. Note that if M is a closable martingale, then $\mathbb{E}[\sup_{t\in I}|M_t|]<\infty$. The other direction does not hold in generality. Therefore we will give an example of a closable martingale which has not an integrable majorant.

Example 3.5.26. On the probability space $((0,1],\mathcal{B}_{(0,1]},\mathbb{P})$ with Lebesgue-Borel measure $\mathbb{P}$ define the random variable

$$M(\omega)=\frac{1}{\omega\log^2(e/\omega)},\qquad \omega\in(0,1].$$

Then M is integrable, because by the fundamental theorem of calculus

$$\mathbb{E}[M]=\int_{(0,1]}\frac{dx}{x\log^2(e/x)}=\lim_{\epsilon\to0}\int_\epsilon^1\frac{dx}{x\log^2(e/x)}=\lim_{\epsilon\to0}\frac{1}{\log(e/x)}\Big|_\epsilon^1=1,$$

because

$$\frac{d}{dx}\log(e/x)^{-1}=(-1)\cdot\log(e/x)^{-2}\cdot\frac{x}{e}\cdot\Big(-\frac{e}{x^2}\Big)=\frac{1}{x\log^2(e/x)}.$$

For $t\in[1,\infty)$ define $\mathcal{F}_t=\{A\cup B\mid A\in\{\varnothing,(0,1/t]\},\,B\in\mathcal{B}_{(1/t,1]}\}$.

(a) $(\mathcal{F}_t)_{t\geq1}$ is a filtration.
Show that $\mathcal{F}_t$ is a σ-algebra for all $t\in[1,\infty)$.

a) We have with $\varnothing\in\{\varnothing,(0,1/t]\}$ and $\varnothing\in\mathcal{B}_{(1/t,1]}$ that $\varnothing=\varnothing\cup\varnothing\in\mathcal{F}_t$.

b) Let $C\in\mathcal{F}_t$, then we get for the complement with De Morgan's laws that $C^c=A^c\cap B^c=\varnothing\in\mathcal{F}_t$.

c) Let $C_n\in\mathcal{F}_t$ for all $n\in\mathbb{N}$. There exist $A_n\in\{\varnothing,(0,1/t]\}$ and $B_n\in\mathcal{B}_{(1/t,1]}$ such that $C_n=A_n\cup B_n$. Then

$$\bigcup_{n\in\mathbb{N}}C_n=\underbrace{\bigcup_{n\in\mathbb{N}}A_n}_{=\varnothing\text{ or }(0,1/t]}\cup\underbrace{\bigcup_{n\in\mathbb{N}}B_n}_{\in\mathcal{B}_{(1/t,1]}}\in\mathcal{F}_t.$$

Show that for $1 < s \leq t \Rightarrow \mathcal{F}_s \subseteq \mathcal{F}_t$:

$$1 < s \leq t \Rightarrow \frac{1}{t} \leq \frac{1}{s} \Rightarrow (1/s, 1] \subseteq (1/t, 1] \text{ and } (0, 1/t] \subseteq (0, 1/s].$$

Therefore for any $B \in \mathcal{B}_{(1/s,1]}$ it follows that $B \in \mathcal{B}_{(1/t,1]}$. Furthermore we have that

$$(0, 1/s] = \underbrace{(0, 1/t]}_{\in \mathcal{F}_t} \cup \underbrace{(1/t, 1/s]}_{\in \mathcal{B}_{(1/t,1]}} \in \mathcal{F}_t.$$

It follows that $\mathcal{F}_s \subseteq \mathcal{F}_t$.

(b) Then we get that

$$\mathbb{E}[M|\mathcal{F}_t](\omega) \overset{\text{a.s.}}{=} \begin{cases} M(\omega) & \text{for } \omega \in (1/t, 1], \\ \frac{t}{\log(et)} & \text{for } \omega \in (0, 1/t], \end{cases} \tag{3.5.27}$$

and that

$$Y(\omega) := \frac{1}{\omega \log(e/\omega)}, \quad \omega \in (0, 1],$$

is the smallest random variable dominating the right-hand side of (3.5.27) pointwise for every $t \geq 1$. Y is not integrable, because of

$$\begin{aligned}
\mathbb{E}[Y] &= \int_{(0,1]} \frac{d\omega}{\omega \log(e/\omega)} = \lim_{\epsilon \to 0} \int_{\epsilon}^{1} \frac{d\omega}{\omega \log(e/\omega)} \\
&\quad \text{substitute: } \frac{e}{\omega} = x \Leftrightarrow -\frac{e}{\omega^2} d\omega = dx \Leftrightarrow \frac{d\omega}{\omega} = -\frac{dx}{x} \\
&= \lim_{\epsilon \to 0} \int_{e/\epsilon}^{e} -\frac{1}{x \log(x)} dx \qquad \left(\frac{d}{dx} \log(|f(x)|) = \frac{f'(x)}{f(x)} \right) \\
&= \lim_{\epsilon \to 0} -\log(|\log(e/\omega)|) \Big|_{\epsilon}^{1} = \lim_{\epsilon \to 0} \Big(-\log(1) + \log(|\log(e/\epsilon)|) \Big) = +\infty.
\end{aligned}$$

Remark 3.5.28. Note that in general $\mathbb{E}[\sup_{t \in I} |Z_t|] < \infty$ does not imply that $\mathbb{E}[\sup_{t \in I} |M_t|] < \infty$. Take a look at the next example.

Example 3.5.29 (Continuation of Example 3.5.26). For every $\tilde{\omega} \in (0, 1]$ which is sufficiently big we define $f_{\tilde{\omega}} : (0, 1] \to \mathbb{R}$ as

$$f_{\tilde{\omega}}(\omega) = \begin{cases} 0 & \text{for } \omega \leq \tilde{\omega}, \\ 1 & \text{for } \omega > \tilde{\omega}, \end{cases}$$

and $Z(\omega) = f_{\tilde{\omega}}(\omega) M(\omega)$. Then we have that $\sup_{\omega \in (0,1]} |Z(\omega)| = \frac{1}{\tilde{\omega} \log^2(e/\tilde{\omega})} < \infty$. Hence with

$$\lim_{x \to 0} \frac{1}{x \log^2(e/x)} \overset{\text{L'Hospital}}{=} \lim_{x \to 0} \frac{1}{2x \log(e/x)} \overset{\text{L'Hospital}}{=} \lim_{x \to 0} \frac{1}{2x} = \infty,$$

we know that $\sup_{\omega \in (0,1]} |M(\omega)| = \sup_{\omega \in (0,1]} M(\omega) = \infty$.

Lemma 3.5.30. *Given a totally ordered countable set I. For every $\gamma \in \mathcal{M}_I^\nu$, the random variable M_γ is well-defined, integrable and satisfies $\mathbb{E}[M_\gamma] = \mathbb{E}[M_t]$ for all $t \in I$, if M satisfies one of the following conditions:*

(a) $\mathbb{E}[\sup_{t\in I} |M_t|] < \infty$,

(b) $\mathbb{E}[\sup_{t\in \mathrm{supp}(\nu)} |M_t|] < \infty$,

(c) *M is closable.*

Proof. If M satisfies one of the conditions given in Lemma 3.5.30, the random variable M_γ is well-defined and integrable for every $\gamma \in \mathcal{M}_I^\nu$. This is shown in Lemma 3.5.22. Now, we have to prove that we also get that $\mathbb{E}[M_\gamma] = \mathbb{E}[M_t]$ for all $t \in I$ and for every $\gamma \in \mathcal{M}_I^\nu$. For first two conditions it follows immediately from Theorem 3.4.1(f).
For the last condition we can repeat the calculation in the proof of Lemma 3.5.22 without absolute values, due to the absolute convergence almost surely of M_γ, and we get

$$\mathbb{E}[M_\gamma] = \mathbb{E}\left[\sum_{t\in I} M_t \gamma_t\right] = \sum_{t\in I} \mathbb{E}[\mathbb{E}[M_\infty \gamma_t | \mathcal{F}_t]] = \sum_{t\in I} \mathbb{E}[M_\infty \gamma_t] = \mathbb{E}\left[M_\infty \sum_{t\in I} \gamma_t\right]$$
$$= \mathbb{E}[M_\infty] = \mathbb{E}[M_t] \quad \forall t \in I,$$

because M is a closable martingale. ❑

The following corollary shows us a turnaround which does not hold in general.

Corollary 3.5.31. *Let $M = (M_t)_{t\in I}$ or $M = (M_t)_{t\in J}$ be a closable martingale. Then the following equivalence holds: $\mathcal{F}_\nu(M)$ contains the identically one function if and only if $\mathbb{E}[M_\gamma]$ is the same real number for all $\gamma \in \mathcal{M}_I^\nu$.*

Proof. If M is closable, we get for every $\gamma \in \mathcal{M}_I^\nu$ that $\mathbb{E}[M_\gamma] = \mathbb{E}[M_\infty] \in \mathbb{R}$, see the proof of Lemma 3.5.30. In the proof of Lemma 3.5.22 it is shown that $\sum_{t\in I} |M_t|\gamma_t$ is integrable. Thus it follows that the identically one function is in $\mathcal{F}_\nu(M)$. ❑

Remark 3.5.32. Note that $\mathbb{E}[\sup_{t\in I} |M_t|] < \infty$ ($\mathbb{E}[\sup_{t\in \mathrm{supp}(\nu)} |M_t|] < \infty$) implies that $\mathbb{E}[M_\gamma] = \mathbb{E}[M_t]$ for every $\gamma \in \mathcal{M}_I$ and $t \in I$ due to Theorem 3.4.1(f). Thereby we use the stopping theorem. What happen, if we only claim that $\mathbb{E}[M_\gamma] = \mathbb{E}[M_{\gamma^*}]$?

Remark 3.5.33. Note that $\mathbb{E}[M_\gamma]$ is the same real number for all $\gamma \in \mathcal{M}_I^\nu$ do not need the stopping theorem, e.g. there are martingales for which hold $\mathbb{E}[M_\gamma] = \mathbb{E}[M_{\gamma^*}]$ for all $\gamma, \gamma^* \in \mathcal{M}_I^\nu$, but $\mathbb{E}[M_\gamma] \neq \mathbb{E}[M_t]$ for every $\gamma \in \mathcal{M}_I^\nu$ and $t \in I$. (For example think about the standard Brownian motion B on the real line $\mathbb{R}$ starting at the origin and the time of hitting a single point different from the starting point 0. Furthermore we know that this hitting time for B has the Lévy distribution.) If we only claim that $\mathbb{E}[M_\gamma] = \mathbb{E}[M_{\gamma^*}]$, what about the proof of Theorem 3.4.1 without using the stopping times? These are still outstanding questions.

A special case is, if M is a closable martingale. Then the second condition in Theorem 3.5.6 has a certain form and we get the following corollary.

Corollary 3.5.34. *Given a totally ordered countable index set $I \subseteq \overline{\mathbb{R}}$ and a probability distribution ν on I. Let $(M_t)_{t\in I}$ be a real-valued closable martingale. Then for an random adapted probability measure $\gamma^* \in \mathcal{M}_I^\nu$ the following conditions are equivalent:*

(a) *γ^* is optimal for all processes $(Z_t)_{t\in I}$ given by*

$$Z_t = f(t)M_t, \quad t \in I, \tag{3.5.35}$$

with $f \in \mathcal{F}_\nu(M)$ (and f is bounded).

(b) *$\gamma^* \in \mathcal{M}_I^\nu$ satisfies*

$$\mathbb{E}[M_s\gamma^*_{>s}] \geq \mathbb{E}[M_s\gamma_{>s}] \quad \text{or} \quad \mathbb{E}[M_s\gamma^*_{\leq s}] \leq \mathbb{E}[M_s\gamma_{\leq s}].$$

for all $s \in I$ and $\gamma \in \mathcal{M}_I^\nu$.

Proof of Corollary 3.5.34. Due to Remark 3.5.19 and Remark 3.5.24 the proof follows from Theorem 3.5.6. We only have to show that

$$\mathbb{E}\left[\sum_{t\in I_{>s}} M_t\gamma_t\right] = \mathbb{E}[M_s\gamma_{>s}]$$

for every $s \in I$ and $\gamma \in \mathcal{M}_I^\nu$. Using that M is closable, we get with Jensen's inequality that

$$|M_t| = |\mathbb{E}[M_\infty|\mathcal{F}_t]| \leq \mathbb{E}[|M_\infty||\mathcal{F}_t].$$

Thus we have

$$\mathbb{E}\left[\sum_{t\in I_{>s}} |M_t|\gamma_t\right] = \mathbb{E}\left[\sum_{t\in I_{>s}} |\mathbb{E}[M_\infty|\mathcal{F}_t]|\gamma_t\right] \leq \mathbb{E}\left[\sum_{t\in I_{>s}} \mathbb{E}[|M_\infty||\mathcal{F}_t]\gamma_t\right] = \mathbb{E}\left[\sum_{t\in I_{>s}} \mathbb{E}[|M_\infty|\gamma_t|\mathcal{F}_t]\right].$$

The last equality results from the fact that γ_t is $\mathcal{F}_t$-measurable. Using monotone convergence it follows that

$$\mathbb{E}\left[\sum_{t\in I_{>s}} \mathbb{E}[|M_\infty|\gamma_t|\mathcal{F}_t]\right] = \sum_{t\in I_{>s}} \mathbb{E}[\mathbb{E}[|M_\infty|\gamma_t|\mathcal{F}_t]] = \sum_{t\in I_{>s}} \mathbb{E}[|M_\infty|\gamma_t] = \mathbb{E}[|M_\infty|\gamma_{>s}] < \infty$$

for every $s \in I$ and $\gamma \in \mathcal{M}_I^\nu$ such that $\sum_{t\in I_{>s}} M_t\gamma_t \in L^1$. Thus everything is well-defined and integrable. Repeating the calculation without absolute values, which is allowed due to the absolute convergence almost surely, we get that

$$\mathbb{E}\left[\sum_{t\in I_{>s}} M_t\gamma_t\right] = \mathbb{E}[M_\infty\gamma_{>s}].$$

Furthermore, we have for every $s \in I$ and $\gamma^*, \gamma \in \mathcal{M}_I^\nu$ that

$$\mathbb{E}[M_\infty\gamma^*_{>s}] \geq \mathbb{E}[M_\infty\gamma_{>s}],$$

$$\mathbb{E}[\mathbb{E}[M_\infty\gamma^*_{>s}|\mathcal{F}_s]] = \mathbb{E}[\mathbb{E}[M_\infty|\mathcal{F}_s]\gamma^*_{>s}] \geq \mathbb{E}[\mathbb{E}[M_\infty|\mathcal{F}_s]\gamma_{>s}] = \mathbb{E}[\mathbb{E}[M_\infty\gamma_{>s}|\mathcal{F}_s]],$$

$$\mathbb{E}[M_s\gamma^*_{>s}] \geq \mathbb{E}[M_s\gamma_{>s}].$$

Because of $\gamma_{>s} = 1 - \gamma_{\leq s}$, for every $\gamma \in \mathcal{M}_I$, it is like $\mathbb{E}[M_s\gamma^*_{\leq s}] \leq \mathbb{E}[M_s\gamma_{\leq s}]$. ❑

Remark 3.5.36. As the proof of Corollary 3.5.34 shows, it is also possible to use the condition $\mathbb{E}[M_\infty \gamma^*_{>s}] \geq \mathbb{E}[M_\infty \gamma_{>s}]$ instead of $\mathbb{E}[M_s \gamma^*_{>s}] \geq \mathbb{E}[M_s \gamma_{>s}]$.

The next corollary gives us an equivalently condition to the second one in Theorem 3.5.6 using the expected shortfall, see Section A.3.

Corollary 3.5.37. *Given a totally ordered countable index set $I \subseteq \overline{\mathbb{R}}$ and a probability distribution ν on I. If $\mathcal{F}_\nu(M)$ includes the identically one function and $\mathbb{E}[M_\gamma] = \mathbb{E}[M_t]$ for all $\gamma \in \mathcal{M}_I^\nu$ and $t \in I$, then an adapted random probability measure $\gamma^* \in \mathcal{M}_I^\nu$ is optimal for all $(Z_t)_{t\in I}$ given in (3.5.4), if for all $s \in I$*

$$\mathbb{E}\Bigg[\sum_{t\in I_{>s}} M_t \gamma_t^*\Bigg] = (1-\nu_{\leq s})\,\mathrm{ES}[M_s; \nu_{\leq s}]. \tag{3.5.38}$$

Proof. The optimality of γ^* follows by Theroem 3.5.6 such that we only need to show that the equation (3.5.38) holds. With the additional assumption that $\mathbb{E}[M_\gamma] = \mathbb{E}[M_s]$ for every $\gamma \in \mathcal{M}_I^\nu$ and $s \in I$ and we get

$$\mathbb{E}[M_s] = \mathbb{E}[M_\gamma] = \mathbb{E}\Bigg[\sum_{t\in I} M_t \gamma_t\Bigg].$$

Therefore

$$\begin{aligned}
\mathbb{E}\Bigg[\sum_{t\in I_{>s}} M_t \gamma_t\Bigg] &= \mathbb{E}[M_s] - \mathbb{E}\Bigg[\sum_{t\in I_{\leq s}} M_t \gamma_t\Bigg] = \mathbb{E}[M_s] - \sum_{t\in I_{\leq s}} \mathbb{E}[\mathbb{E}[M_s \gamma_t | \mathcal{F}_t]] \\
&= \mathbb{E}\Bigg[M_s\Bigg(1 - \sum_{t\in I_{\leq s}} \gamma_t\Bigg)\Bigg] = \mathbb{E}[M_s(1-\gamma_{\leq s})] \\
&\leq \sup_{1-\gamma_{\leq s}\in \mathcal{F}^1_{\nu_{\leq s}, M_s}} \mathbb{E}[M_s(1-\gamma_{\leq s})] \overset{\text{Lemma A.3.3}}{=} (1-\nu_{\leq s})\,\mathrm{ES}[M_s;\nu_{\leq s}] \\
&= \mathbb{E}\Bigg[\sum_{t\in I_{>s}} M_t \gamma_t^*\Bigg].
\end{aligned}$$

❑

Remark 3.5.39. With $\mathbb{E}[M_\gamma] = \mathbb{E}[M_s]$ for every $\gamma \in \mathcal{M}_I^\nu$ and $s \in I$ we have

$$\begin{aligned}
\mathbb{E}\Bigg[\sum_{t\in I_{>s}} M_t \gamma_t\Bigg] = \mathbb{E}[M_s(1-\gamma_{\leq s})] &\leq \sup_{1-\gamma_{\leq s}\in \mathcal{F}^1_{\nu_{\leq s}, M_s}} \mathbb{E}[M_s(1-\gamma_{\leq s})] \\
&\overset{\text{Lemma A.3.3}}{=} (1-\nu_{\leq s})\,\mathrm{ES}[M_s;\nu_{\leq s}].
\end{aligned}$$

In particular, it follows that

$$\mathbb{E}\Bigg[\sum_{t\in I_{>s}} M_t \gamma_t^*\Bigg] \leq (1-\nu_{\leq s})\,\mathrm{ES}[M_s;\nu_{\leq s}].$$

The following statements remain in similar form as in [33]:

Remark 3.5.40. Cf. [33, Remark 5.58]:
Equivalently a stopping time $\tau^* \in \mathcal{T}_I^\nu$ is optimal, if for all $t \in I$

$$\mathbb{E}[M_{\tau^*}\mathbb{1}_{\{\tau^*>t\}}] = (1-\mathbb{P}(\tau^* \leq t))\,\mathrm{ES}[M_t;\mathbb{P}(\tau^* \leq t)].$$

Remark 3.5.41. Cf. [33, Remark 5.59]:
In the setting of Corollary 3.5.37 a stopping time $\tau^* \in \mathcal{T}_I^\nu$ is optimal, if for all $t \in I$

$$\mathbb{E}[M_{\tau^*}\mathbb{1}_{\{\tau^*>t\}}] = \mathbb{E}[M_t\mathbb{1}_{\{\tau^*>t\}}] = \|M_t\|_\infty \mathbb{P}(\tau^* > t).$$

This implies that up to a null set $\{\tau^* > t\}$ is contained in $\{M_t = \|M_t\|_\infty\}$ for all $t \in I$, i.e., $\mathbb{P}(\{\tau^* > t\} \setminus \{M_t = \|M_t\|_\infty\}) = 0$ for all $t \in I$. This representation will be useful in the following sections.

Corollary 3.5.37 gives us an upper bound for our problem, which is not achieved in any case. Using a different approach, which is not directly connected to the one of Corollary 3.5.37, we want to deduce another upper bound, which will be achieved from an adapted random probability measure. For this we will use $\mathcal{F}_{\delta_s,M_s}^{1-\gamma_{<s}}$ which includes the past of γ.
Remember that

$$\mathcal{F}_{\delta,X}^{Y} := \{f:\Omega \to [0,1] \mid f \text{ measurable}, \mathbb{E}[fY] = \mathbb{E}[f_{\delta,X}Y]\}$$

for $\delta \in [0,1]$ and two real-valued random variables X and Y, which satisfy $Y \geq 0$, $\mathbb{E}[Y] < \infty$ and $\mathbb{E}[|X|] < \infty$. Then we have that $\mathbb{E}[f_{\delta,X}XY]$ is well-defined and

$$\sup_{f \in \mathcal{F}_{\delta,X}^{Y}} \mathbb{E}[fXY] = \mathbb{E}[f_{\delta,X}XY].$$

Given $s \in I$, we want to maximize $\mathbb{E}[M_s(1-\gamma_{\leq s})]$. There exists $\delta_s \in [0,1]$ such that

$$\mathbb{E}[f_{\delta_s,M_s}(1-\gamma_{<s})] = 1-\nu_{\leq s}.$$

Note that

$$\mathbb{E}[M_s(1-\gamma_{\leq s})] = \mathbb{E}[M_s(1-\gamma_{<s})f] \tag{3.5.42}$$

with

$$f := \begin{cases} 1-\frac{\gamma_s}{1-\gamma_{<s}} & \text{on } \{\gamma_{<s} < 1\}, \\ 0 & \text{otherwise,} \end{cases}$$

because

$$\mathbb{E}[M_s(1-\gamma_{<s})f] = \mathbb{E}\left[M_s\left(1-\gamma_{<s}-(1-\gamma_{<s})\frac{\gamma_s}{(1-\gamma_{<s})}\right)\right] = \mathbb{E}[M_s(1-\gamma_{\leq s})].$$

Furthermore, we get with the choice $X = M_s$ and $Y = 1-\gamma_{<s}$ that

$$\begin{aligned}\mathbb{E}[Yf] = \mathbb{E}[(1-\gamma_{<s})f] &= \mathbb{E}\left[1-\gamma_{<s}-(1-\gamma_{<s})\frac{\gamma_s}{(1-\gamma_{<s})}\right] \\ &= \mathbb{E}[1-\gamma_{\leq s}] = 1-\nu_{\leq s} = \mathbb{E}[f_{\delta_s,M_s}(1-\gamma_{<s})] = \mathbb{E}[Yf_{\delta_s,M_s}],\end{aligned}$$

such that it follows that $f \in \mathcal{F}^{1-\gamma_{<s}}_{\delta_s,M_s}$. Using Lemma A.3.3 we get

$$\mathbb{E}[M_s(1-\gamma_{\leq s})] = \mathbb{E}[M_s(1-\gamma_{<s})f] \leq \sup_{g\in\mathcal{F}^{1-\gamma_{<s}}_{\delta_s,M_s}} \mathbb{E}[M_s(1-\gamma_{<s})g] = \mathbb{E}[M_s(1-\gamma_{<s})f_{\delta_s,M_s}].$$

For every $\gamma \in \mathcal{M}^{\nu}_{I}$ we have the inequality

$$\mathbb{E}[M_s(1-\gamma_{\leq s})] \leq \mathbb{E}[M_s(1-\gamma_{<s})f_{\delta_s,M_s}]. \tag{3.5.43}$$

For $I = \mathbb{N}_0$ we want to construct the process $(\gamma_t)_{t\in I}$ recursively such that for every $s \in I$ it hold the equality in (3.5.43), i.e.,

$$\mathbb{E}[M_s(1-\gamma_{\leq s})] = \mathbb{E}[M_s(1-\gamma_{<s})f_{\delta_s,M_s}]. \tag{3.5.44}$$

Lemma 3.5.45. *For $I = \mathbb{N}_0$ we define*

$$\gamma_s := \begin{cases} 1 - f_{\delta_0,M_0} & \text{for } s = 0, \\ (1-f_{\delta_s,M_s})(1-\gamma_{<s}) & \text{for } s \in \mathbb{N}, \end{cases}$$

and $\gamma_{<0} = 0$. Then we get

$$\gamma_s = (1-f_{\delta_s,M_s})\prod_{t=0}^{s-1} f_{\delta_t,M_t}, \quad \forall s \in \mathbb{N}, \tag{3.5.46}$$

$$1-\gamma_{\leq s} = \prod_{t=0}^{s} f_{\delta_t,M_t}, \quad \forall s \in \mathbb{N}, \tag{3.5.47}$$

$$\mathbb{E}[M_s(1-\gamma_{\leq s})] = \mathbb{E}[M_s(1-\gamma_{<s})f_{\delta_s,M_s}] = \mathbb{E}\left[M_s\prod_{t=0}^{s} f_{\delta_t,M_t}\right], \quad \forall s \in \mathbb{N}_0. \tag{3.5.48}$$

Proof. 1.) We show the equations (3.5.46) and (3.5.47) by induction.

(a) Basis: Show that the statements hold for $s = 1$. We have that

$$\begin{aligned}\gamma_1 &= (1-f_{\delta_1,M_1})(1-\gamma_{<1}) = (1-f_{\delta_1,M_1})(1-\gamma_0) = (1-f_{\delta_1,M_1})f_{\delta_0,M_0} \\ &= (1-f_{\delta_1,M_1})\prod_{t=0}^{0} f_{\delta_t,M_t}\end{aligned}$$

and

$$\begin{aligned}1-\gamma_{\leq 1} &= 1-\gamma_1-\gamma_0 = 1-(1-f_{\delta_1,M_1})f_{\delta_0,M_0} - (1-f_{\delta_0,M_0}) \\ &= 1 - f_{\delta_0,M_0} + f_{\delta_1,M_1}f_{\delta_0,M_0} - (1-f_{\delta_0,M_0}) = f_{\delta_1,M_1}f_{\delta_0,M_0} = \prod_{t=0}^{1} f_{\delta_t,M_t}.\end{aligned}$$

(b) Inductive step: Show that if the equation holds for γ_s , then it also holds for γ_{s+1}. This can be done as follows:

$$\gamma_{s+1} = (1-f_{\delta_s,M_s})(1-\gamma_{<s+1}) = (1-f_{\delta_s,M_s})(1-\gamma_{\leq s}) = (1-f_{\delta_s,M_s})\prod_{t=0}^{s} f_{\delta_t,M_t}$$

and

$$1-\gamma_{\le s+1} = 1-\gamma_{<s+1}-\gamma_{s+1} = (1-\gamma_{\le s})-\gamma_{s+1}$$
$$= \prod_{t=0}^{s} f_{\delta_t,M_t} - (1-f_{\delta_s,M_s})\prod_{t=0}^{s} f_{\delta_t,M_t} = \prod_{t=0}^{s+1} f_{\delta_t,M_t}.$$

Thus the statements are true for γ_{s+1}.

2.) We show by induction that $\mathbb{E}[M_s(1-\gamma_{\le s})] = \mathbb{E}[M_s(1-\gamma_{<s})f_{\delta_s,M_s}]$ for all $s \in \mathbb{N}_0$, i.e., the equation (3.5.48).

(a) Basis: Show that the statement holds for $s=0$. For the left-hand side of the equation we have

$$\mathbb{E}[M_0(1-\gamma_{\le 0})] = \mathbb{E}[M_0(1-\gamma_0)] = \mathbb{E}[M_0 f_{\delta_0,M_0}].$$

For the right-hand side of the equation we get

$$\mathbb{E}[M_0(1-\gamma_{<0})f_{\delta_0,M_0}] = \mathbb{E}[M_0(1-0)f_{\delta_0,M_0}] = \mathbb{E}[M_0 f_{\delta_0,M_0}].$$

The two sides are equal, so the statement is true for $s=0$.

(b) Inductive step: Show that if the equation holds for s, then also holds for $s+1$. This can be done as follows. For the left-hand side of the equation we have

$$\mathbb{E}[M_{s+1}(1-\gamma_{\le s+1})] = \mathbb{E}\left[M_{s+1}\prod_{t=0}^{s+1} f_{\delta_t,M_t}\right].$$

For the right-hand side of the equation we get

$$\mathbb{E}[M_{s+1}(1-\gamma_{<s+1})f_{\delta_{s+1},M_{s+1}}] = \mathbb{E}[M_{s+1}(1-\gamma_{\le s})f_{\delta_{s+1},M_{s+1}}]$$
$$= \mathbb{E}\left[M_{s+1}\cdot\prod_{t=0}^{s} f_{\delta_t,M_t}\cdot f_{\delta_{s+1},M_{s+1}}\right] = \mathbb{E}\left[M_{s+1}\prod_{t=0}^{s+1} f_{\delta_t,M_t}\right].$$

The two sides are equal, so the statement is true for $s+1$. Thus it has been shown that equality holds in (3.5.43) for this special γ. ❑

In this case we maximize the value $\mathbb{E}[M_s(1-\gamma_{\le s})] = \mathbb{E}[M_s(1-\gamma_{<s})f]$ over the set $\in \mathcal{F}_{\delta_s,M_s}^{1-\gamma_{<s}}$ which bear the past of γ in mind.

3.5.3. The Binomial Model

The results of [33, Section 5.8] can be corrected, generalized and extended in the following way.
For $I = \{0\} \cup J$ with $J \subseteq \mathbb{N}$ and $n = |J|$ (number of elements of J), let $X = (X_t)_{t \in J}$ be an independent process of $\{0,1\}$-valued random variables and let the filtration be given by

$$\mathcal{F}_t = \begin{cases} \{\emptyset, \Omega\} & \text{for } t = 0, \\ \sigma(X_s \mid s \in J_{\le t}) & \text{for } t \in J, \end{cases}$$

where $J_{\le t} := \{s \in J \mid s \le t\}$ and $J_{<t} := \{s \in J \mid s < t\}$. We set $p_t = \mathbb{P}(X_t = 1)$ for all $t \in J$. Let Z be the underlying price process or a process which describes a special payoff. This process Z is modeled by $Z_0 > 0$ and

$$Z_t = Z_0 \prod_{s \in J_{\le t}} u_s^{X_s} d_s^{1-X_s}, \quad t \in J, \tag{3.5.49}$$

with $u_t > 1 \ge d_t > 0$. In this model the price process Z could be recursively rewritten as

$$Z_t = Z_{t-1} u_t^{X_t} d_t^{1-X_t}, \quad t \in J,$$

such that the increments of the process are given by

$$\Delta Z_t := Z_t - Z_{t-1} = Z_{t-1}(u_t^{X_t} d_t^{1-X_t} - 1), \quad t \in J.$$

In addition, depending on the choice of parameters, the following applies

Lemma 3.5.50. *In the setting of the binomial model with Z given in (3.5.49), we get that*

(a) Z is a submartingale iff $p_t \ge \frac{1-d_t}{u_t-d_t}$ for all $t \in J$.

(b) Z is a supermartingale iff $p_t \le \frac{1-d_t}{u_t-d_t}$ for all $t \in J$.

(c) Z is a martingale iff $p_t = \frac{1-d_t}{u_t-d_t}$ for all $t \in J$.

Proof. We only show the first case, because the others follow analogously. The process Z is a submartingale if and only if $\mathbb{E}[Z_t] \ge \mathbb{E}[Z_{t-1}]$ for all $t \in J$. Furthermore it holds that

$$\mathbb{E}[Z_t] = (u_t p_t + d_t(1-p_t))\mathbb{E}[Z_{t-1}].$$

Therefore we get for all $t \in J$ that

$$\begin{aligned} \mathbb{E}[Z_t] &\ge \mathbb{E}[Z_{t-1}] \\ \mathbb{E}[Z_{t-1}](u_t p_t + d_t(1-p_t)) &\ge \mathbb{E}[Z_{t-1}] \\ u_t p_t + d_t(1-p_t) &\ge 1 \\ (u_t - d_t)p_t &\ge 1 - d_t \\ p_t &\ge \frac{1-d_t}{u_t-d_t}. \end{aligned}$$

❑

Let the Doob decomposition of Z be given by $Z = M + A$ with a martingale M and a predictable process A. Then the increments of the predictable process A are given by

$$\Delta A_t := \mathbb{E}[\Delta Z_t | \mathcal{F}_{t-1}] = Z_{t-1}\Big(\mathbb{E}\Big[u_t^{X_t} d_t^{1-X_t} \,|\, \mathcal{F}_{t-1}\Big] - 1\Big) = Z_{t-1}(u_t p_t + (1-p_t)d_t - 1), \quad t \in J.$$

Remark 3.5.51 (Special cases).

- Let be $Z_0 > 0$, $u_t \equiv u$, $d_t \equiv d$ and $p_t \equiv p \in (0,1)$ for all $t \in J$. The assumption that $p_t \equiv p$ for all $t \in J$ implies that $X = (X_t)_{t\in J}$ is an independent process of identically distributed $\{0,1\}$-valued random variables. Then the process Z is given by
$$Z_t = Z_0 u^{N_t} d^{n_t - N_t}, \quad t \in J, \tag{3.5.52}$$
with $u > 1 \geq d > 0$, $n_t = |J_{\leq t}|$ and $N_t = \sum_{s \in J_{\leq t}} X_s$.
- Let be $I = \{0, \dots, T\}$, then $n_t = t$ for all $t \in J$ and we have that
$$\mathbb{E}[Z_t] = \mathbb{E}[Z_{t-1}](up + d(1-p)) = \mathbb{E}[Z_0](up + d(1-p))^t = Z_0(up + d(1-p))^t.$$
- If $u_t \equiv u$, $d_t \equiv d$ for all $t \in J$, the cases $1 \geq u > d > 0$ and $u > d \geq 1$ are out of our interest, because it is clear when Z must be stopped. In the first case we stop the process immediately, because Z is (strictly) monotone decreasing. In the second case Z is (strictly) monotone increasing and we stop at the maturity T. This also applies for a single time step.

Now, we want to pick up the considerations in [33, Section 5.8], correct and expand them. Therefore we consider a binomial model on $I = \{0, \dots, T\}$ with given constants $Z_0 > 0$, $u_t \equiv u$, $d_t \equiv d$ with $u > 1 \geq d > 0$ and $p_t \equiv p \in (0,1)$ for all $t \in I \setminus \{0\} =: J$.
In our reflections, we refer back to stopping times. We assume that $\mathcal{T}_I^\nu \neq \emptyset$ and an optimal stopping time exists and is denote by τ^*. Then we have that

$$V_{\mathcal{T}}^\nu(Z) = \sup_{\tau \in \mathcal{T}_I^\nu} \mathbb{E}[Z_\tau] = \mathbb{E}[Z_{\tau^*}] = \sum_{t \in I} \mathbb{E}[Z_t \mathbb{1}_{\{\tau^* = t\}}].$$

If the process Z and a stopping time τ or an adapted random probability measure γ are independent, we have that

$$V_{\text{ind}}^\nu(Z) = \sum_{t \in I} \mathbb{E}[Z_t]\nu_t.$$

As in [33] discussed, $V_{\text{ind}}^\nu(Z) \leq V_{\mathcal{T}}^\nu(Z)$ and if Z is a submartingale, we get that $V_{\mathcal{T}}^\nu(Z) \leq \mathbb{E}[Z_T]$ by [33, Lemma 4.37]. We want to consider the following proposition from [33].

Proposition 3.5.53. *Cf. [33, Proposition 5.61]:*
In the setting for the binomial model stated above we now assume $I = \{0, \dots, T\}$. Further we have to assume that the distribution ν is given by

$$\nu_t = \begin{cases} 0 & \text{if } t = 0, \\ \frac{1}{2^t} & \text{if } t \in \{1, \dots, T-1\}, \\ \frac{1}{2^{T-1}} & \text{if } t = T. \end{cases}$$

Then the optimal stopping time is given by

$$\tau^* = T \wedge \min\{t \in I \mid Z_t < Z_{t-1}\}.$$

This proposition is for the special choice $p = \frac{1}{2}$ which is not explicitly specified and it generates the so called "symmetric" case. Furthermore it is illustrated with [33, Example 5.62]. This proposition fails for each choice of u and d such that $u + d \leq 2$.

Example 3.5.54. Counterexamples for [33, Proposition 5.61]:
We will now consider a binomial model with given constants $Z_0 = 1$, u, d with $u > 1 \geq d > 0$ and $p = \frac{1}{2}$. We assume that the distribution ν is given as in Proposition 3.5.53. For each choice of u and d such that $u+d \leq 2$ we get a counterexample for this proposition. Further let $I = \{0,\dots,5\}$.

(u,d)	$(\frac{3}{2},\frac{1}{4})$	$(2,\frac{1}{2})$	$(2,\frac{1}{5})$
$V^{\nu}_{\text{ind}}(Z)$	0.7813	1.6031	1.2110
$V^{\nu}_{T}(Z)$	0.6187	2.25	1.5
$\mathbb{E}[Z_5]$	0.5129	3.0518	1.6105

To get the inequality $V^{\nu}_{T}(Z) \leq \mathbb{E}[Z_T]$, we need that Z is an submartingale. By Lemma 3.5.50, we get that Z is a submartingale if $p \geq \frac{1-d}{u-d}$, i.e., for $p = 1/2$ if $u + d \geq 2$. In the case that Z is a supermartingale, iff $p \leq \frac{1-d}{u-d}$, we have

$$\mathbb{E}[Z_T] = \mathbb{E}[Z_T]\cdot 1 = \mathbb{E}[Z_T]\sum_{t\in I}\nu_t = \sum_{t\in I}\mathbb{E}[Z_T]\nu_t \leq \sum_{t\in I}\mathbb{E}[Z_t]\nu_t = V^{\nu}_{\text{ind}}(Z),$$

because for any supermartingale Z it holds that $\mathbb{E}[Z_t] \leq \mathbb{E}[Z_{t-1}]$ for all $t \in I$ and in particular, $\mathbb{E}[Z_T] \leq \mathbb{E}[Z_t]$ for all $t \in I$. The equality holds only for martingales, i.e., in the case $p = \frac{1-d}{u-d}$. In general we get the following proposition:

Proposition 3.5.55. *In the setting of Proposition 3.5.53 with $I = \{0,\dots,T\}$ and $p \geq \frac{1-d}{u-d}$ we have that $\mathbb{E}[Z_{\tau^*}] \leq \mathbb{E}[Z_T]$.*

Proof. We know by Lemma 3.5.50 that the process Z is a submartingale for $p \geq \frac{1-d}{u-d}$. Using Lemma 3.4.5 we get for every $\gamma \in \mathcal{M}^{\nu}_{I}$ that $\mathbb{E}[Z_\gamma] \leq \mathbb{E}[Z_T]$. It is obvious that the stopping time τ^* given in Proposition 3.5.56 satisfies $\mathcal{L}(\tau^*) = \nu$. Furthermore, τ^* can be identified with $\gamma \in \mathcal{M}^{\nu}_{I}$ given by $\gamma_t(\omega) = \mathbb{1}_{\tau^*(\omega)}(t)$ for all $t \in J$. ❑

We consider an restricted optimization problem. The stopping time or adapted random probability measure follows a given distribution and can depend on the underlying process of payoff. We are concretely interested in the deduction of estimations for the worst-case scenario for this problem, that means, the supremum over the expected payoffs. The stopping time $\tau = T \wedge \min\{s \in I \mid Z_s < Z_{s-1}\}$ might be a good candidate for the optimal strategy, because it is the mathematical description of the greedy strategy. We now want to clarify the question for which problems this special strategy is optimal. This strategy τ stops our process Z in the following way:

$$\begin{array}{ll}
\tau = 0 & Z_\tau = Z_0 \\
\tau = t \text{ with } t \in \{1,\dots,T-1\} & Z_\tau = Z_0 \cdot u^{t-1}\cdot d \\
\tau = T & Z_\tau = \begin{cases} Z_0 \cdot u^T & \text{with probability } p, \\ Z_0\cdot u^{T-1}\cdot d & \text{with probability } 1-p. \end{cases}
\end{array}$$

The distribution of the stopping time τ is given by

$$\mathbb{P}(\tau = t) = \begin{cases} 0, & \text{if } t = 0, \\ p^{t-1}(1-p), & \text{if } t \in \{1, \ldots, T-1\}, \\ p^{T-1}, & \text{if } t = T. \end{cases}$$

Thus we get

$$\begin{aligned} \mathbb{E}[Z_\tau] &= \sum_{t \in I} \mathbb{E}[Z_t \mathbb{1}_{\{\tau = t\}}] \\ &= Z_0 \cdot \nu_0 + Z_0 d \cdot \nu_1 + Z_0 u d \cdot \nu_2 + \ldots + (Z_0 u^{T-1} d(1-p) + Z_0 u^T p) \cdot \nu_T. \end{aligned}$$

If we assume that the process Z and the stopping time τ are independent, we get

$$V^{\nu}_{\text{ind}}(Z) = \sum_{t \in I} \mathbb{E}[Z_t] \nu_t = \sum_{t \in I} Z_0 (up + d(1-p))^t \nu_t.$$

The following Proposition gives us a generalized and corrected version of [33, Proposition 5.61].

Proposition 3.5.56. *In the setting for the binomial model stated above we now assume $I = \{0, \ldots, T\}$. Let be $p \geq \frac{1-d}{u-d}$ and the distribution ν is given by*

$$\nu_t = \begin{cases} 0, & t = 0, \\ p^{t-1}(1-p), & \text{if } t \in \{1, \ldots, T-1\}, \\ p^{T-1}, & \text{if } t = T. \end{cases}$$

Then the optimal stopping time is given by

$$\tau^* = T \wedge \min\{t \in I \,|\, Z_t < Z_{t-1}\}.$$

Remark 3.5.57. Note that the optimal stopping time in Proposition 3.5.56 has the distribution ν, i.e., $\mathcal{L}(\tau^*) = \nu$.

Proof of Proposition 3.5.56. The result of this proposition follows from Corollary 3.5.37. For this we need to represent Z as $Z_t = f(t) M_t$ for $t \in I$ with a martingale $(M_t)_{t \in I}$ and a non-decreasing deterministic function $f : I \to \mathbb{R}$ which satisfy certain conditions. We set $f(0) := 1$ and $M_0 := Z_0$ as well as for each $t \in J$

$$f(t) := \prod_{s=1}^{t} \mathbb{E}[u^{X_s} d^{1-X_s}]$$

and

$$M_t := Z_0 \prod_{s=1}^{t} \frac{u^{X_s} d^{1-X_s}}{\mathbb{E}[u^{X_s} d^{1-X_s}]}.$$

The inequality $p \geq \frac{1-d}{u-d}$ is equivalent to $up + (1-p)d \geq 1$. Therefore it is immediately clear that f is a non-decreasing deterministic function. Furthermore using that X_s, $s \in J$, are independent and Z_0 is $\mathcal{F}_0$-measurable, we have for every $t \in I$ that

$$\mathbb{E}[M_t] = \mathbb{E}\left[Z_0 \prod_{s=1}^{t} \frac{u^{X_s} d^{1-X_s}}{\mathbb{E}[u^{X_s} d^{1-X_s}]} \right] = Z_0 \prod_{s=1}^{t} \frac{\mathbb{E}[u^{X_s} d^{1-X_s}]}{\mathbb{E}[u^{X_s} d^{1-X_s}]} = Z_0,$$

such that $M = (M_t)_{t \in I}$ is a martingale and M is non-negative because $Z_0 \geq 0$ and $u, d > 0$. Furthermore we know that for every $t \in I$

$$|M_t| = M_t \leq Z_0 \frac{u^T}{\mathbb{E}[u^{X_s} d^{1-X_s}]} \quad \text{and} \quad |Z_t| = Z_t \leq Z_0 u^T$$

such that we have $\mathbb{E}[\sup_{t \in I} |M_t|] < \infty$ and $\mathbb{E}[\sup_{t \in I} |Z_t|] < \infty$. Due to the structure of the stopping time, for all $t \in I$ we have

$$\{\tau^* > t\} = \{M_t = \|M_t\|_\infty\} = \{Z_t = \|Z_t\|_\infty\},$$

which implies optimality of τ^* by Remark 3.5.41. ❑

Example 3.5.58 (Examples for Proposition 3.5.56)**.** We will now consider a binomial model with given constants $Z_0 = 1$, u, d with $u > 1 \geq d > 0$ and $p \in [0, 1]$. Let $I = \{0, \dots, 5\}$. Further we assume that the distribution ν is given as in Proposition 3.5.56. As in [33] discussed, it should hold that $V^{\text{ind}}(\nu) \leq V_T^\nu(Z)$ and if Z is a submartingale, we get that $V_T^\nu(Z) \leq \mathbb{E}[Z_T]$ by [33, Lemma 4.37].

(a) $Z_0 = 1, u = 2, d = \frac{1}{2}$

p	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{3}{4}$
$V_{\text{ind}}^\nu(Z)$	0.7143	0.8401	1	1.6031	3.5	5.8254
$V_T^\nu(Z)$	0.6265	0.7578	1	2.25	5.8210	9.2422
$\mathbb{E}[Z_5]$	0.2373	0.51291	1	3.0518	7.5938	11.3310

$\frac{1-d}{u-d} = \frac{1}{3}$

(b) $Z_0 = 1, u = 3, d = \frac{1}{3}$

p	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{3}{4}$
$V_{\text{ind}}^\nu(Z)$	0.7447	1	1.3708	3.3925	13.3328	28.4010
$V_T^\nu(Z)$	0.5694	1	2.1111	9.7917	35.4444	61.4427
$\mathbb{E}[Z_5]$	0.2846	1	2.7274	12.8601	41.9330	69.1646

$\frac{1-d}{u-d} = \frac{1}{4}$

(c) $Z_0 = 1, u = 2, d = \frac{1}{5}$

p	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{3}{4}$
$V_{\text{ind}}^\nu(Z)$	0.4545	0.5821	0.7276	1.2110	2.7505	4.8012
$V_T^\nu(Z)$	0.2531	0.3219	0.4790	1.5	4.8568	8.2531
$\mathbb{E}[Z_5]$	0.0312	0.1160	0.3277	1.6105	5.3782	8.9466

$\frac{1-d}{u-d} = \frac{4}{9}$

The following proposition gives an condition for ν such that we get a better lower bound for the value of the expected payoff $V_{\mathcal{T}}^{\nu}(Z)$ by using the greedy strategy.

Proposition 3.5.59. *In the setting for the binomial model stated above we now assume* $I = \{0,\dots,T\}$ *and* $p \geq \frac{1-d}{u-d}$. *If the distribution* ν *satisfies the following condition:*

$$\sum_{t=1}^{T-1} (u^{t-1}d - (up + d(1-p))^t)\nu_t \geq ((up + d(1-p))^T - u^{T-1}(up + d(1-p)))\nu_T,$$

we get that $V_{ind}^{\nu}(Z) \leq \mathbb{E}[Z_{\tau^*}] \leq V_{\mathcal{T}}^{\nu}(Z)$ *for* $\tau^* = T \wedge \min\{t \in I \,|\, Z_t < Z_{t-1}\}$.

Proof. It follows by

$$\sum_{t=1}^{T-1} (u^{t-1}d - (up + d(1-p))^t)\nu_t \geq ((up + d(1-p))^T - u^{T-1}(up + d(1-p)))\nu_T$$

with $e := up + d(1-p)$ that

$$\begin{aligned}
0 &\geq \sum_{t=1}^{T-1} (e^t - u^{t-1}d)\nu_t + (e^T - u^{T-1}e)\nu_T \\
0 &\geq \sum_{t=1}^{T-1} (Z_0 e^t - Z_0 u^{t-1}d)\nu_t + (Z_0 e^T - Z_0 u^{T-1}e)\nu_T \\
0 &\geq \sum_{t=0}^{T-1} Z_0 e^t \nu_t + Z_0 e^T \nu_T - (Z_0\nu_0 + \sum_{t=1}^{T-1} Z_0 u^{t-1} d\nu_t + Z_0 u^{T-1} e\nu_T) \\
0 &\geq \sum_{t=0}^{T} (\mathbb{E}[Z_t] - \mathbb{E}[Z_t \mathbb{1}_{\{\tau^* = t\}}])\nu_t \\
0 &\geq V_{\text{ind}}^{\nu}(Z) - \mathbb{E}[Z_{\tau^*}].
\end{aligned}$$

In the martingale case, i.e., $up + d(1-p) = 1$, the assumption is given by

$$\sum_{t=1}^{T-1} (u^{t-1}d - 1)\nu_t \geq (1 - u^{T-1})\nu_T.$$

It is clear that $V_{\mathcal{T}}^{\nu}(Z) = \sup_{\tau \in \mathcal{T}_I^{\nu}} \mathbb{E}[Z_\tau] \geq \mathbb{E}[Z_{\tau^*}]$, because τ^* is an element of $\mathcal{T}_I^{\nu}$. ❑

In the most general setting for the binomial model the strategy τ stops our process Z, which is given by (3.5.52), in the following way

$$\begin{array}{ll}
\tau = 0 & Z_\tau = Z_0 \\
\tau = t \text{ with } t \in I \setminus \{0, T\} & Z_\tau = Z_0 \cdot \left(\prod_{s \in J_{<t}} u_s \right) \cdot d \\
\tau = T & Z_\tau = \begin{cases} Z_0 \cdot \left(\prod_{s \in J_{<t}} u_s \right) & \text{with probability } p, \\ Z_0 \cdot \left(\prod_{s \in J_{<t}} u_s \right) \cdot d & \text{with probability } 1-p. \end{cases}
\end{array}$$

The distribution of the stopping time τ is given by

$$\mathbb{P}(\tau = t) = \begin{cases} 0 & \text{if } t = 0, \\ (1-p_t) \cdot \prod_{s \in J_{<t}} p_s & \text{if } t \in I \setminus \{0, T\}, \\ \prod_{s \in J_{<t}} p_s & \text{if } t = T. \end{cases}$$

We obtain the following proposition:

Proposition 3.5.60. *In the setting for the general binomial model stated above we now assume $I = \{0\} \cup J \subseteq \mathbb{N}_0$, J is finite, $T = \max(J)$ and $p_t \geq \frac{1-d_t}{u_t - d_t}$ for all $t \in J$. Further we have to assume that the distribution ν is given by*

$$\nu_t = \begin{cases} 0 & \textit{if } t = 0, \\ (1-p_t) \cdot \prod_{s \in J_{<t}} p_s & \textit{if } t \in I \setminus \{0, T\}, \\ \prod_{s \in J_{<t}} p_s & \textit{if } t = T. \end{cases}$$

Then the optimal stopping time is given by

$$\tau^* = T \wedge \min\{t \in I \mid Z_t < Z_{t-1}\}.$$

Proof. The proof works similar to that of Proposition 3.5.56. ❑

Analogous to the Proposition 3.5.59 we get

Proposition 3.5.61. *In the setting for the binomial model stated above we now assume $I = \{0, \dots, T\}$, $u_t > 1 \geq d_t \geq 0$ and $p_t \in [0,1]$ with $d_t(1-p_t) + u_t p_t \geq 1$ for all $t \in I \setminus \{0\}$. Then $Z = (Z_t)_{t \in I}$ given in (3.5.52) is a submartingale with $Z_0 \geq 0$.*
Let ν be a distribution on I with

$$\nu_t \geq \nu_{\geq t}(1-p_t), \quad \forall t \in J.$$

Then each τ^* with $\mathcal{L}(\tau^*) = \nu$ and

$$\{\tau^* = t\} \supset \{\tau^* \leq t, Z_t < Z_{t-1}\}, \quad \forall t \in J$$

is optimal.

Special Payout: Call-Option Now, a short example of special payouts will be given. The payout, which we will denote by $Z = (Z_t)_{t\in I}$, at each time point $t \in I$ is given by

$$Z_t = (S_t - K)^+$$

with strike $K \in \mathbb{R}$, the underlying price process $S = (S_t)_{t\in I}$ and $x^+ := \max\{0, x\}$. We assume that the process Z is given in such a way that the processes M and A of the Doob decomposition satisfy the necessary conditions. By construction, the process Z is a submartingale, such that $Z_t = M_t + A_t$ for each $t \in I$ with a martingale M and a predictable, increasing process $A = (A_t)_{t\in I}$ starting at $A_0 = 0$. Furthermore we have that

$$\mathbb{E}[Z_\tau] = \mathbb{E}[M_\tau] + \mathbb{E}[A_\tau] = \mathbb{E}[M_0] + \sum_{t\in I} \mathbb{E}[A_t \mathbb{1}_{\tau=t}] \quad \text{and} \quad \mathbb{E}[Z_\gamma] = \mathbb{E}[M_\gamma] + \mathbb{E}[A_\gamma].$$

Now, we consider the distribution-constrained optimization problem. For that we know that $\mathbb{E}[M_\gamma]$ is constant for all $\gamma \in \mathcal{M}_I^\nu$ and $\mathbb{E}[A_\gamma] = \sum_{t\in I} \mathbb{E}[\Delta A_t \gamma_{\geq t}]$ with the increments $\Delta A_t = A_t - A_{t-1}$. In the binomial model these increments are given by

$$\begin{aligned}
\Delta A_t = \mathbb{E}[Z_t \mid \mathcal{F}_{t-1}] - Z_{t-1} &= \mathbb{E}[(S_t - K)^+ | \mathcal{F}_{t-1}] - (S_{t-1} - K)^+ \\
&= S_{t-1}\, \mathbb{E}\left[\left(u^{X_t} d^{1-X_t} - \frac{K}{S_{t-1}} \right)^+ \,\middle|\, \mathcal{F}_{t-1} \right] - (S_{t-1} - K)^+ \\
&= S_{t-1}\, \mathbb{E}\left[\left(u^{X_t} d^{1-X_t} - \frac{K}{S_{t-1}} \right)^+ \right] - (S_{t-1} - K)^+.
\end{aligned}$$

The last equality follows by Proposition A.2.7. We can use this formulation to determine a strategy to yield the worst case scenario. If we have the simulated paths for Z, we will stop at every time $t \in I$ the part of the simulated paths with the smallest values determined by $\Delta A_{t+1} = \mathbb{E}[Z_{t+1}|\mathcal{F}_t] - Z_t$. The share in stopped paths depends on the adapted random probability measure γ. Similar can be obtained for the put options. We will use this in the next section.

3.6. Examples for Applications in Actuarial Mathematics

In the introduction we already noted that there are several applications of the problem in financial and actuarial mathematics. In this section we discuss applications of the problem in actuarial mathematics, especially unit-linked life insurances with guarantee. We reproduce examples from [33, Chapter 9] more detailed and add additional ones. The considerations are intentionally very detailed.

Imagine that you are a unit-linked life insurer and want to insure a married couple. Before we can do this, we consider the insurance for one person. For this we survey the discrete time interval $I = \{0, \dots, T\}$ with $T \in \mathbb{N}$. Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in I}, \mathbb{P})$ be a filtered probability space. Furthermore let $S = (S_t)_{t\in I}$ be the stock price process at the financial market, $G = (G_t)_{t\in I}$ the deterministic process of the guarantee and $Z = (Z_t)_{t\in I} \in \mathcal{L}^1(\mathbb{P})$ be a real-valued adapted process or special payout. The process of the payouts $(Z_t)_{t\in I}$ is defined by

$$Z_t = \max\{G_t, S_t\} = S_t + (G_t - S_t)^+, \quad \text{for all } t \in I.$$

The family of the insured person will get the insurance benefit Z_τ at an random time point τ after paying advance premiums. In this case the random time point τ is an stopping time and the insurance companies are interested in the expected value $\mathbb{E}[Z_\tau]$ of the payouts of the contract. Thus we would consider an classical optimal stopping problem with value $V_T(Z)$. Furthermore τ is modeled as the minimum of the maturity T and the future lifetime T_x of the insured person, where x indicate the age at conclusion of contract. Thus this stopping time has the distribution which is given through the life table or through the termination of a contract.

In order to determine the distribution ν of the stopping time τ we need to use a life table. The probability that a x-year old person will survive the next n years is denoted by ${}_np_x$. Conversely, we denote by ${}_nq_x = 1 - {}_np_x$ the probability that a x-year old person will die within the next n years. For $n = 1$ we write p_x and q_x. At first we set $\nu_0 = 0$. This is a reasonable assumption, since there will not be a payoff at the initiation time of the contract and the person to be insured is alive. For modeling one single unit-linked life insurance contract with payoff at the end of the year of death of the insured or at the end of the contract, we set ν_t equal to the probability that a x-year old person survives $t-1$ years and dies within the t-th year for $t \in \{1, \dots, T-1\}$. Finally, we have that

$$\nu_t = \begin{cases} 0 & \text{for } t = 0, \\ {}_{t-1}p_x \cdot q_{x+t-1} & \text{for } t \in \{1, \dots, T-1\}, \\ {}_{T-1}p_x & \text{for } t = T. \end{cases} \tag{3.6.1}$$

Importantly, we have to choose $\nu_T = {}_{T-1}p_x$ to have $\sum_{t=0}^{T} \nu_t = 1$. With the given distribution ν, we consider then a distribution-constrained optimal stopping problem OptStop$^\tau$ with value $V_T^\nu(Z)$. It is also possible to replace the stopping time by an adapted random probability measure and consider OptStop$^\gamma$.

As already described, in practice it is usually assumed that financial risks and biometric risks are independent. Using these assumption it follows that

$$\mathbb{E}[Z_\tau] = \sum_{t=0}^{T} \mathbb{E}[\mathbb{1}_{\{\tau=t\}} Z_t] \overset{\text{indep.}}{=} \sum_{t=0}^{T} \mathbb{P}(\tau = t)\mathbb{E}[Z_t],$$

where $\mathbb{P}(\tau = t) = \nu_t$ is determined by the life table and $\mathbb{E}[Z_t]$ by models from mathematical finance. We have denoted that value by $V^{\nu}_{\text{ind}}(Z)$. In the introduction it is already presented that the assumption of independence is questionable and that the approaches of models with dependence are searched. Since $(Z_t)_{t\in I}$ depends on the financial market and there is an restriction of distribution for the stopping time τ, the problem equates to the problem of one-time optimal stopping under distribution restriction described in Chapter 2.
Let us go back to the insurance of a married couple. A description of how the distribution of the stopping time can be found using a life table is given in (3.6.1). In addition to the assumption of independence of financial risks and biometric risks, in this case the independence of the physical and emotional health of the partners from each other is often assumed. Then we calculate the expected values for each individual person by means of the above considerations and add them together. But with some common sense, it is clear that this is not the case. The couple lives in the same environment and is strongly connected. For example, both can get injured in a possible car accident. Furthermore the broken heart syndrome is also known and studied since a long time in medicine, see [26]. They found that the mortality rate of bereaved close relatives is much greater within a year of bereavement compared with a control group. As a consequence, health can drastically deteriorate when one's partner dies. Therefore, it is not reasonable to assume independence of the times of death of either partner.
Moreover, using the adapted random probability measure a portfolio of similar contracts can be modeled. If a stopping time τ is used for modeling a life insurance contract for one person, then the adapted random probability measure can be used to model a married couple. Our married couple consists of person A and person B with the corresponding stopping times τ_A and τ_B. Then the group can be modeled using Remark 3.2.2 and the adapted random probability measure γ given by

$$\gamma_t(\omega) = w_A \mathbb{1}_{\{\tau_A(\omega)\}}(t) + w_B \mathbb{1}_{\{\tau_B(\omega)\}}(t) \qquad \text{for } t \in I,\, \omega \in \Omega$$

with non-negative weights w_A and w_B defining a probability measure on $\{A, B\}$. Note that τ_j in $\mathcal{T}^{\nu}_I$ for every $j \in \{A, B\}$ implies $\gamma \in \mathcal{M}^{\nu}_I$. If one prefers to model a whole portfolio of $N \in \mathbb{N}$ homogeneous contracts using adapted random probability measures, one can choose $\mathbb{E}[\gamma_t] = \nu_t$ with ν_t defined as before for $t \in I$. Assume that the process Z models the evolution of the underlying fund for one single contract. Then $X = N \cdot Z$ models the evolution of the entire portfolio.
Health insurance contracts are often modeled similar to life insurance contracts. This implies that the problem could also be used for modeling a health insurance contract. Especially the claims amount per risk could be modeled stochastically.
We will illustrate the computation of the values $V_{\mathcal{T}}(Z)$, $V^{\nu}_{\mathcal{T}}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ for a unit-linked life insurance with and without guarantee. We use [33, Lemma 5.37] and [33, Theorem 5.25]. In order to derive the values for the distribution ν we will use the values q_x given in the Austrian annuity table 2005, which was presented in [37].

3.6.1. Example without Guarantee

First, we look at examples without guarantee, such that we have $G_t = 0$ and $Z_t = S_t$ for all $t \in I$. The following example corresponds to [33, Example 9.2].

Example 3.6.2 (Uniform distribution). In order to be able to use results from [33] we need to assume that the process Z consists of independent random variables. We will assume that Z is an i.i.d. process with $Z_t \sim U(a,b)$ for all $t \in I$. Let the distribution ν be given from a life table as explained above, see (3.6.1). First of all, we will compute the quantiles and expected shortfalls needed. For a random variable $X \sim U(a,b)$ the δ-quantile of X is given by

$$q_\delta(X) = \delta(b-a) + a.$$

Then the expected shortfall of X for a given $\delta \in (0,1)$ is given by

$$\begin{aligned}\mathrm{ES}[X;\delta] &= \frac{1}{\mathbb{P}(X > q_\delta(X))}\mathbb{E}[X \mathbb{1}_{\{X > q_\delta(X)\}}] = \frac{1}{1-\delta}\int_{q_\delta(X)}^{b} \frac{x}{b-a}\,dx \\ &= \frac{1}{2(b-a)(1-\delta)}\left(b^2 - (q_\delta(X))^2\right).\end{aligned}$$

Note that $\mathrm{ES}[X;\delta] = 0$ for $\delta = 1$ and $\mathrm{ES}[X;\delta] = \mathbb{E}[X] = \frac{a+b}{2}$ for $\delta = 0$.

Computation of $V_{\mathrm{ind}}^{\nu}(Z)$: Under the assumption of independence between the process Z and the stopping time τ or the adapted random probability measure γ we have $V_{\mathrm{ind}}^{\nu}(Z) = \frac{a+b}{2}$ for all maturities $T \in \mathbb{N}$, because of

$$\begin{aligned}V_{\mathrm{ind}}^{\nu}(Z) = \mathbb{E}[Z_\tau] &= \sum_{t \in I} \mathbb{E}[Z_t \mathbb{1}_{\{\tau = t\}}] \quad \text{(independence between } Z \text{ and } \tau) \\ &= \sum_{t \in I} \underbrace{\mathbb{E}[Z_t]}_{=\frac{a+b}{2}} \underbrace{\mathbb{E}[\mathbb{1}_{\{\tau = t\}}]}_{=\nu_t} \quad (Z_t \sim U(a,b) \text{ for all } t) \\ &= \left(\frac{a+b}{2}\right) \underbrace{\sum_{t \in I} \nu_t}_{=1} = \frac{a+b}{2}\end{aligned} \tag{3.6.3}$$

or rather

$$\begin{aligned}V_{\mathrm{ind}}^{\nu}(Z) = \mathbb{E}[Z_\gamma] = \mathbb{E}\Big[\sum_{t \in I} Z_t \gamma_t\Big] & \quad \text{(independence between } Z \text{ and } \gamma) \\ = \sum_{t \in I} \underbrace{\mathbb{E}[Z_t]}_{=\frac{a+b}{2}} \underbrace{\mathbb{E}[\gamma_t]}_{=\nu_t} & \quad (Z_t \sim U(a,b) \text{ for all } t) \\ = \left(\frac{a+b}{2}\right) \underbrace{\sum_{t \in I} \nu_t}_{=1} = \frac{a+b}{2}. &\end{aligned} \tag{3.6.4}$$

Computation of $V_T^\nu(Z)$: In order to be able to use Lemma A.3.7, respectively [33, Lemma 5.37], we have to use $\delta_t = 1 - \frac{\nu_t}{1-\nu_0-\ldots-\nu_{t-1}} = 1 - \frac{\nu_t}{\nu_{\geq t}} = \frac{\nu_{>t}}{\nu_{\geq t}}$ for each $t \in I$ for the computation of the expected shortfall, such that

$$V_T^\nu(Z) = \sum_{t=0}^{T} \nu_t \cdot \mathrm{ES}[Z_t; \delta_t] = \sum_{t=0}^{T} \nu_t \cdot \mathrm{ES}\left[Z_t; 1 - \frac{\nu_t}{\nu_{\geq t}}\right].$$

Computation of $V_T(Z)$: The computation of $V_T(Z)$ is described in Example 3.6.6 with $G = 0$.

Sample calculation: For a sample calculation we assume now that $Z_t \sim U(0,2)$ for all $t \in I$. We want to compute the price for a unit-linked life insurance contract for a 20-year old male person, with different maturities. We get the following values:

T	10	20	40	60	80
$V_{\mathrm{ind}}^\nu(Z)$	1.0	1.0	1.0	1.0	1.0
$V_T^\nu(Z)$	1.0071	1.0156	1.0736	1.3229	1.8866
$V_T(Z)$	1.7222	1.8398	1.9122	1.9393	1.9536

Figure 3.3 shows the evolution of the values $V_T(Z)$, $V_T^\nu(Z)$ and $V_{\mathrm{ind}}^\nu(Z)$ for different maturities. We see that the difference between $V_T^\nu(Z)$ and $V_{\mathrm{ind}}^\nu(Z)$ becomes higher for larger maturities, while the difference between $V_T(Z)$ and $V_T^\nu(Z)$ becomes smaller.

Example 3.6.5 (Log-normal distribution)**.** We assume that Z is an i.i.d. process with $Z_t \sim Log\mathcal{N}(\mu, \sigma^2)$ for all $t \in I$. Let the distribution ν be given as explained above, see (3.6.1). First we will compute the quantiles and expected shortfalls needed. For a random variable $X \sim Log\mathcal{N}(\mu, \sigma^2)$ the δ-quantile of X is given by

$$q_\delta(X) = \exp(\mu + u_{(\delta)} \cdot \sigma),$$

where $u_{(\delta)}$ denotes the δ-quantile of the standard normal distribution. The cumulative distribution function of the standard normal distribution will be denoted by Φ. Then the expected shortfall of X is

$$\mathrm{ES}[X;\delta] = \begin{cases} \mathbb{E}[X] = \exp(\mu + \frac{\sigma^2}{2}) & \text{for } \delta = 0, \\ \exp\left(\frac{\sigma^2}{2} + \mu\right)\Phi\left(\sigma - \frac{\ln(q_\delta(X)) - \mu}{\sigma}\right) & \text{for } \delta \in (0,1), \\ 0 & \text{for } \delta = 1. \end{cases}$$

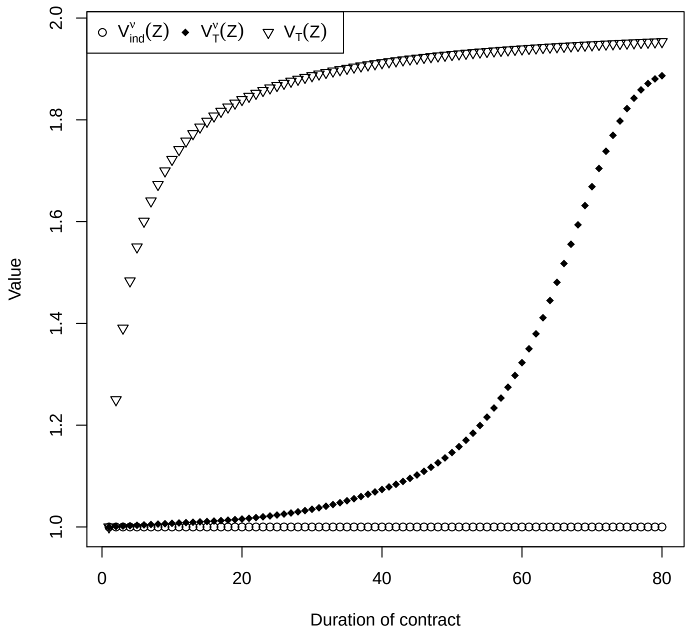


Figure 3.3.: The values $V_{\mathcal{T}}(Z)$, $V_{\mathcal{T}}^{\nu}(Z)$ and $V_{\text{ind}}^{\nu}(Z)$ for a unit-linked life insurance contract for a 20-year old male person for different maturities with an uniformly distributed underlying process on $[0,2]$.

For $\delta \in (0,1)$, it follows from

$$
\begin{aligned}
\mathrm{ES}[X;\delta] &= \frac{1}{\mathbb{P}(X > q_\delta(X))}\mathbb{E}[X\mathbb{1}_{\{X>q_\delta(X)\}}] \\
&= \frac{1}{1-\delta}\int_{q_\delta(X)}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}\sigma x}\exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right)dx \\
&\text{substitution: } y = \frac{\ln(x)-\mu}{\sigma}; \quad dy = \frac{dx}{\sigma x} \\
&= \frac{1}{\sqrt{2\pi}}\int_{u_{(\delta)}}^{\infty} \exp(y\sigma+\mu)\exp\Big(-\frac{y^2}{2}\Big)dy \\
&= \frac{1}{\sqrt{2\pi}}\exp(\frac{\sigma^2}{2}+\mu)\int_{u_{(\delta)}-\sigma}^{\infty} \exp\Big(-\frac{y^2}{2}\Big)dy \\
&\text{if } X \sim Log\mathcal{N}(\mu,\sigma^2), \text{ then } Y = \frac{\ln(X)-\mu}{\sigma} \sim \mathcal{N}(0,1) \\
&= \exp\Big(\frac{\sigma^2}{2}+\mu\Big)(1-\Phi(u_{(\delta)}-\sigma)) \quad (\Phi(-x) = 1-\Phi(x)) \\
&= \exp\Big(\frac{\sigma^2}{2}+\mu\Big)\Phi(\sigma-u_{(\delta)}) = \exp\Big(\frac{\sigma^2}{2}+\mu\Big)\Phi\left(\sigma - \frac{\ln(q_\delta(X))-\mu}{\sigma}\right).
\end{aligned}
$$

Computation of $V^{\nu}_{\text{ind}}(Z)$: Under the assumption of independence between the process Z and the stopping time τ or the adapted random probability measure γ we have

$$V^{\nu}_{\text{ind}}(Z) = \sum_{t \in I} \mathbb{E}[Z_t]\, \nu_t = \exp\left(\mu + \frac{\sigma^2}{2}\right)$$

for all maturities $T \in \mathbb{N}$.

Computation of $V^{\nu}_{T}(Z)$: In order to be able to use Lemma A.3.7 we have to use $\delta_t = 1 - \frac{\nu_t}{\nu_{\geq t}}$ for each $t \in I$ for the computation of the expected shortfall, so that

$$V^{\nu}_{T}(Z) = \sum_{t=0}^{T} \nu_t \cdot \mathrm{ES}[Z_t; \delta_t] = \sum_{t=0}^{T} \nu_t \cdot \mathrm{ES}\left[Z_t; 1 - \frac{\nu_t}{\nu_{\geq t}}\right].$$

Computation of $V_T(Z)$: The computation of $V_T(Z)$ is described in Example 3.6.11 for $G = 0$.

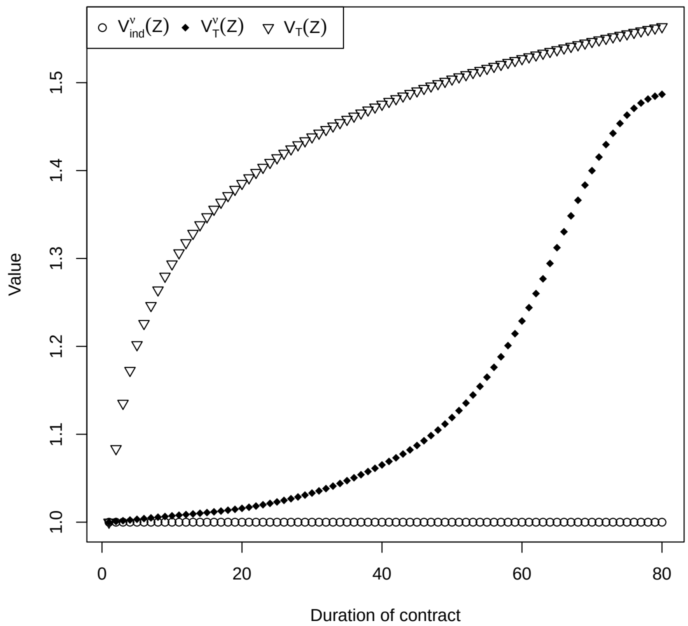


Figure 3.4.: The values $V_T(Z)$, $V^{\nu}_{T}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ for a unit-linked life insurance contract for a 20-year old male person for different maturities with a log-normally distributed underlying process with $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$.

Sample calculation: Now we want to compute the price for a unit-linked life insurance contract for a 20-year old male person with different maturities. Furthermore let $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$ such that $\mathbb{E}[Z_t] = \exp\left(\mu + \frac{\sigma^2}{2}\right) = 1$ for all $t \in I$. If σ describes an index, it is reasonable to choose σ lower than 30%. For this unit-linked life insurance we get the following values:

T	10	20	40	60	80
$V^{\nu}_{\text{ind}}(Z)$	1.0	1.0	1.0	1.0	1.0
$V^{\nu}_{\mathcal{T}}(Z)$	1.0072	1.0158	1.0651	1.2288	1.4867
$V_{\mathcal{T}}(Z)$	1.2938	1.3853	1.4755	1.5273	1.5636

Figure 3.4 shows the evolution of the values $V_{\mathcal{T}}(Z)$, $V^{\nu}_{\mathcal{T}}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ for different maturities. We get a similar behavior as in Figure 3.3. Note the serious difference between $V^{\nu}_{\text{ind}}(Z)$ and $V^{\nu}_{\mathcal{T}}(Z)$ for long-term maturities and the one between $V^{\nu}_{\mathcal{T}}(Z)$ and $V_{\mathcal{T}}(Z)$ for mid-term maturities.

3.6.2. Example with Guarantee

Nowadays, the payoff of the last examples is not absolutely realistic for an insurance contract. We are also interested in insurance contracts including a guarantee. Such a contract can often not be perfectly hedged. To model a unit-linked life insurance contract including a guarantee, let the process $S = (S_t)_{t\in I}$ model the underlying fund and let $G = (G_t)_{t\in I}$ model the guaranteed value. Then the payoff of the insurance contract, which we will denote by $Z = (Z_t)_{t\in I}$, at each time point $t \in I$ is given by $Z_t = \max\{S_t, G_t\}$. This payoff may then be represented as the sum of the fund and the value of a put option by writing $Z_t = S_t + (G_t - S_t)^+$ with $x^+ = \max\{x, 0\}$. The payoff will be modeled very simple by an i.i.d. process. This allows us to use Lemma A.3.7, see [33, Lemma 5.37].
In the following example we will now extend Example 3.6.2 by a guarantee. It corresponds to [33, Example 9.4].

Example 3.6.6 (Uniform distribution)**.** Let $S = (S_t)_{t\in I}$ be an i.i.d. process with $S_t \sim U(a,b)$ for all $t \in I$ and let $G = (G_t)_{t\in I}$ be a deterministic process with $G_t \in [a,b]$ for all $t \in I$. The process G will model the guaranteed value. Let the process $Z = (Z_t)_{t\in I}$ be given by $Z_t = \max\{S_t, G_t\}$ for all $t \in I$ as described above. Let the distribution ν be given as explained above, see (3.6.1). For each $t \in I$ and $x \in \mathbb{R}$ the distribution of Z_t is given by

$$\mathbb{P}(Z_t \le x) = \begin{cases} 0 & \text{if } x < G_t, \\ \mathbb{P}(S_t \le G_t) & \text{if } x = G_t, \\ \mathbb{P}(S_t \le G_t) + \mathbb{P}(G_t < S_t \le x) & \text{if } G_t < x < b, \\ 1 & \text{if } x \ge b. \end{cases}$$

Furthermore, we have that

$$\mathbb{P}(S_t \le G_t) = \int_a^{G_t} \frac{1}{b-a}\,dx = \frac{G_t - a}{b-a} \quad \text{and}$$
$$\mathbb{P}(G_t < S_t \le x) = \mathbb{P}(S_t \le x) - \mathbb{P}(S_t \le G_t) = \frac{x-a}{b-a} - \frac{G_t - a}{b-a} = \frac{x - G_t}{b-a}.$$

Then it follows that

$$\mathbb{P}(Z_t \le x) = \begin{cases} 0 & \text{if } x < G_t, \\ \frac{x-a}{b-a} & \text{if } G_t \le x < b, \\ 1 & \text{if } x \ge b. \end{cases}$$

For every $t \in I$ the δ-quantile of Z_t is therefore given by

$$q_\delta(Z_t) = \begin{cases} -\infty & \text{for } \delta = 0, \\ G_t & \text{for } 0 < \delta \le \mathbb{P}(S_t \le G_t), \\ \delta(b-a)+a & \text{for } \mathbb{P}(S_t \le G_t) < \delta \le 1, \end{cases}$$

and the expected value of Z_t is given by

$$\begin{aligned} \mathbb{E}[Z_t] &= \mathbb{E}[\max\{S_t, G_t\}] = \mathbb{E}[G_t \mathbb{1}_{\{S_t > G_t\}}] + \mathbb{E}[S_t \mathbb{1}_{\{S_t \ge G_t\}}] \\ &\quad (G \text{ is a deterministic process}) \\ &= G_t\, \mathbb{P}(S_t \le G_t) + \int_{G_t}^{b} \frac{s}{b-a}\, ds = \frac{1}{b-a}\left(G_t^2 - aG_t + \frac{1}{2}(b^2 - G_t^2)\right) \\ &= \frac{1}{b-a}\left(\frac{1}{2}(G_t^2 + b^2) - aG_t\right). \end{aligned}$$

For simplification we will assume that $G_t = G$ for all $t \in I$, such that Z is also an i.i.d. process.

Computation of $V_{\text{ind}}^{\nu}(Z)$: Under the assumption from above and the independence of the process Z and the stopping time τ or the adapted random probability measure γ we have that $V_{\text{ind}}^{\nu}(Z) = \mathbb{E}[Z_0]$ for all maturities $T \in \mathbb{N}$, compare (3.6.3) and (3.6.4). Therefore the value of $V_{\text{ind}}^{\nu}(Z)$ is given by

$$V_{\text{ind}}^{\nu}(Z) = \frac{1}{b-a}\left(\frac{1}{2}(G^2 + b^2) - aG\right).$$

Computation of $V_{\mathcal{T}}^{\nu}(Z)$: To compute the value $V_{\mathcal{T}}^{\nu}(Z)$ we need the expected shortfall of Z_t for all $t \in I$. It is computed as

$$\mathrm{ES}[Z_t;\delta] = \begin{cases} \frac{1}{b-a}\left(\frac{1}{2}(G_t^2 + b^2) - aG_t\right) & \text{for } \delta = 0, \\ \frac{1}{2(b-a)(1-\delta)}(b^2 - (q_\delta(Z_t))^2) & \text{for } \delta \in (0,1), \\ 0 & \text{for } \delta = 1. \end{cases}$$

In order to be able to use Lemma A.3.7, respectively [33, Lemma 5.37], we have to use $\delta_t = 1 - \frac{\nu_t}{\nu_{\ge t}} = \frac{\nu_{>t}}{\nu_{\ge t}}$ for each $t \in I$. Then we could determine $V_{\mathcal{T}}^{\nu}(Z)$ as

$$V_{\mathcal{T}}^{\nu}(Z) = \sum_{t=0}^{T} \nu_t \cdot \mathrm{ES}[Z_t;\delta_t] = \sum_{t=0}^{T} \nu_t \cdot \mathrm{ES}\left[Z_t; 1 - \frac{\nu_t}{\nu_{\ge t}}\right].$$

Computation of $V_T(Z)$: The value $V_T(Z)$ coincides for a non-negative process Z with the value of a standard American option without any hedging possibilities. The pricing of American options or optimal stopping problems are well known problems in the literature. The calculation results from the Snell envelope $U = (U_t)_{t\in I}$ of Z which is given by

$$U_t = \begin{cases} Z_T & \text{if } t = T, \\ \max\{Z_t, \mathbb{E}[U_{t+1}|\mathcal{F}_t]\} & \text{otherwise,} \end{cases} \tag{3.6.7}$$

and the corresponding process of values $(V_t)_{t\in I}$ which is given by

$$V_t = \sup\{\, \mathbb{E}[Z_{\tau_t}] \mid \tau_t : \Omega \to \{t, \dots, T\} \text{ stopping time } \}.$$

Then the value $V_T(Z)$ is determined by $V_T(Z) = V_0$.
Using the recursive scheme (3.6.7) of the Snell envelope and the assumptions from above, yields $U_T = Z_T$ and $U_{T-1} = \max\{Z_{T-1}, \mathbb{E}[U_T]\}$, because

$$\begin{aligned} \mathbb{E}[U_T|\mathcal{F}_{T-1}] &= \mathbb{E}[Z_T|\mathcal{F}_{T-1}] \text{ (independence between } Z_T \text{ and } \mathcal{F}_{T-1}) \\ &= \mathbb{E}[Z_T] = \mathbb{E}[U_T]. \end{aligned}$$

For each $t \in \{0, \dots, T-1\}$ we get recursively that $\mathbb{E}[U_t|\mathcal{F}_{t-1}] = \mathbb{E}[U_t]$ and consequently $U_t = \max\{Z_t, \mathbb{E}[U_{t+1}]\}$. Furthermore it holds

$$\begin{aligned} \mathbb{E}[U_t] &= \mathbb{E}\Big[\max\{Z_t, \mathbb{E}[U_{t+1}]\}\Big] = \mathbb{E}\Big[Z_t \mathbb{1}_{\{Z_t > \mathbb{E}[U_{t+1}]\}}\Big] + \mathbb{E}\Big[\mathbb{E}[U_{t+1}]\mathbb{1}_{\{Z_t \le \mathbb{E}[U_{t+1}]\}}\Big] \\ &= \mathbb{E}\Big[Z_t \mathbb{1}_{\{Z_t > \mathbb{E}[U_{t+1}]\}}\Big] + \mathbb{E}[U_{t+1}]\mathbb{P}(Z_t \le \mathbb{E}[U_{t+1}]). \end{aligned} \tag{3.6.8}$$

We know for every $t \in I$ and $K \in \mathbb{R}$ with $K \neq G_t$ that

$$\begin{aligned} \mathbb{E}\Big[Z_t \mathbb{1}_{\{Z_t > K\}}\Big] &= \mathbb{E}\Big[\max\{S_t, G_t\}\mathbb{1}_{\{\max\{S_t,G_t\} > K\}}\Big] = \mathbb{E}\Big[S_t \mathbb{1}_{\{S_t > G_t\}}\mathbb{1}_{\{S_t > K\}}\Big] + \mathbb{E}\Big[G_t \mathbb{1}_{\{G_t > K\}}\mathbb{1}_{\{S_t \le G_t\}}\Big] \\ &\quad G_t \text{ is deterministic} \\ &= \mathbb{E}\Big[S_t \mathbb{1}_{\{S_t > \max\{G_t, K\}\}}\Big] + G_t \mathbb{1}_{\{G_t > K\}}\mathbb{E}\Big[\mathbb{1}_{\{S_t \le G_t\}}\Big] \\ &= \mathbb{E}\Big[S_t \mathbb{1}_{\{S_t > \max\{G_t, K\}\}}\Big] + G_t \cdot \max\left\{\frac{G_t - K}{|G_t - K|}, 0\right\} \cdot \mathbb{P}(S_t \le G_t) \\ &= \frac{1}{b-a}\left(\frac{1}{2}(b^2 - (\max\{G_t, K\})^2\right) + G_t \cdot \max\left\{\frac{G_t - K}{|G_t - K|}, 0\right\} \cdot \frac{G_t - a}{b-a}. \end{aligned}$$

The process of values $(V_t)_{t\in I}$ can also be calculated recursively. For the first steps it holds that $V_T = \mathbb{E}[Z_{\tau_T}] = \mathbb{E}[Z_T]$ and $V_{T-1} = \mathbb{E}[Z_{\tau_{T-1}}] = \mathbb{E}[U_{\tau_{T-1}}] = \mathbb{E}[U_{\tau_{T-1}\wedge T}] = \mathbb{E}[U_{T-1}]$, because $U_{\tau_{T-1}}$ is a martingale, τ_{T-1} is optimal and $\tau_{T-1} \ge T-1$. For each $t \in I$ we get then

$$V_t = \mathbb{E}[Z_{\tau_t}] = \mathbb{E}[U_t].$$

Finally, it follows that

$$
\begin{aligned}
V_t = \mathbb{E}[U_t] &= \mathbb{E}\Big[Z_t \mathbb{1}_{\{Z_t > \mathbb{E}[U_{t+1}]\}}\Big] + \mathbb{E}[U_{t+1}]\mathbb{P}(Z_t \leq \mathbb{E}[U_{t+1}]) \\
&= \mathbb{E}\Big[S_t \mathbb{1}_{\{S_t > \max\{G_t, \mathbb{E}[U_{t+1}]\}\}}\Big] + G_t \mathbb{1}_{\{G_t > \mathbb{E}[U_{t+1}]\}}\mathbb{P}(S_t \leq G_t) + \mathbb{E}[U_{t+1}]\mathbb{P}(Z_t \leq \mathbb{E}[U_{t+1}]) \qquad (3.6.9) \\
&= \frac{1}{b-a}\left(\frac{1}{2}(b^2 - (\max\{G_t, \mathbb{E}[U_{t+1}]\})^2) + G_t \cdot (G_t - a) \cdot \max\left\{\frac{G_t - \mathbb{E}[U_{t+1}]}{|G_t - \mathbb{E}[U_{t+1}]|}, 0\right\}\right) \\
&\quad + \mathbb{E}[U_{t+1}]\mathbb{P}(Z_t \leq \mathbb{E}[U_{t+1}]). \qquad (3.6.10)
\end{aligned}
$$

Note that in equation (3.6.9) the term $G_t \mathbb{1}_{G_t \geq \mathbb{E}[U_{t+1}]}\mathbb{P}(S_t \leq G_t)$ or $\mathbb{E}[U_{t+1}]\mathbb{P}(Z_t \leq \mathbb{E}[U_{t+1}])$ is zero. If $G = (G_t)_{t\in I}$ dominates $(\mathbb{E}[U_t])_{t\in I}$, we get that the last term $\mathbb{E}[U_{t+1}]\mathbb{P}(Z_t \leq \mathbb{E}[U_{t+1}])$ of equation (3.6.9) is zero for all $t \in T$. We get this through the following considerations. We know that the Snell envelope $U = (U_t)_{t\in\{0,\ldots,T\}}$ is a supermartingale in general, i.e., $\mathbb{E}[U_t|\mathcal{F}_s] \leq U_s$ for all $s \leq t$. It holds for all $t \in \{0,\ldots,T\}$

$$\mathbb{E}[Z_T] = \mathbb{E}[U_T] \leq \ldots \leq \mathbb{E}[U_t] \leq \ldots \leq \mathbb{E}[U_0].$$

In the case of uniform distribution and for every constant guarantee G we have that

$$
\begin{aligned}
G &\geq \mathbb{E}[U_T] = \mathbb{E}[Z_T] = \frac{1}{b-a}\left(\frac{1}{2}(G^2 + b^2) - aG\right) \\
0 &\geq \frac{1}{2(b-a)}G^2 - \frac{b}{b-a}G + \frac{1}{2(b-a)}b^2 \\
0 &\geq G^2 - 2bG + b^2 = (G-b)^2,
\end{aligned}
$$

such that $0 = (G-b)^2$, if $G = b$. It follows for $G_t = G$ for all $t \in I$ (note: Z is an i.i.d. process) and $G \geq \mathbb{E}[U_T] = \mathbb{E}[Z_T]$ (note: $\mathbb{E}[Z_T] = \mathbb{E}[Z_t]\forall t$) that $V_T(Z) = V^{\nu}_{\text{ind}}(Z)$, because

$$
\begin{aligned}
V_T(Z) = V_0 &= \mathbb{E}[S_0 \mathbb{1}_{\{S_0 > \max\{G, \mathbb{E}[U_1]\}\}}] + G\mathbb{1}_{\{G > \mathbb{E}[U_1]\}}\mathbb{P}(S_t \leq G) \\
&= \mathbb{E}[S_0 \mathbb{1}_{\{S_0 > G\}}] + G\,\mathbb{P}(S_0 \leq G) = \mathbb{E}[Z_0] = V^{\nu}_{\text{ind}}(Z).
\end{aligned}
$$

If $G = b$, we get $V_T(Z) = V^{\nu}_{\text{ind}}(Z)$ for all maturities.

Sample calculation: Assume that we again want to compute the price for a unit-linked life insurance contract for a 20-year old male person with different maturities. Furthermore $S_t \sim U(0,2)$ and $G_t = 1$ for all $t \in I$. For this unit-linked life insurance we get the following values:

T	10	20	40	60	80
$V^{\nu}_{\text{ind}}(Z)$	1.25	1.25	1.25	1.25	1.25
$V^{\nu}_T(Z)$	1.2553	1.2617	1.3052	1.4908	1.8908
$V_T(Z)$	1.7415	1.8462	1.9142	1.9403	1.9541

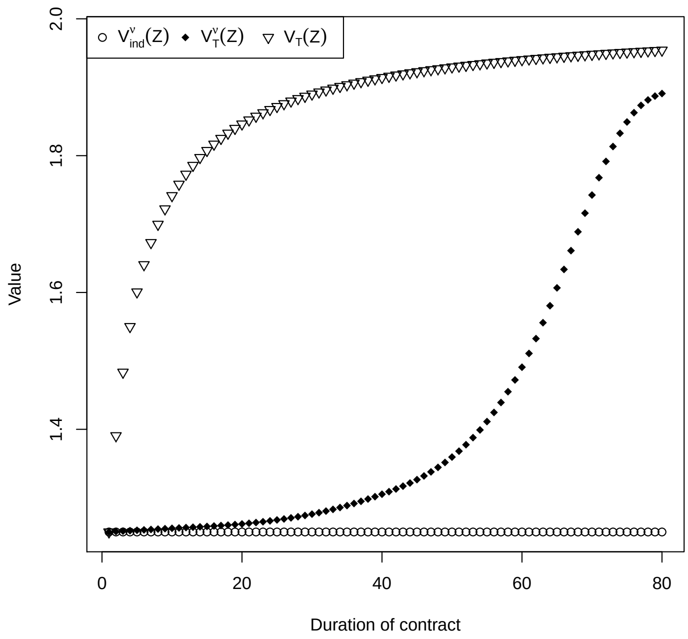


Figure 3.5.: The values $V_{\mathcal{T}}(Z)$, $V^{\nu}_{\mathcal{T}}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ for a unit-linked life insurance contract with guarantee 1 for a 20-year old male person for different maturities with an uniformly distributed underlying process on $[0,2]$.

We get another solution as in [33], but there is the same trend as in Figure 3.5.
Figure 3.5 shows the evolution of the values $V_{\mathcal{T}}(Z)$, $V^{\nu}_{\mathcal{T}}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ for different maturities. We see that the difference between $V^{\nu}_{\mathcal{T}}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ becomes higher for larger maturities, while the difference between $V_{\mathcal{T}}(Z)$ and $V^{\nu}_{\mathcal{T}}(Z)$ becomes smaller for larger maturities. Note that including the guarantee increases the values of $V_{\mathcal{T}}(Z)$, $V^{\nu}_{\mathcal{T}}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ from the beginning, compared to Figure 3.3 it does not increase the difference between $V^{\nu}_{\mathcal{T}}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$.

We still consider the values for unit-linked life insurance contracts with guarantee 1.5 and 0.5. For this unit-linked life insurance contract with $S_t \sim U(0,2)$ and $G_t = 1.5$ for all $t \in I$ we get the following values:

T	10	20	40	60	80
$V^{\nu}_{\text{ind}}(Z)$	1.5625	1.5625	1.5625	1.5625	1.5625
$V^{\nu}_{\mathcal{T}}(Z)$	1.5656	1.5693	1.5946	1.7007	1.8774
$V_{\mathcal{T}}(Z)$	1.7884	1.8637	1.9198	1.9430	1.9558

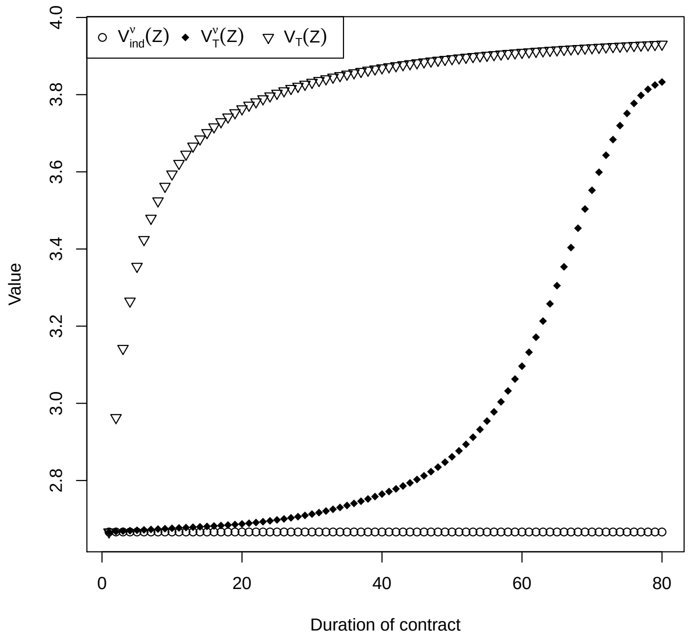


Figure 3.6.: The values $V_{\mathcal{T}}(Z)$, $V_{\mathcal{T}}^{\nu}(Z)$ and $V_{\text{ind}}^{\nu}(Z)$ for a unit-linked life insurance contract with guarantee 2 for a 20-year old male person for different maturities with an uniformly distributed underlying process on $[1,4]$.

For this unit-linked life insurance contract with $S_t \sim U(0,2)$ and $G_t = 0.5$ for all $t \in I$ we get the following values:

T	10	20	40	60	80
$V_{\text{ind}}^{\nu}(Z)$	1.0625	1.0625	1.0625	1.0625	1.0625
$V_{\mathcal{T}}^{\nu}(Z)$	1.0692	1.0771	1.1315	1.3648	1.8877
$V_{\mathcal{T}}(Z)$	1.7261	1.8410	1.9126	1.9395	1.9537

We want to consider another unit-linked life insurance contract for a 20-year old male person. Now, the process $S_t \sim U(1,4)$ and $G_t = 2$ for all $t \in I$. Then we get the following values:

T	10	20	40	60	80
$V_{\text{ind}}^{\nu}(Z)$	2.6667	2.6667	2.6667	2.6667	2.6667
$V_{\mathcal{T}}^{\nu}(Z)$	2.6762	2.6875	2.7648	3.0962	3.8327
$V_{\mathcal{T}}(Z)$	3.5943	3.7632	3.8694	3.9095	3.9307

The corresponding results for the different maturities are shown in Figur 3.6.

In the following example we will now extend Example 3.6.5 by a guarantee.

Example 3.6.11 (Log-normal distribution)**.** We assume that $S = (S_t)_{t\in I}$ is an i.i.d. process with $S_t \sim \mathrm{Log}\,\mathcal{N}(\mu,\sigma^2)$ for all $t \in I$. Let $G = (G_t)_{t\in I}$ be a deterministic process with $G_t \in \mathbb{R}_+$ for all $t \in I$ which models the guaranteed value. Moreover, the process $Z = (Z_t)_{t\in I}$ is again given by $Z_t = S_t + (G_t - S_t)^+$ for all $t \in I$. Let the distribution ν be given as explained above, see (3.6.1). First we will compute the distribution of Z_t for each $t \in I$ and $x \in \mathbb{R}$. For this we have that

$$\mathbb{P}(Z_t \le x) = \begin{cases} 0 & \text{if } x < G_t, \\ \mathbb{P}(S_t \le G_t) & \text{if } x = G_t, \\ \mathbb{P}(S_t \le G_t) + \mathbb{P}(G_t < S_t \le x) & \text{if } G_t < x. \end{cases}$$

Furthermore, the cumulative distribution function of the standard normal distribution will be denoted by Φ and $u_{(\delta)}$ denote the δ-quantile of the standard normal distribution. Then we have that

$$\begin{aligned} \mathbb{P}(S_t \le G_t) &= \int_0^{G_t} \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right) dx \\ &\text{substitution: } z = \frac{\ln(x)-\mu}{\sigma} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{\ln(G_t)-\mu}{\sigma}} \exp\left(-\frac{z^2}{2}\right) dz = \Phi\left(\frac{\ln(G_t)-\mu}{\sigma}\right) \end{aligned}$$

and $\mathbb{P}(G_t < S_t \le x) = \mathbb{P}(S_t \le x) - \mathbb{P}(S_t \le G_t) = \Phi\left(\frac{\ln(x)-\mu}{\sigma}\right) - \Phi\left(\frac{\ln(G_t)-\mu}{\sigma}\right)$. It follows that

$$\mathbb{P}(Z_t \le x) = \begin{cases} 0 & \text{if } x < G_t, \\ \Phi\left(\frac{\ln(x)-\mu}{\sigma}\right) & \text{if } G_t \le x. \end{cases}$$

For every $t \in I$ the δ-quantile of Z_t is therefore given by

$$q_\delta(Z_t) = \begin{cases} -\infty & \text{for } \delta = 0, \\ G_t & \text{for } 0 < \delta \le \mathbb{P}(S_t \le G_t), \\ \exp(\mu + u_{(\delta)} \cdot \sigma) & \text{for } \mathbb{P}(S_t \le G_t) < \delta \le 1. \end{cases}$$

For simplicity we will now assume that G is deterministic, for example G_t is constant for all $t \in I$.

Computation of $V^\nu_{\mathrm{ind}}(Z)$: Under the assumption from above and the independence of the process Z and the stopping time τ or the adapted random probability mesure γ we have that $V^\nu_{\mathrm{ind}}(Z) = \sum_{t=1}^T \mathbb{E}[Z_t]\nu_t$ for all maturities $T \in \mathbb{N}$, compare (3.6.3) and (3.6.4). Given that $S_t \sim \mathrm{Log}\,\mathcal{N}(\mu,\sigma^2)$ for every $t \in I$ the expected value of Z_t for each $t \in I$ is given by

$$\begin{aligned} \mathbb{E}[Z_t] &= \mathbb{E}[\max\{S_t, G_t\}] = \mathbb{E}[G_t \mathbb{1}_{\{S_t \le G_t\}}] + \mathbb{E}[S_t \mathbb{1}_{\{S_t > G_t\}}] \\ &= G_t\,\mathbb{P}(S_t \le G_t) + \int_{G_t}^{\infty} \frac{x}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right) dx \\ &= G_t \cdot \Phi\left(\frac{\ln(G_t)-\mu}{\sigma}\right) + \exp\left(\frac{\sigma^2}{2} + \mu\right) \Phi\left(-\frac{\ln(G_t)-\mu}{\sigma} + \sigma\right). \end{aligned} \tag{3.6.12}$$

Furthermore, if $G_t = G$ for all $t \in I$ and a constant $G \in \mathbb{R}_+$, then Z is also an i.i.d. process and $V^{\nu}_{\text{ind}}(Z) = \mathbb{E}[Z_0]$ for all maturities $T \in \mathbb{N}$.

Computation of $V^{\nu}_{\mathcal{T}}(Z)$: To compute the value $V^{\nu}_{\mathcal{T}}(Z)$ we need the expected shortfall of Z_t for all $t \in I$. For $\delta \in (0,1)$ it is computed as

$$\begin{aligned}\mathrm{ES}[Z_t;\delta] &= \frac{1}{(1-\delta)} \exp\left(\frac{\sigma^2}{2} + \mu\right)\Phi(\sigma - u_{(\delta)}) \\ &= \frac{1}{(1-\delta)} \exp\left(\frac{\sigma^2}{2} + \mu\right)\Phi\left(\sigma - \frac{\ln(q_\delta(Z_0)) - \mu}{\sigma}\right).\end{aligned}$$

Note that for each $t \in I$ $\mathrm{ES}[Z_t;0] = G_t \cdot \Phi\left(\frac{\ln(G_t)-\mu}{\sigma}\right) + \exp\left(\frac{\sigma^2}{2} + \mu\right)\Phi\left(-\frac{\ln(G_t)-\mu}{\sigma} + \sigma\right)$ and $\mathrm{ES}[Z_t;1] = 0$. In order to be able to use Lemma A.3.7, respectively [33, Lemma 5.37], we have to use $\delta_t = 1 - \frac{\nu_t}{\nu_{\geq t}}$ for each $t \in I$. Then we could determine $V^{\nu}_{\mathcal{T}}(Z)$ as

$$V^{\nu}_{\mathcal{T}}(Z) = \sum_{t=0}^{T} \nu_t \cdot \mathrm{ES}[Z_t;\delta_t] = \sum_{t=0}^{T} \nu_t \cdot \mathrm{ES}\left[Z_t; 1 - \frac{\nu_t}{\nu_{\geq t}}\right].$$

Computation of $V_{\mathcal{T}}(Z)$: The value $V_{\mathcal{T}}(Z)$ coincides for a non-negative process Z with the value of a standard American option without any hedging possibilities. The computation of $V_{\mathcal{T}}(Z)$ could be described analogously to the previous example. The calculation results also from the Snell envelope $U = (U_t)_{t\in I}$ of Z and its process of values $V = (V_t)_{t\in I}$. Using the definition (3.6.7) of the Snell envelope and the assumption of Z we have for each $t \in \{0,\dots,T-1\}$ that $U_t = \max\{Z_t, \mathbb{E}[U_{t+1}]\}$ and $U_T = Z_T$. It hold for each $t \in I$ that

$$V_t = \mathbb{E}[Z_{\tau_t}] = \mathbb{E}[U_t].$$

The value $V_{\mathcal{T}}(Z)$ is determined by $V_{\mathcal{T}}(Z) = V_0$. Furthermore it holds analogously to (3.6.8) that

$$\mathbb{E}[U_t] = \mathbb{E}[Z_t \mathbb{1}_{\{Z_t > \mathbb{E}[U_{t+1}]\}}] + \mathbb{E}[U_{t+1}]\mathbb{P}(Z_t \leq \mathbb{E}[U_{t+1}]).$$

We know for every $t \in I$ and $K \in \mathbb{R}$ with $K \neq G_t$ that

$$\begin{aligned}\mathbb{E}[Z_t \mathbb{1}_{\{Z_t>K\}}] &= \mathbb{E}[S_t \mathbb{1}_{\{S_t>G_t\}}\mathbb{1}_{\{S_t>K\}}] + \mathbb{E}[G_t \mathbb{1}_{\{G_t>K\}}\mathbb{1}_{\{S_t\leq G_t\}}] \\ &\quad G_t \text{ is deterministic} \\ &= \mathbb{E}[S_t \mathbb{1}_{\{S_t>\max\{G_t,K\}\}}] + G_t \mathbb{1}_{\{G_t>K\}}\mathbb{E}[\mathbb{1}_{\{S_t\leq G_t\}}] \\ &= \mathbb{E}[S_t \mathbb{1}_{\{S_t>\max\{G_t,K\}\}}] + G_t \cdot \max\left\{\frac{G_t-K}{|G_t-K|}, 0\right\} \cdot \mathbb{P}(S_t \leq G_t) \\ &= \exp\left(\frac{\sigma^2}{2} + \mu\right)\Phi\left(\frac{\mu - \ln(\max\{G_t,K\})}{\sigma} + \sigma\right) \\ &\qquad + G_t \cdot \max\left\{\frac{G_t-K}{|G_t-K|}, 0\right\} \cdot \Phi\left(\frac{\ln(G_t)-\mu}{\sigma}\right).\end{aligned}$$

Then we have that

$$\begin{aligned} V_t = \mathbb{E}[U_t] = &\exp\left(\frac{\sigma^2}{2}+\mu\right)\Phi\left(-\frac{\ln(\max\{G_t,\mathbb{E}[U_{t+1}]\})-\mu}{\sigma}+\sigma\right) \\ &+ G_t\cdot\max\left\{\frac{G_t-\mathbb{E}[U_{t+1}]}{|G_t-\mathbb{E}[U_{t+1}]|},0\right\}\cdot\Phi\left(\frac{\ln(G_t)-\mu}{\sigma}\right) \\ &+ \mathbb{E}[U_{t+1}]\mathbb{P}(Z_t\le\mathbb{E}[U_{t+1}]). \end{aligned}$$

More specifically, we have for all $t\in I$ that if

(a) $G_t \ge \mathbb{E}[U_{t+1}]$

$$V_t = \exp\left(\frac{\sigma^2}{2}+\mu\right)\Phi\left(-\frac{\ln(G_t)-\mu}{\sigma}+\sigma\right)+G_t\cdot\Phi\left(\frac{\ln(G_t)-\mu}{\sigma}\right),$$

(b) $G_t \le \mathbb{E}[U_{t+1}]$

$$V_t = \exp\left(\frac{\sigma^2}{2}+\mu\right)\Phi\left(-\frac{\ln(\mathbb{E}[U_{t+1}])-\mu}{\sigma}+\sigma\right)+\mathbb{E}[U_{t+1}]\cdot\Phi\left(\frac{\ln(\mathbb{E}[U_{t+1}])-\mu}{\sigma}\right).$$

Finally, it follows that

$$\begin{aligned} V_t = \mathbb{E}[U_t] = &\exp\left(\frac{\sigma^2}{2}+\mu\right)\Phi\left(-\frac{\ln(\max\{G_t,\mathbb{E}[U_{t+1}]\})-\mu}{\sigma}+\sigma\right) \\ &+ \max\{G_t,\mathbb{E}[U_{t+1}]\}\cdot\Phi\left(\frac{\ln(\max\{G_t,\mathbb{E}[U_{t+1}]\})-\mu}{\sigma}\right). \end{aligned} \tag{3.6.13}$$

The value $V_{\mathcal{T}}(Z)$ is determined by V_0. Analogously to Example 3.6.6, if $G_t = G$ for all $t\in I$ (note: Z is then also an i.i.d. process) and $G \ge \mathbb{E}[U_T] = \mathbb{E}[Z_T]$ (note: $\mathbb{E}[Z_T] = \mathbb{E}[Z_t]\,\forall t$), we have that $V_{\mathcal{T}}(Z) = V^{\nu}_{\text{ind}}(Z)$. Figure 3.9 shows for which choice G dominates the expected values of the Snell envelope $(\mathbb{E}[U_t])_{t\in I}$, in the case that $S_t \sim \text{Log}\mathcal{N}(\mu,\sigma^2)$ with $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$.

Sample calculation: Assume that we again want to compute the price for a unit-linked life insurance contract for a 20-year old male person with different maturities. Furthermore let $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$, so that $\exp\left(\mu+\frac{\sigma^2}{2}\right) = 1$ for all $t\in I$. Let $S_t \sim \text{Log}\mathcal{N}(\mu,\sigma^2)$ and $G_t = 1$ for all $t\in I$. Then we get for this unit-linked life insurance the following values:

T	10	20	40	60	80
$V^{\nu}_{\text{ind}}(Z)$	1.0836	1.0836	1.0836	1.0836	1.0836
$V^{\nu}_{\mathcal{T}}(Z)$	1.0903	1.0981	1.1426	1.2850	1.4881
$V_{\mathcal{T}}(Z)$	1.3064	1.3917	1.4786	1.5294	1.5652

Figure 3.7 shows the evolution of the values $V_{\mathcal{T}}(Z)$, $V^{\nu}_{\mathcal{T}}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ for different maturities. In this figure we see that the difference between $V^{\nu}_{\mathcal{T}}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ becomes higher for larger maturities, while the difference between $V_{\mathcal{T}}(Z)$ and $V^{\nu}_{\mathcal{T}}(Z)$ becomes smaller for larger maturities. We get analogously results as in the previous example with the uniformly distributed underlying process.

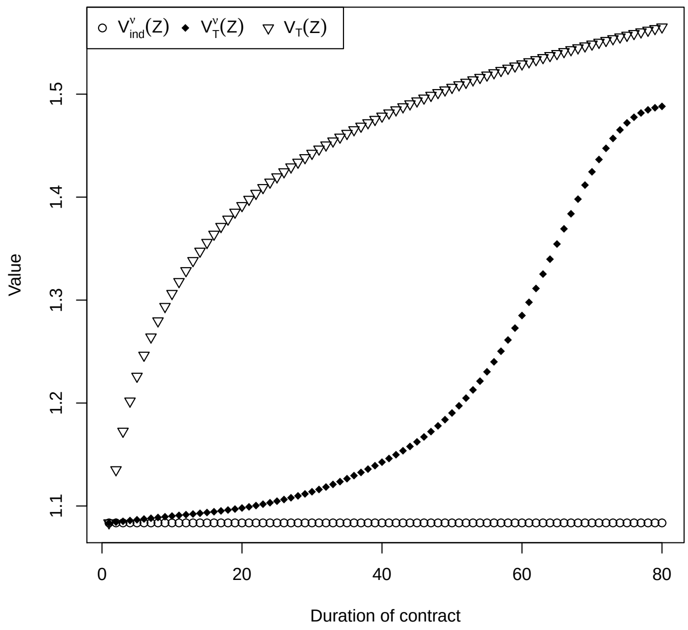


Figure 3.7.: The values $V_{\mathcal{T}}(Z)$, $V^{\nu}_{\mathcal{T}}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ for a unit-linked life insurance contract with guarantee 1 for a 20-year old male person for different maturities with a log-normally distributed underlying process with $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$.

We consider the same unit-linked life insurance contract with other guarantees. The underlying process S is again log-normally distributed with $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$. Then for $G_t = 0.5$ for all $t \in I$ we get the following values:

T	10	20	40	60	80
$V^{\nu}_{\text{ind}}(Z)$	1.0000	1.0000	1.0000	1.0000	1.0000
$V^{\nu}_{\mathcal{T}}(Z)$	1.0072	1.0158	1.0652	1.2288	1.4867
$V_{\mathcal{T}}(Z)$	1.2938	1.3853	1.4755	1.5273	1.5636

For $G_t = 4$ for all $t \in I$ we get the following values:

T	10	20	40	60	80
$V^{\nu}_{\text{ind}}(Z)$	4.0000	4.0000	4.0000	4.0000	4.0000
$V^{\nu}_{\mathcal{T}}(Z)$	3.9715	3.9375	3.7046	2.6867	0.0665
$V_{\mathcal{T}}(Z)$	4.0000	4.0000	4.0000	4.0000	4.0000

Note that the choice $G_t = 4$ for all $t \in I$ illustrates the case that $V_T(Z) = V^{\nu}_{\text{ind}}(Z)$, see Figure 3.9. Figure 3.9 shows for which choice G dominates the expected values of the Snell envelope $(\mathbb{E}[U_t])_{t \in I}$, in the case that $S_t \sim \text{Log}\,\mathcal{N}(\mu, \sigma^2)$ with $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$.

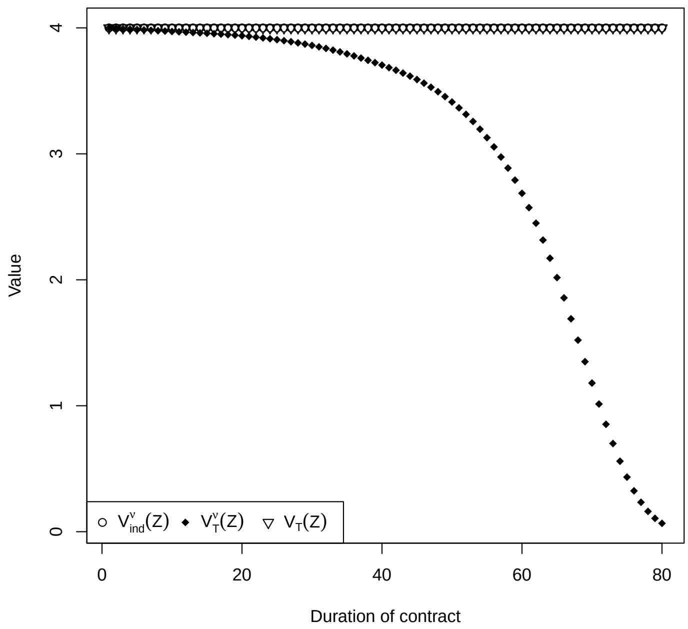


Figure 3.8.: The values $V_T(Z)$, $V^{\nu}_T(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ for a unit-linked life insurance contract with guarantee 4 for a 20-year old male person for different maturities with a log-normally distributed underlying process with $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$.

The table for $G_t = 4$, $t \in I$, and Figure 3.8 shows the evolution of the values $V_T(Z)$, $V^{\nu}_T(Z)$ and $V^{\nu}_{\text{ind}}(Z)$. It shows a totally different behavior, such that it is obviously that the choice of the guarantee is important and the choice should be well-considered.

Finally, we consider another unit-linked life insurance contract. The underlying process S is again log-normally distributed with $\sigma = 0.18$ and $\mu = 1$. Then for $G_t = 0.9$, $t \in I$, we get the following values:

T	10	20	40	60	80
$V^{\nu}_{\text{ind}}(Z)$	2.7627	2.7627	2.7627	2.7627	2.7627
$V^{\nu}_T(Z)$	2.7790	2.7982	2.9103	3.2859	3.8862
$V_T(Z)$	3.4495	3.6587	3.8626	3.9790	4.0603

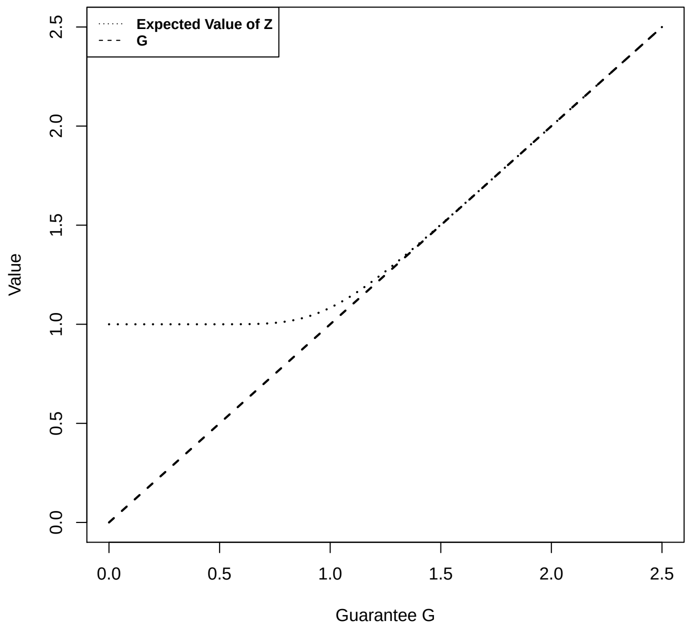


Figure 3.9.: Let $S_t \sim \text{Log}\mathcal{N}(\mu, \sigma^2)$ with $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$ and $G_t = G$ for all $t \in I$. The relation between G and $\mathbb{E}[Z_t]$, $t \in I$, is given.

3.6.3. Comparison

In this subsection we want to compare the two cases of a unit-linked life insurance contract without and with guarantee separately, see Subsection 3.6.1 and Subsection 3.6.2. For comparison we consider an uniformly distributed underlying process S^1 and a log-normally distributed process S^2. Both processes should have the same expected value. For example in the case without guarantee, we will assume that $S^1 = Z^1$ is an i.i.d. process with $S_t^1 \sim U(a,b)$ for all $t \in I$ as in Example 3.6.2 and $S^2 = Z^2$ is an i.i.d. process with $S_t^2 \sim \text{Log}\mathcal{N}(\mu, \sigma^2)$ for all $t \in I$ as in Example 3.6.5. We demand that $\mathbb{E}[Z_t^1] = \mathbb{E}[Z_t^1]$ for all $t \in I$. Because of this condition, note that the value $V_{\text{ind}}^{\nu}(Z^1)$ is equal to the value $V_{\text{ind}}^{\nu}(Z^2)$. The analog applies to the case of a unit-linked life insurance contract with guarantee. The given distribution ν is defined by (3.6.1) and is calculated via the values q_x given in the Austrian annuity table 2005, which was presented in [37].

Example 3.6.14 (Without Guarantee)**.** We want to compute the price for a unit-linked life insurance contract for a 20-year old male person with different maturities. For a sample calculation we assume now that $Z_t^1 \sim U(0,2)$ for all $t \in I$, cf. Example 3.6.2 . Then we have that $\mathbb{E}[Z_t^1] = 1$ for all $t \in I$. Furthermore, let $Z_t^2 \sim \text{Log}\mathcal{N}(\mu, \sigma^2)$ for all $t \in I$ with $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$ such that $\mathbb{E}[Z_t^2] = \exp\left(\mu + \frac{\sigma^2}{2}\right) = 1$ for all $t \in I$, cf. Example 3.6.5. Thus, the required condition $\mathbb{E}[Z_t^1] = \mathbb{E}[Z_t^2]$ is fulfilled for all $t \in I$.

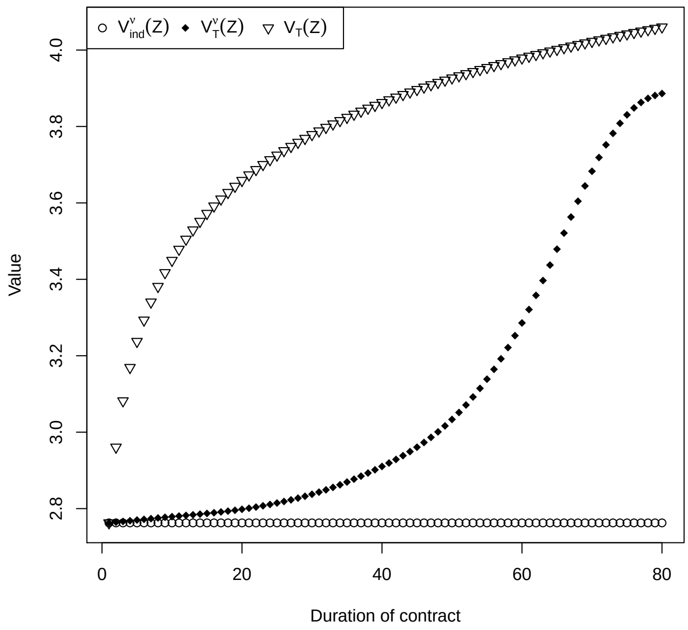


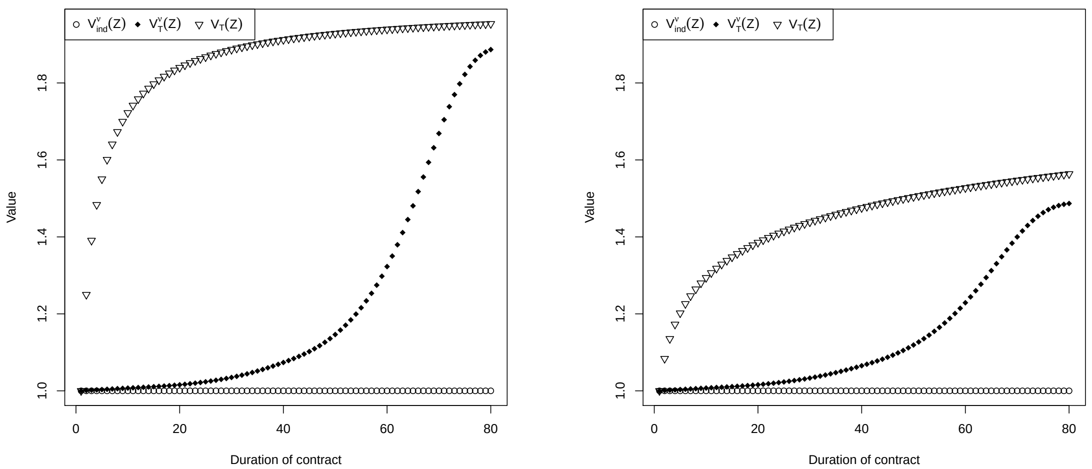
Figure 3.10.: The values $V_{\mathcal{T}}(Z)$, $V^{\nu}_{\mathcal{T}}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ for a unit-linked life insurance contract with guarantee 0.9 for a 20-year old male person for different maturities with a log-normally distributed underlying process with $\sigma = 0.18$ and $\mu = 1$.

We get the following values:

T		10	20	40	60	80
Uniform	$V^{\nu}_{\text{ind}}(Z^1)$	1.0000	1.0000	1.0000	1.0000	1.0000
	$V^{\nu}_{\mathcal{T}}(Z^1)$	1.0071	1.0156	1.0736	1.3229	1.8866
	$V_{\mathcal{T}}(Z^1)$	1.7222	1.8398	1.9122	1.9393	1.9536
Log-normal	$V^{\nu}_{\text{ind}}(Z^2)$	1.0000	1.0000	1.0000	1.0000	1.0000
	$V^{\nu}_{\mathcal{T}}(Z^2)$	1.0072	1.0158	1.0651	1.2288	1.4867
	$V_{\mathcal{T}}(Z^2)$	1.2938	1.3853	1.4755	1.5273	1.5636

In the following Figure 3.11, the corresponding results are shown graphically.

Example 3.6.15 (With Guarantee $G = 1$)**.** Now, we want to compute the price for a unit-linked life insurance contract with guarantee $G = 1$ for a 20-year old male person with different maturities. For a sample calculation we assume that $S^2_t \sim \text{Log}\,\mathcal{N}(\mu, \sigma^2)$ for all $t \in I$ with $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$, cf. Example 3.6.11, and $S^1_t \sim U(0, b_t)$ for all $t \in I$, cf. Example 3.6.6. In order to fulfill the required condition $\mathbb{E}[Z^1_t] = \mathbb{E}[Z^2_t]$ for all $t \in I$, we have to choose b_t as a solution of the equation $0 = b_t^2 - 2b_t\mathbb{E}[Z^2_t] + 1$.



(a) Uniformly distributed underlying process on $[0,2]$.

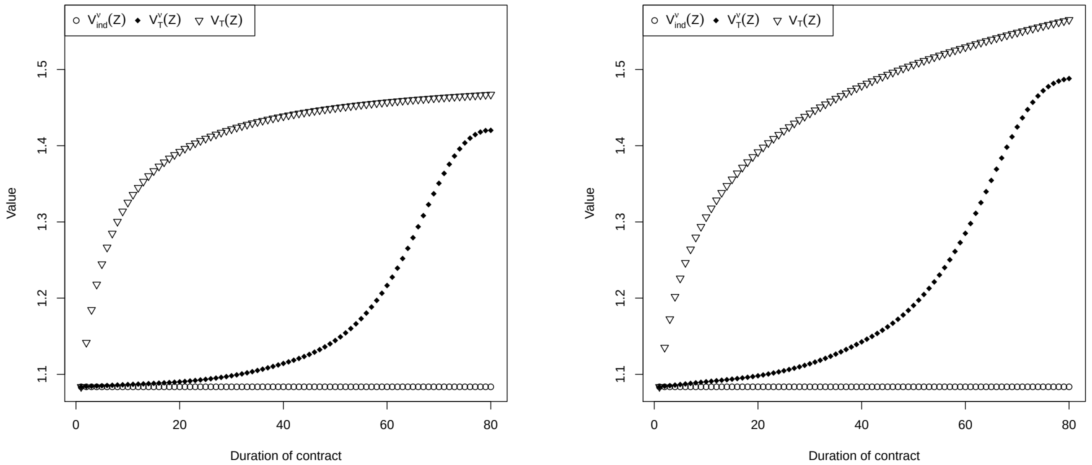
(b) Log-normally distributed underlying process with $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$.

Figure 3.11.: The values $V_{\mathcal{T}}(Z)$, $V^{\nu}_{\mathcal{T}}(Z)$ and $V^{\nu}_{\text{ind}}(Z)$ for a unit-linked life insurance contract without a guarantee for a 20-year old male person for different maturities.

We get the following values:

T		10	20	40	60	80
Uniform	$V^{\nu}_{\text{ind}}(Z^1)$	1.0836	1.0836	1.0836	1.0836	1.0836
	$V^{\nu}_{\mathcal{T}}(Z^1)$	1.0866	1.0901	1.1143	1.2166	1.4202
	$V_{\mathcal{T}}(Z^1)$	1.3255	1.3923	1.4387	1.4572	1.4672
Log-normal	$V^{\nu}_{\text{ind}}(Z^2)$	1.0836	1.0836	1.0836	1.0836	1.0836
	$V^{\nu}_{\mathcal{T}}(Z^2)$	1.0903	1.0981	1.1426	1.2850	1.4881
	$V_{\mathcal{T}}(Z^2)$	1.3064	1.3917	1.4786	1.5294	1.5652

In the following Figure 3.12, the corresponding results are shown graphically.



(a) Uniformly distributed underlying process on $[0, 1.501047]$.

(b) Log-normally distributed underlying process with $\sigma = 0.21$ and $\mu = -\frac{\sigma^2}{2}$.

Figure 3.12.: The values $V_{\mathcal{T}}(Z)$, $V_{\mathcal{T}}^{\nu}(Z)$ and $V_{\text{ind}}^{\nu}(Z)$ for a unit-linked life insurance contract without guarantee $G = 1$ for a 20-year old male person for different maturities.

3.6.4. Examples of Strategies

Now, we want to consider a model for insurance contracts in which the random variables of the underlying process are not independent such that we have not a formula for the value $V_{\mathcal{T}}^{\nu}(Z)$. We assume that we work with already discounted values to slightly simplify the considerations.

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be the given filtrated probability space and $I = \{0, \dots, T\}$ the observed discrete time interval with $T \in \mathbb{N}$. The given distribution ν is defined by (3.6.1) and is calculated via the values q_x for a 20-year old male person given in the Austrian annuity table 2005, which was presented in [37].

Furthermore let $X = (X_t)_{t \in I}$ be a process of independent and log-normally distributed random variables with expected value 1, such that the starting value $X_0 = 1$ and $X_t \sim \operatorname{Log}\mathcal{N}(\mu_t, \sigma_t^2)$ with $-2\mu_t = \sigma_t^2$ for each $t \in I \setminus \{0\}$. Note that $\mathbb{E}[X_t] = \exp(\mu_t + \sigma_t^2/2) = 1$, if $-2\mu_t = \sigma_t^2$ for all $t \in I \setminus \{0\}$. The filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in I}$ is given by $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_t = \sigma(X_1, X_2, \dots, X_t)$ for each $t \in I \setminus \{0\}$.

Then we consider the development $S = (S_t)_{t \in I}$ of the underlying fund which is modeled as product of the random variables of the process X. This means that the process S is defined by

$$S_t = X_0 \cdot \ldots \cdot X_t \qquad \text{for each } t \in I.$$

In this model the process S could be recursively rewritten as $S_0 = X_0 = 1$ and $S_t = X_t \cdot S_{t-1}$ for all $t \in I \setminus \{0\}$.

Remark 3.6.16 (The underlying process). The underlying process $S = (S_t)_{t \in I}$ is modeled as a product of independent, log-normally distributed random variables X_t, $t \in I$, with expected value 1. By Lemma A.4.5 we know for every $t \in I \setminus \{0\}$ with $\mu = (\mu_1, \dots, \mu_t) \in \mathbb{R}^t$,

$C = \operatorname{diag}(\sigma_1^2, \dots, \sigma_t^2) \in \mathbb{R}^{t\times t}$ and $p = (1, \dots, 1) \in \mathbb{R}^t$ that

$$S_t = \prod_{u=0}^{t} X_u \sim \operatorname{Log}\mathcal{N}\left(\sum_{u=1}^{t} \mu_u, \sum_{u=1}^{t} \sigma_u^2\right).$$

Thus we have that $S_t \sim \operatorname{Log}\mathcal{N}(\bar{\mu}_t, \bar{\sigma}_t^2)$ with $\bar{\mu}_t = \sum_{u=1}^t \mu_u$ and $\bar{\sigma}_t^2 = \sum_{u=1}^t \sigma_u^2$ for all $t \in I \setminus \{0\}$. Furthermore, it follows that $\mathbb{E}[S_t] = \exp\left(\bar{\mu}_t + \bar{\sigma}_t^2/2\right) = 1$ for all $t \in I \setminus \{0\}$. That means, the underlying process consists of random variables that are also log-normally distributed with expected value 1, but in this case *not* independent and *not* identically distributed. In particular, the process S is a martingale and the discrete version of a geometric Brownian motion.

Additionally, let $(G_t)_{t\in I}$ be the deterministic process of the guarantee. Then the payoff of the insurance contract, which we will denote by $Z = (Z_t)_{t\in I}$, at each time point $t \in I$ is defined as

$$Z_t = \max\{G_t, S_t\} = S_t + (G_t - S_t)^+.$$

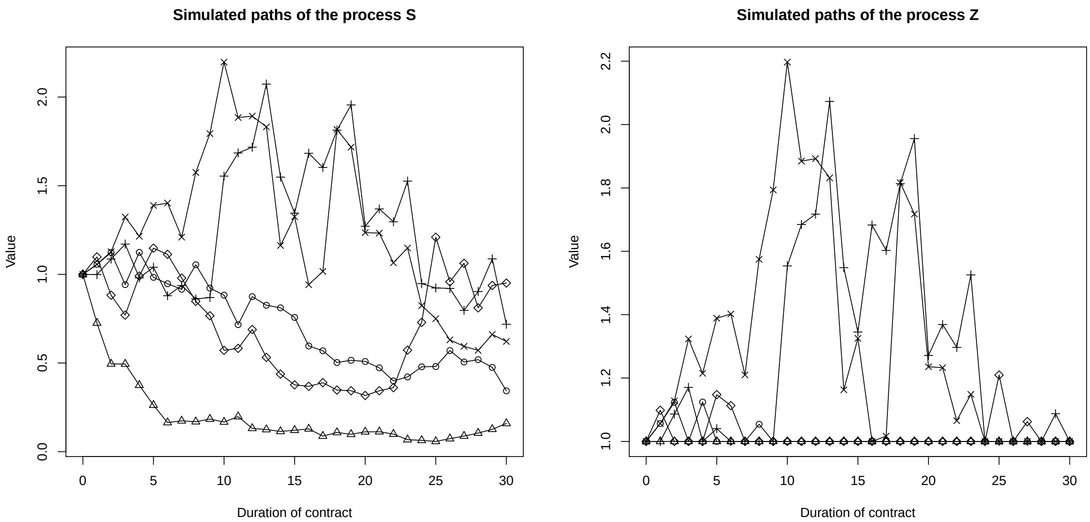


Figure 3.13.: Simulated paths of the process S and the corresponding ones of the process Z with $\sigma_t = 0.21$, $-2\mu_t = \sigma_t^2$ and $G_t = 1$ for all $t \in I$.

In Figure 3.13, five simulated paths of the underlying process S and the corresponding ones of Z are shown with constant guarantee $G_t = 1$ for all $t \in I$. We are interested in the value $V_{\mathcal{T}}^{\nu}(Z)$, $V_{\mathcal{T}}(Z)$ and $V_{\text{ind}}^{\nu}(Z)$. For the value $V_{\text{ind}}^{\nu}(Z)$, we can use the results of Example 3.6.11, because of the above considerations. Then we get the following:

Computation of $V_{\text{ind}}^{\nu}(Z)$: Under the assumption from above and the independence of the process Z and the stopping time τ or the adapted random probability measure γ we have that

$$V_{\text{ind}}^{\nu}(Z) = \sum_{t=1}^{T} \mathbb{E}[Z_t]\nu_t$$

for all maturities $T \in \mathbb{N}$. This payoff is represented as the sum of the fund and the value of a put option. Using (3.6.12) and Remark 3.6.16, then the expected value of Z_t is given for each $t \in I$ by

$$\begin{aligned}\mathbb{E}[Z_t] &= \mathbb{E}[\max\{S_t, G_t\}] = \mathbb{E}[G_t \mathbb{1}_{\{S_t \leq G_t\}}] + \mathbb{E}[S_t \mathbb{1}_{\{S_t > G_t\}}] \\ &= G_t \cdot \Phi\left(\frac{\ln(G_t) - \bar{\mu}_t}{\bar{\sigma}_t}\right) + \exp\left(\frac{\bar{\sigma}_t^2}{2} + \bar{\mu}_t\right) \cdot \Phi\left(-\frac{\ln(G_t) - \bar{\mu}_t}{\bar{\sigma}_t} + \bar{\sigma}_t\right).\end{aligned}$$

Computation of $V_T^\nu(Z)$: Due to the construction of S, the variables of Z are also dependent. Therefore we can not use the results of Example 3.6.11. Instead, we want to find a suitable strategy that closely approximates the value $V_T^\nu(Z)$.
As we have already seen, we can write the payoff as the sum of the fund and the payoff of a put option. We will denote this payoff of the corresponding put option by $P = (P_t)_{t \in I}$ with

$$P_t = (G_t - S_t)^+, \quad t \in I.$$

Remark 3.6.17 (Put Option). By construction the process P is a submartingale.

(a) We assume that the process P is given in such a way that the processes M and A of the Doob decomposition satisfy the necessary conditions. Then we have for each $t \in I$ that $P_t = M_t + A_t$ with a martingale M and a predictable, increasing process $A = (A_t)_{t \in I}$ starting at $A_0 = 0$. This knowledge will influence the choice of the strategies.

(b) Let $(G_t)_{t \in I}$ be a non-decreasing process. Then, in addition, for all $t \in I \setminus \{T\}$ we have that

$$P_t = (G_t - S_t)^+ \leq \mathbb{E}[(G_t - S_{t+1})^+ | \mathcal{F}_t] \leq \mathbb{E}[(G_{t+1} - S_{t+1})^+ | \mathcal{F}_t], \quad \text{a.s.} \tag{3.6.18}$$

Furthermore, we have that $\mathbb{E}[Z_\tau] = \mathbb{E}[S_\tau] + \mathbb{E}[P_\tau] = 1 + E[P_\tau]$, such that we can reduce the problem of finding an optimal strategy for the payoff of the process Z to find an optimal one for the payoff of the corresponding put option.

Computation of $V_T^\nu(P)$: Now, we want to find a suitable strategy that closely approximates the value $V_T^\nu(P)$. To get such an lower bound for the value $V_T^\nu(P)$, we simulated $n_{\text{sim}} \in \mathbb{N}$ different possible paths of S for any fixed maturity $T \in \mathbb{N}$ and consider different strategies. The motivation behind the choice of strategies will be explained later in the detailed consideration of the respective strategy. The different strategies should be used to stop the paths of P in a proper way. Iteratively starting from time 0, at every time $t \in I$, those remaining simulated paths with the smallest values determined by strategies. We will consider the following strategies:

1. $B_t = \mathbb{E}[P_{t+1} | \mathcal{F}_t] - P_t$
2. $C_t = \mathbb{E}[P_T | \mathcal{F}_t] - P_t$
3. $D_t = \sum_{u=t+1}^T \frac{\nu_u}{1-\nu_0-\ldots-\nu_t} (\mathbb{E}[P_u | \mathcal{F}_t] - P_t)$

The share of stopped paths at $t \in I$ depends on the given probability ν_t.

Strategy 1: At every time $t \in I$ the part of the simulated paths with the smallest values of B_t will be stopped. The share of stopped paths depends on the probability of the stopping time τ.
The random variable B_t describes the difference between the expected value in the next time step under the given information and the current value at time t. We stop those paths with the smallest value. To clarify the motivation behind the chosen strategy, we use the Remark 3.6.17. Given the Doob decomposition of P, i.e., we have for each $t \in I$ that $P_t = M_t + A_t$. Then we can rewrite the strategy as

$$B_t = \mathbb{E}[P_{t+1}|\mathcal{F}_t] - P_t = \mathbb{E}[M_{t+1}|\mathcal{F}_t] - M_t + \mathbb{E}[A_{t+1}|\mathcal{F}_t] - A_t = A_{t+1} - A_t = \Delta A_{t+1}.$$

The process A is predictable and increasing, so that we stop those paths that have the smallest growth in the next time step.
The values B_t for every $t \in I$ are given by

$$\begin{aligned} B_t &= \mathbb{E}[P_{t+1}|\mathcal{F}_t] - P_t = \mathbb{E}[(G_{t+1} - S_{t+1})^+ | \mathcal{F}_t] - (G_t - S_t)^+ \\ &= \mathbb{E}[(G_{t+1} - S_t \cdot X_{t+1})^+ | \mathcal{F}_t] - (G_t - S_t)^+ \\ &= S_t \, \mathbb{E}\left[\left(\frac{G_{t+1}}{S_t} - X_{t+1}\right)^+ \Big| \mathcal{F}_t\right] - (G_t - S_t)^+ . \end{aligned}$$

Using Proposition A.2.7 we only have to observe the following function for the computation

$$H(K,t) := \mathbb{E}\Big[\big(K - X_t\big)^+\Big] \tag{3.6.19}$$

for all $K \in \mathbb{R}_+$ and $t \in I$. Furthermore, we obtain that

$$\mathbb{E}\Big[\big(K - X_t\big)^+\Big] = K\, \mathbb{E}[\mathbb{1}_{\{X_t \le K\}}] - \mathbb{E}[X_t \mathbb{1}_{\{X_t \le K\}}].$$

If we look at the two terms individually and use the given conditions, we get for each $t \in I$ that

$$\begin{aligned} \mathbb{E}[\mathbb{1}_{\{X_t \le K\}}] &= \mathbb{P}(X_t \le K) = \mathbb{P}(\exp(\sigma_t Y + \mu_t) \le K) \quad \text{with } Y \sim \mathcal{N}(0,1) \\ &= \mathbb{P}\Big(Y \le \frac{\ln(K) - \mu_t}{\sigma_t}\Big) = \Phi\Big(\frac{\ln(K) - \mu_t}{\sigma_t}\Big) \end{aligned} \tag{3.6.20}$$

and

$$\begin{aligned} \mathbb{E}[X_t \mathbb{1}_{\{X_t \le K\}}] &= \int_0^K x \cdot \frac{1}{\sqrt{2\pi}\sigma_t x} \exp\Big(-\frac{(\ln(x) - \mu_t)^2}{2\sigma_t^2}\Big) dx \\ &\quad \text{substitution: } y = \frac{\ln(x) - \mu_t}{\sigma_t}; \quad dy = \frac{dx}{\sigma_t x} \\ &= \frac{1}{\sqrt{2\pi}} \exp\Big(\frac{\sigma_t^2}{2} + \mu_t\Big) \int_{-\infty}^{\frac{\ln(K)-\mu_t}{\sigma_t}} \exp\Big(-\frac{(y - \sigma_t)^2}{2}\Big) dy \\ &\quad \text{if } X_t \sim \operatorname{Log}\mathcal{N}(\mu_t, \sigma_t^2), \text{ then } Y = \frac{\ln(X_t) - \mu_t}{\sigma_t} \sim \mathcal{N}(0,1) \\ &= \exp\Big(\frac{\sigma_t^2}{2} + \mu_t\Big) \Phi\Big(\frac{\ln(K) - \mu_t}{\sigma_t} - \sigma_t\Big). \end{aligned} \tag{3.6.21}$$

Hence, the formula (3.6.19) could be written as

$$H(K,t) = K\cdot\Phi\Big(\frac{\ln(K)-\mu_t}{\sigma_t}\Big) - \exp\Big(\frac{\sigma_t^2}{2}+\mu_t\Big)\Phi\Big(\frac{\ln(K)-\mu_t}{\sigma_t}-\sigma_t\Big). \tag{3.6.22}$$

Finally, it follows for every $t\in I$ that

$$\begin{aligned}
B_t &= S_t\cdot H\left(\frac{G_{t+1}}{S_t}, t+1\right) - (G_t-S_t)^+\\
&= -(G_t-S_t)^+ + G_{t+1}\cdot\Phi\Big(\frac{\ln(G_{t+1}/S_t)-\mu_{t+1}}{\sigma_{t+1}}\Big)\\
&\quad - S_t\cdot\exp\Big(\frac{\sigma_{t+1}^2}{2}+\mu_{t+1}\Big)\cdot\Phi\Big(\frac{\ln(G_{t+1}/S_t)-\mu_{t+1}}{\sigma_{t+1}}-\sigma_{t+1}\Big).
\end{aligned}$$

Note that $\mathbb{E}[X_t] = \exp\big(\mu_t+\sigma_t^2/2\big) = 1$, if $-2\mu_t = \sigma_t^2$ for all $t\in I$.

Strategy 2: At every time $t\in I$ the part of the simulated paths with the smallest values of $C_t = \mathbb{E}[P_T|\mathcal{F}_t] - P_t$ will be stopped. The share of stopped paths depends on the probability of the stopping time τ.
The random variable C_t describes the difference between the expected value at the maturity under the given information at time t and the current value at t. We stop those paths with the smallest value. Given the Doob decomposition of P, we stop those paths that have the smallest expected growth of A until the maturity.
The values C_t for all $t\in I$ are given by

$$\begin{aligned}
C_t &= \mathbb{E}[P_T|\mathcal{F}_t] - P_t = \mathbb{E}[(G_T-S_T)^+|\,\mathcal{F}_t] - (G_t-S_t)^+\\
&= \mathbb{E}[(G_T - S_t\cdot X_{t+1}\cdot\ldots\cdot X_T)^+|\,\mathcal{F}_t] - (G_t-S_t)^+\\
&= S_t\cdot\mathbb{E}\Big[\Big(\frac{G_T}{S_t} - X_{t+1}\cdot\ldots\cdot X_T\Big)^+\Big|\,\mathcal{F}_t\Big] - (G_t-S_t)^+.
\end{aligned}$$

Analogously to Strategy 1, by using Proposition A.2.7 we observe initially the following function

$$\tilde{H}(K,t) := \mathbb{E}\Big[\Big(K - \prod_{u=t+1}^T X_u\Big)^+\Big] \tag{3.6.23}$$

for all $K\in\mathbb{R}_+$ and $t\in I$. Furthermore, we obtain that

$$\mathbb{E}\Big[\Big(K-\prod_{u=t+1}^T X_u\Big)^+\Big] = K\,\mathbb{E}\Big[\mathbb{1}_{\{\prod_{u=t+1}^T X_u\le K\}}\Big] - \mathbb{E}\Big[\prod_{u=t+1}^T X_u\mathbb{1}_{\{\prod_{u=t+1}^T X_u\le K\}}\Big].$$

By Lemma A.4.5 we know with $\mu = (\mu_{t+1},\ldots,\mu_T)\in\mathbb{R}^{T-t}$, $C = \operatorname{diag}(\sigma_{t+1}^2,\ldots,\sigma_T^2)\in\mathbb{R}^{(T-t)\times(T-t)}$ and $p = (1,\ldots,1)\in\mathbb{R}^{T-t}$ that

$$\tilde{X}_{t+1} = \prod_{u=t+1}^T X_u \sim \operatorname{Log}\mathcal{N}\Big(\underbrace{\sum_{u=t+1}^T \mu_u}_{=\tilde{\mu}_{t+1}}, \underbrace{\sum_{u=t+1}^T \sigma_u^2}_{=\tilde{\sigma}_{t+1}^2}\Big).$$

With the computation of Strategy 1 we get that for each $t \in I$

$$\mathbb{E}[\mathbb{1}_{\{\tilde{X}_{t+1} \leq K\}}] = \mathbb{P}(\tilde{X}_{t+1} \leq K) = \Phi\Big(\frac{\ln(K) - \tilde{\mu}_{t+1}}{\tilde{\sigma}_{t+1}}\Big)$$

and

$$\mathbb{E}[\tilde{X}_{t+1}\mathbb{1}_{\{\tilde{X}_{t+1} \leq K\}}] = \exp\Big(\frac{\tilde{\sigma}^2_{t+1}}{2} + \tilde{\mu}_{t+1}\Big)\Phi\Big(\frac{\ln(K) - \tilde{\mu}_{t+1}}{\tilde{\sigma}_{t+1}} - \tilde{\sigma}_{t+1}\Big).$$

Finally, the formula (3.6.23) could be written as

$$\tilde{H}(K,t) = K \cdot \Phi\Big(\frac{\ln(K) - \tilde{\mu}_{t+1}}{\tilde{\sigma}_{t+1}}\Big) - \exp\Big(\frac{\tilde{\sigma}^2_{t+1}}{2} + \tilde{\mu}_{t+1}\Big)\Phi\Big(\frac{\ln(K) - \tilde{\mu}_{t+1}}{\tilde{\sigma}_{t+1}} - \tilde{\sigma}_{t+1}\Big). \tag{3.6.24}$$

Finally, we get

$$\begin{aligned} C_t &= S_t \cdot \mathbb{E}\Big[\Big(\frac{G_T}{S_t} - \prod_{u=t+1}^{T} X_u\Big)^+ \Big| \mathcal{F}_t\Big] - (G_t - S_t)^+ \\ &= -(G_t - S_t)^+ + G_T \cdot \Phi\Big(\frac{\ln(G_T/S_t) - \tilde{\mu}_{t+1}}{\tilde{\sigma}_{t+1}}\Big) \\ &\quad - S_t \cdot \exp\Big(\frac{\tilde{\sigma}^2_{t+1}}{2} + \tilde{\mu}_{t+1}\Big) \cdot \Phi\Big(\frac{\ln(G_T/S_t) - \tilde{\mu}_{t+1}}{\tilde{\sigma}_{t+1}} - \tilde{\sigma}_{t+1}\Big). \end{aligned}$$

If $-2\mu_t = \sigma_t^2$ for all $t \in I$, then $\mathbb{E}[X_t] = \exp\big(\mu_t + \sigma_t^2/2\big) = 1$ and it also follows $\mathbb{E}[\tilde{X}_{t+1}] = 1$ for every $t \in I$, because of the independence of the X_t's, $t \in I$.

Strategy 3: At every time $t \in I$ the part of the simulated paths with the smallest values of $\sum_{u=t+1}^{T} \frac{\nu_u}{1-\nu_{\leq t}}(\mathbb{E}[P_u|\mathcal{F}_t] - P_t)$ will be stopped. The share of stopped paths depends on the probability of the stopping time τ.
The random variable D_t describes the weighted sum of the differences between the expected value at a time point $u \in \{t+1,\dots,T\}$ under the given information at time t and the current value at t. We stop those paths with the smallest value. Given the Doob decomposition of P, we stop those paths that have the smallest expected growth of A until the maturity.
The values $\mathbb{E}[P_u|\mathcal{F}_t] - P_t$ for all $t \in I$ and for all $u \in \{t+1,\dots,T\}$ are given analogously to Strategy 2 by

$$\begin{aligned} \mathbb{E}[P_u|\mathcal{F}_t] - P_t &= \mathbb{E}[(G_u - S_u)^+ \mid \mathcal{F}_t] - (G_t - S_t)^+ \\ &= \mathbb{E}[(G_u - S_t \cdot X_{t+1} \cdot \ldots \cdot X_u)^+ \mid \mathcal{F}_t] - (G_t - S_t)^+ \\ &= S_t \cdot \mathbb{E}\Big[\Big(\frac{G_u}{S_t} - X_{t+1} \cdot \ldots \cdot X_u\Big)^+ \Big| \mathcal{F}_t\Big] - (G_t - S_t)^+. \end{aligned}$$

Analogously to Strategy 1 and 2 we observe initially the following function

$$\check{H}(K,t,u) := \mathbb{E}\Big[\Big(K - \prod_{j=t+1}^{u} X_j\Big)^+\Big] \tag{3.6.25}$$

for all $K \in \mathbb{R}_+$, $t \in I$ and $u \in \{t+1,\dots,T\}$, because of Proposition A.2.7. Furthermore, by Lemma A.4.5 we know with $\mu = (\mu_{t+1},\dots,\mu_u) \in \mathbb{R}^{u-t}$, $C = \operatorname{diag}(\sigma_{t+1}^2,\dots,\sigma_u^2) \in \mathbb{R}^{(u-t)\times(u-t)}$ and $p = (1,\dots,1) \in \mathbb{R}^{u-t}$ for all $u \in \{t+1,\dots,T\}$ that

$$\breve{X}_{t+1,u} = \prod_{j=t+1}^{u} X_j \sim \operatorname{Log}\mathcal{N}\Big(\underbrace{\sum_{j=t+1}^{u}\mu_j}_{=\breve{\mu}_{t+1,u}}, \underbrace{\sum_{j=t+1}^{u}\sigma_j^2}_{=\breve{\sigma}_{t+1,u}^2}\Big).$$

With the similar consideration as for Strategy 2 the formula (3.6.25) could finally be written as

$$\breve{H}(K,t,u) = K\cdot\Phi\Big(\frac{\ln(K)-\breve{\mu}_{t+1,u}}{\breve{\sigma}_{t+1,u}}\Big) - \exp\Big(\frac{\breve{\sigma}_{t+1,u}^2}{2}+\breve{\mu}_{t+1,u}\Big)\Phi\Big(\frac{\ln(K)-\breve{\mu}_{t+1,u}}{\breve{\sigma}_{t+1,u}} - \breve{\sigma}_{t+1,u}\Big) \tag{3.6.26}$$

for all $K \in \mathbb{R}_+$, $t \in I$ and $u \in \{t+1,\dots,T\}$.
If $-2\mu_t = \sigma_t^2$ for all $t \in I$, then $\mathbb{E}[X_t] = \exp\big(\mu_t + \sigma_t^2/2\big) = 1$ and it also follows $\mathbb{E}[\breve{X}_{t+1,u}] = 1$ for every $t \in I$, because of the independence of the X_t's, $t \in I$. Finally, we get

$$\begin{aligned}
\mathbb{E}[P_u|\mathcal{F}_t] - P_t &= S_t\cdot\mathbb{E}\Big[\Big(\frac{G_u}{S_t} - \prod_{j=t+1}^{u} X_j\Big)^+\Big|\,\mathcal{F}_t\Big] - (G_t - S_t)^+ \\
&= -(G_t-S_t)^+ + G_u\cdot\Phi\Big(\frac{\ln(G_u/S_t)-\breve{\mu}_{t+1,u}}{\breve{\sigma}_{t+1,u}}\Big) \\
&\quad - S_t\cdot\exp\Big(\frac{\breve{\sigma}_{t+1,u}^2}{2}+\breve{\mu}_{t+1,u}\Big)\cdot\Phi\Big(\frac{\ln(G_u/S_t)-\breve{\mu}_{t+1,u}}{\breve{\sigma}_{t+1,u}} - \breve{\sigma}_{t+1,u}\Big)
\end{aligned}$$

for all $t \in I$ and $u \in \{t+1,\dots,T\}$. It holds that

$$\begin{aligned}
D_t &= \sum_{u=t+1}^{T}\frac{\nu_u}{1-\nu_0-\ldots-\nu_t}\big(\mathbb{E}[P_u|\mathcal{F}_t]-P_t\big) \\
&= -(G_t-S_t)^+ + \sum_{u=t+1}^{T}\frac{\nu_u}{1-\nu_{\le t}}G_u\,\Phi\Big(\frac{\ln(G_u/S_t)-\breve{\mu}_{t+1,u}}{\breve{\sigma}_{t+1,u}}\Big) \\
&\quad - S_t\cdot\sum_{u=t+1}^{T}\frac{\nu_u}{1-\nu_{\le t}}\exp\Big(\frac{\breve{\sigma}_{t+1,u}^2}{2}+\breve{\mu}_{t+1,u}\Big)\Phi\Big(\frac{\ln(G_u/S_t)-\breve{\mu}_{t+1,u}}{\breve{\sigma}_{t+1,u}} - \breve{\sigma}_{t+1,u}\Big).
\end{aligned}$$

Sample calculation: From now on and for our later simulation, we will assume the following:

Assumption 3.6.27 (Simulation)**.** For all $t \in I$ we assume that

(a) the guaranteed value should be fixed over time, i.e., $G_t = G$ for all $t \in I$ and a constant $G \in \mathbb{R}_+$.

(b) $\sigma_t = \sigma$ with a constant $\sigma \in \mathbb{R}_+$ and $-2\mu_t = \sigma_t^2$ for the log-normally distributed random variables.

Using Remark 3.6.16 and Assumption (b), we have that $S_t \sim \text{Log}\mathcal{N}(\mu t, t\sigma^2)$ for all $t \in I \setminus \{0\}$. Furthermore, it follows that $\mathbb{E}[S_t] = \exp\left(t\mu + t\sigma^2/2\right) = 1$ for all $t \in I \setminus \{0\}$. The payoff of the insurance contract is represented as the sum of the fund and the payoff of a put option. By Assumption (a) we have that $Z_t = S_t + (G - S_t)^+$ and $P_t = (G - S_t)^+$ for all $t \in I$. As we already known, we can reduce the problem of finding an optimal strategy for the payoff of the process Z by finding an optimal one for the payoff of the corresponding put option, such that we are interested in the value $V_T^\nu(P)$, $V_T(P)$ and $V_{\text{ind}}^\nu(P)$.

Computation of $V_{\text{ind}}^\nu(P)$: Using the notation from above and the independence of the process P and the stopping time τ or the adapted random probability measure γ we have that

$$V_{\text{ind}}^\nu(P) = \sum_{t=1}^{T} \mathbb{E}[P_t]\nu_t$$

for all maturities $T \in \mathbb{N}$. Using Assumption 3.6.27, Remark 3.6.16 and (3.6.21) the expected value of P_t is given for each $t \in I$ by

$$\begin{aligned}\mathbb{E}[P_t] &= \mathbb{E}[(G - S_t)^+] = \mathbb{E}[(G - S_t)\mathbb{1}_{\{S_t \le G\}}] \\ &= G \cdot \Phi\left(\frac{\ln(G) - t\mu}{\sqrt{t}\sigma}\right) - \exp\left(\frac{t\sigma^2}{2} + t\mu\right)\Phi\left(\frac{\ln(G) - t\mu}{\sqrt{t}\sigma} - \sqrt{t}\sigma\right).\end{aligned}$$

Furthermore, we have that $V_{\text{ind}}^\nu(Z) = 1 + V_{\text{ind}}^\nu(P)$. In the simulation, we will also determine $V_{\text{ind}}^\nu(\hat{P})$ which is the estimated value from the data. For this, the expected value of P_t, $t \in I$, is calculated as arithmetic mean.

Computation of $V_T^\nu(P)$: Using Assumption 3.6.27 we get the following formulas for the different strategies.
For all $t \in I \setminus \{0\}$, we get for Strategy 1 that

$$B_t = G \cdot \Phi\left(\frac{2\ln(G/S_t) + \sigma^2}{2\sigma}\right) - S_t \cdot \Phi\left(\frac{2\ln(G/S_t) - \sigma^2}{2\sigma}\right) - (G - S_t)^+. \tag{3.6.28}$$

Using Assumption (b), we have for all $t \in I \setminus \{0\}$ that $\tilde{\sigma}_{t+1}^2 = (T-t)\sigma^2$ and $\tilde{\mu}_{t+1} = -\sigma^2(T-t)/2$ and thus for Strategy 2 that

$$C_t = G \cdot \Phi\left(\frac{2\ln(G/S_t) + (T-t)\sigma^2}{2\sigma\sqrt{T-t}}\right) - S_t \cdot \Phi\left(\frac{2\ln(G/S_t) - (T-t)\sigma^2}{2\sigma\sqrt{T-t}}\right) - (G - S_t)^+. \tag{3.6.29}$$

Using Assumption (b), we have for all $t \in I \setminus \{0\}$ that $\breve{\mu}_{t+1,u} = -\sigma^2(u-t)/2$ and $\breve{\sigma}_{t+1,u}^2 = (u-t)\sigma^2$ and thus for Strategy 3 that

$$\begin{aligned}D_t = &-(G - S_t)^+ + G \cdot \sum_{u=t+1}^{T} \frac{\nu_u}{1 - \nu_{\le t}} \cdot \Phi\left(\frac{2\ln(G/S_t) + (u-t)\sigma^2}{2\sigma\sqrt{u-t}}\right) \\ &- S_t \cdot \sum_{u=t+1}^{T} \frac{\nu_u}{1 - \nu_{\le t}} \cdot \Phi\left(\frac{2\ln(G/S_t) - (T-t)\sigma^2}{2\sigma\sqrt{T-t}}\right).\end{aligned} \tag{3.6.30}$$

Computation of $V_T(P)$: Analogously to Example 3.6.11 the calculation results also from the Snell envelope $U = (U_t)_{t\in I}$ of P and its process of expected values $V = (V_t)_{t\in I}$. Keep in mind that the process P consists of dependent variables. Using the Definition of the Snell envelope given in (3.6.7) and $0 \le P_t = (G - S_t)^+ \le G$ for all $t \in I$, we have that $0 \le U_T = P_T \le G$. Furthermore, then it yields recursively for all $t \in \{0,\dots,T-1\}$ that

$$0 \le U_t = \max\Big\{P_t, \mathbb{E}[U_{t+1}|\mathcal{F}_t]\Big\} \le G,$$

because of $0 \le P_t \le G$ and $0 \le U_{t+1} \le G$ for the considered time point t. It follows also that $0 \le \mathbb{E}[U_{t+1}|\mathcal{F}_t] \le G$ for each $t \in I \setminus \{T\}$.
It holds for each $t \in I$ that $V_t = \mathbb{E}[U_t]$ and the value $V_T(P)$ is determined by $V_T(P) = V_0$. To calculate the expected value of U_t for all $t \in I$, we have to determine U_t itself.

(i) For $t = T$, we have that $U_T = P_T$ and

$$\mathbb{E}[U_T] = \mathbb{E}[P_T] = G\cdot\Phi\left(\frac{\ln(G) - T\mu}{\sqrt{T}\sigma}\right) - \exp\left(\frac{T\sigma^2}{2} + T\mu\right)\Phi\left(\frac{\ln(G) - T\mu}{\sqrt{T}\sigma} - \sqrt{T}\sigma\right).$$

(ii) For $t = T-1$, we have that $U_{T-1} = \max\Big\{P_{T-1}, \mathbb{E}[U_T|\mathcal{F}_{T-1}]\Big\}$, where

$$\begin{aligned}\mathbb{E}[U_T|\mathcal{F}_{T-1}] = \mathbb{E}[P_T|\mathcal{F}_{T-1}] &= S_{T-1}\,\mathbb{E}\left[\left(\frac{G}{S_{T-1}} - X_T\right)^+\Big|\mathcal{F}_{T-1}\right] = S_{T-1}\cdot H\left(\frac{G}{S_{T-1}}, T\right)\\ &= G\cdot\Phi\Big(\frac{\ln(G/S_{T-1}) - \mu_T}{\sigma_T}\Big) - S_{T-1}\cdot\Phi\Big(\frac{\ln(G/S_{T-1}) - \mu_T}{\sigma_T} - \sigma_T\Big).\end{aligned}$$

Using that P is a submartingale and (3.6.18) we get that $U_{T-1} = \mathbb{E}[U_T|\mathcal{F}_{T-1}]$.

(iii) For all $t < T-1$, we get iteratively that $U_t = \mathbb{E}[U_{t+1}|\mathcal{F}_t]$ because of the tower property of the conditional expected value and P is a submartingal. Using parts of the calulation in Strategy 2 and (3.6.24), we obtain the following:

$$\begin{aligned}\mathbb{E}[U_{t+1}|\mathcal{F}_t] = \mathbb{E}[\mathbb{E}[U_{t+2}|\mathcal{F}_{t+1}]|\mathcal{F}_t] &= \mathbb{E}[U_{t+2}|\mathcal{F}_t] = \ldots = \mathbb{E}[U_T|\mathcal{F}_t]\\ &= \mathbb{E}[(G - S_T)^+|\mathcal{F}_t] = S_t\cdot\mathbb{E}\left[\left(\frac{G}{S_t} - X_{t+1}\cdot\ldots\cdot X_T\right)^+\Big|\mathcal{F}_t\right] = S_t\cdot\breve{H}\Big(\frac{G}{S_t}, t, T\Big)\\ &= G\cdot\Phi\Big(\frac{\ln(G/S_t) - \breve{\mu}_{t+1,T}}{\breve{\sigma}_{t+1,T}}\Big) - S_t\cdot\Phi\Big(\frac{\ln(G/S_t) - \breve{\mu}_{t+1,T}}{\breve{\sigma}_{t+1,T}} - \breve{\sigma}_{t+1,T}\Big)\\ &= G\cdot\Phi\Big(\frac{2\ln(G/S_t) + (T-t)\sigma^2}{2\sigma\sqrt{T-t}}\Big) - S_t\cdot\Phi\Big(\frac{2\ln(G/S_t) - (T-t)\sigma^2}{2\sigma\sqrt{T-t}}\Big).\end{aligned}$$

Using the tower property of the conditional expectation again, we have that $(U_t)_{t\in I}$ is a martingale and $V_T(P) = V_0 = \mathbb{E}[U_0] = \mathbb{E}[U_T]$.

Finally, the results of the simulation are to be presented. We assume that the guaranteed value G is 1 and σ is 0.21 over time. The calculations were made for $n_{\text{sim}} = 100,000$ paths. Note that the simulated paths are generated at the beginning of the simulation and are the same for all calculations. Then we get the following values:

T	10	20	30	40	50	60	80
$V^{\nu}_{\text{ind}}(P)$	0.2595	0.3596	0.4313	0.4867	0.5302	0.5624	0.5835
$V^{\nu}_{\text{ind}}(\hat{P})$	0.2593	0.3588	0.4302	0.4865	0.5304	0.5614	0.5831
$V_{\mathcal{T}}(P)$	0.2601	0.3613	0.4348	0.4934	0.5422	0.5840	0.6523
Strategy 1	0.2598	0.3603	0.4334	0.4928	0.5416	0.5812	0.6407
Strategy 2	0.2598	0.3603	0.4334	0.4928	0.5419	0.5820	0.6428
Strategy 3	0.2598	0.3603	0.4333	0.4924	0.5361	0.5646	0.5827

The following figures additionally illustrate the results. Figure 3.14 shows the evolution of the approximated values by the different strategies for $V^{\nu}_{\mathcal{T}}(P)$, $V_{\mathcal{T}}(P)$ and $V^{\nu}_{\text{ind}}(P)$ for the different maturities. The differences are sometimes very small, especially between the calculated values through the strategies.

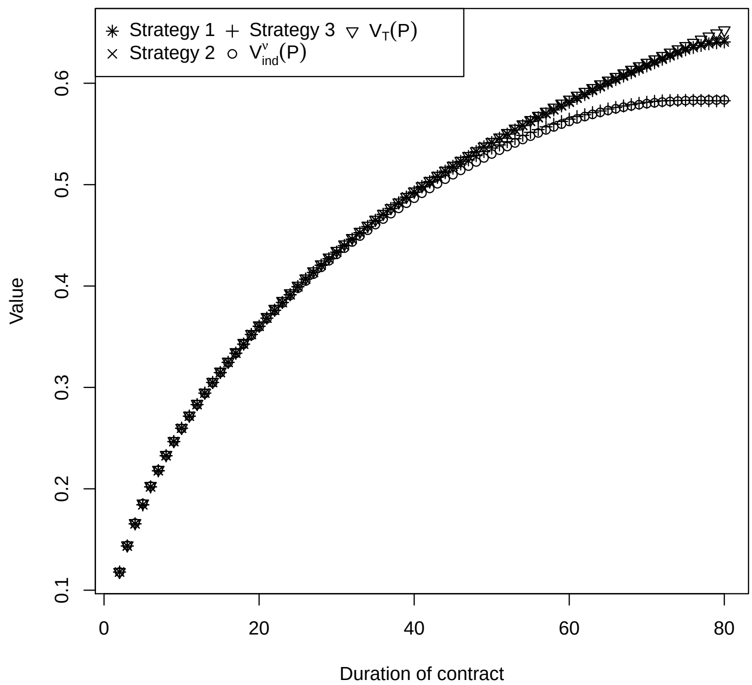


Figure 3.14.: Simulation with $\sigma_t = 0.21$, $-2\mu_t = \sigma_t^2$ and $G_t = 1$ for all $t \in I$.

With longer term of the contract, the differences become more visible. To make it more visible, the differences to the value $V_{\text{ind}}^{\nu}(P)$ were calculated and shown in Figure 3.15. Note that the value $V_{\text{ind}}^{\nu}(P)$ is explicit and does not depend on the data. Therefore, it is taken as a reference value. The value $V_{\text{ind}}^{\nu}(\hat{P})$ estimated from the data was calculated and given to evaluate the quality of the data and calculations. The differences vary, but they are low, see Figure 3.15 and Figure 3.16.

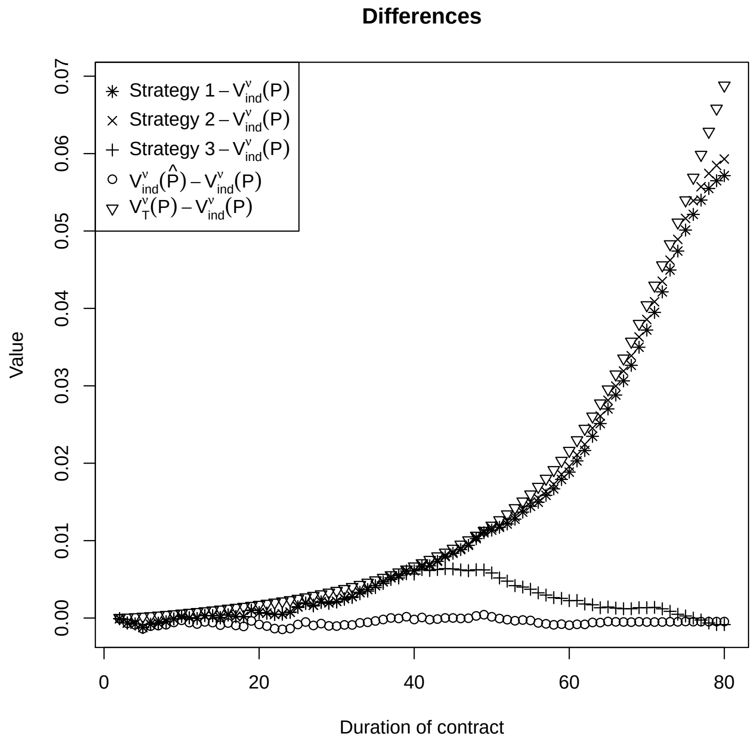


Figure 3.15.: Simulation with $\sigma_t = 0.21$, $-2\mu_t = \sigma_t^2$ and $G_t = 1$ for all $t \in I$.

The values of Strategy 1 and 2 behave relative to the value similar $V_{\text{ind}}^{\nu}(P)$ to the Example 3.6.11. The behavior of the third strategy does not meet the expectations. The highest values are thus obtained with the strategy that at every time $t \in I$ the part of the simulated paths with the smallest values of $\mathbb{E}[P_T|\mathcal{F}_t] - P_t$ will be stopped, see also Figure 3.17. We stop those paths with the smallest difference between the expected value at the maturity under the given information at time t and the current value at t, i.e., those paths that have the smallest expected growth until the maturity.

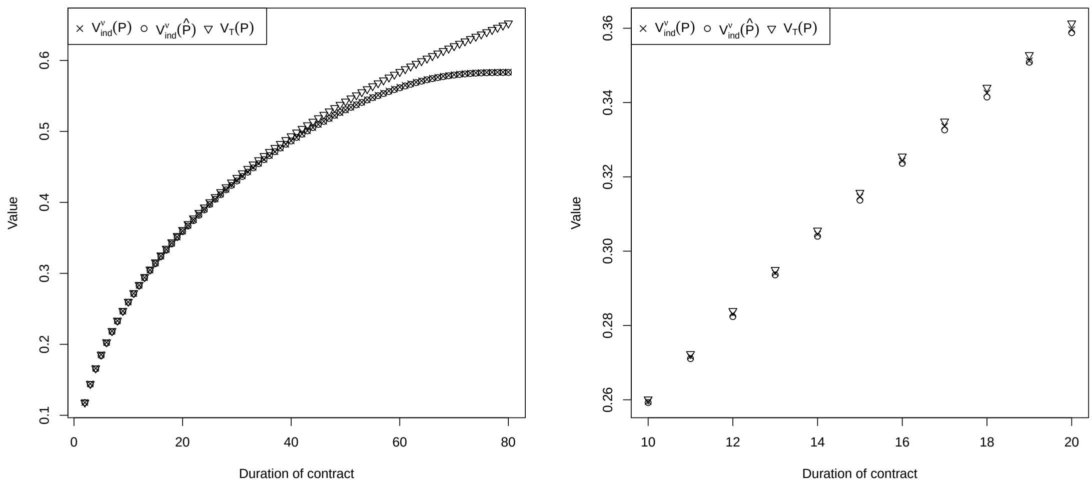


Figure 3.16.: Simulation with $\sigma_t = 0.21$, $-2\mu_t = \sigma_t^2$ and $G_t = 1$ for all $t \in I$.

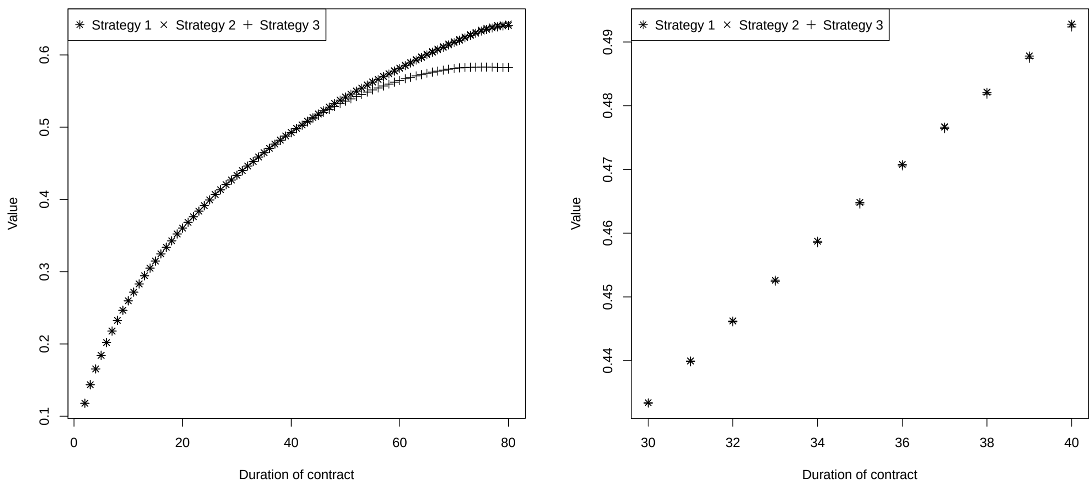


Figure 3.17.: Simulation with $\sigma_t = 0.21$, $-2\mu_t = \sigma_t^2$ and $G_t = 1$ for all $t \in I$.

4

Randomized Stopping Time

As we have already mentioned, there are many situations in financial and actuarial mathematics where it is questionable whether the assumed independence of two stochastic components is always justified. The recent research proves an increasing interest in the topic of distribution-constrained optimal stopping problems as in [8, 10, 9, 14, 38]. This chapter deals with another general framework to show how some of these situations can be handled without the assumption of independence. As seen, there is the view of the distribution-constrained optimal stopping problem in discrete time as a restricted optimization problem where we replace the stopping times by adapted random probability measures, see Chapter 3. Now, we want to consider another possibility. These problems can also be formulated as optimal transport problems based on the theory of optimal transport. The existence of optimizers can be shown, see Chapter 7 in Part II for the continuous time case and this chapter for the discrete time case.
We will follow the conventions of optimal transport and especially the one of [73]. To do so, we have to slightly reformulate our starting problem. An attempt of modeling the dependency can be made by using measures related to stochastic processes which mimic ordinary stopping times and are, in fact, a generalization of the latter. An informal way of posing this problem is the following:
Given a payoff function c, which may depend on the values of the stochastic process up to a time t and in the theory of optimal transport is called the cost function, we seek to maximize

$$\tau \mapsto \mathbb{E}[c((Z_t)_{t\leq\tau}, \tau)],$$

where τ is not an ordinary stopping time in the filtration generated by Z, i.e., it does not stop the process at a time $\tau(\omega)$, but rather $\tau(\omega)$ is a sub-probability measure on the time domain by itself. Essentially this can be formalized in three different ways:

(a) As an optimal stopping problem where adapted random probability measures are used instead of ordinary stopping times.

(b) As an optimal transport problem by reformulating it by means of randomized stopping times.

(c) As an ordinary optimal stopping problem on a larger probability space, cf. [13, Lemma 3.11].

It should be clear that this type of problems naturally arises from ordinary optimal stopping problems, where additional dependencies have to be modeled. In addition to the existence, questions about different geometric optimality criteria – so-called monotonicity principles – are of interest. If we deal with the theory of optimal transport, we come into contact with the two common basic concepts, cyclical monotonicity and Kantorovich duality. The cyclical monotonicity is a geometric property. An optimal plan should be c-cyclically monotone, i.e., it is concentrated on a c-cyclically monotone set and you can not improve the cost by rerouting mass along some cycle. It is impossible to perturb it and get something more economical. Informally, a c-cyclically monotone plan is a plan that can not be improved. The converse property is much less obvious, i.e., a c-cyclically monotone plan should be optimal. Maybe it is possible to get something better by radically changing the plan as only rerouting mass along some cycle. In this chapter we will see soon that it holds true under certain conditions.
Inspired by this classical c-monotonicity which shows that optimality is an attribute of the support of a coupling, other different monotonicity principles have been developed in the area of martingale optimal transport problems, cf. [9, 10]. To test if a randomized stopping time is a possible candidate for optimality in the considered problem, different monotonicity criteria were developed. In this context, the so-called c-cyclical monotonicity as in [73] deserves a special mention, which is in fact a geometric property of the support of an optimal transport plan. In the initial form the monotonicity was shown only for couplings which do not have to satisfy additional adaptivity constraints. Zaev introduced (c, W)-cyclical monotonicity in [76, Theorem 3.6], which enhances the notion with constraints, denoted by W. Contrary to the classical c-monotonicity, the (c, W)-monotonicity of a support of a randomized stopping time is a necessary optimality condition, but in general not sufficient. In independent work, Beiglböck and Griessler found a closely related monotonicity principle which includes the result [76, Theorem 3.6] as a special case, see [11, Theorem 1.4].
The remainder of this chapter is organized as follows: in Section 4.2 the maximization problems $\textsc{OptStop}^\gamma$ and $\textsc{OptStop}^\pi$ are formally introduced. We give the notions of adapted random probability measures ($\mathcal{M}_I$), couplings (Cpl) and randomized stopping times (RST), and explore their relations. The subsequent Subsection 3.2 especially draws the connection between $\textsc{OptStop}^\gamma$ and $\textsc{OptStop}^\pi$. The following considerations then focus on the problem $\textsc{OptStop}^\pi$. The existence of a maximizer of $\textsc{OptStop}^\pi$ is shown in Section 4.3 utilizing Prokhorov's Theorem and [13, Lemma 2.3]. Based on the theory of optimal transport and recent results [76] in this area, duality in the sense of Kantorovich is deduced in Section 4.4. In Section 4.5 examples are investigated and optimal maximizers are determined. Finally, Section 4.6 shortly sketches different monotonicity principles and uses one to show optimality of the maximizer introduced in the previous section. Some further considerations are mentioned.

This chapter was written in collaboration with Gudmund Pammer, who received his diploma thesis from it, see [57].

4.1. Notational Conventions

Since this topic can be considered separately, some important notations are repeated. In principle we will follow the conventions of [73].
In this chapter, we consider a discrete time domain. Its index set is again denoted by I. Typical examples for an infinite index set are $\mathbb{N}$ and for a finite $\{1,\ldots,T\}$, $T \in \mathbb{N}$. For better readability and simplification we choose $I = \{1,\ldots,T\}$ with $T \in \mathbb{N}$ and $T < \infty$ or $T = \infty$ which represents the finite or infinite index set. For $t \in I$ we define the set $I_{<t} := \{s \in I | s < t\}$ of all times before t, the set $I_{\leq t} := \{s \in I | s \leq t\}$ of all times up to t, the set $I_{\geq t} := \{s \in I | s \geq t\}$ of all times from t on, and the set $I_{>t} := \{s \in I | s > t\}$ of all times after t. We also use $I_{<t} = [0,t)$, $I_{\leq t} = [0,t]$ and $I_{>t} = (t,T]$ as a representation, where 0 symbolizes the starting point and T the maturity.
Given a topological space $(X,\mathcal{T})$, we denote its Borel-σ-algebra with $\mathcal{B}(X) = \sigma(\mathcal{T})$, the interior of a set $A \subseteq X$ with $\mathrm{int}(A)$ and its boundary with $\partial(A)$. The space of all Borel-measurable functions from X into $\mathbb{R}$ is denoted by $B(X)$ and its subspace of all bounded, Borel-measurable functions by $B_b(X)$.
Typically, we will work with (sub-)probability measures on the Polish space $\mathbb{R}^I$. To facilitate the notation of projections onto particular subspace of $\mathbb{R}^I$, which we may define as

$$\mathbb{R}^I =: \prod_{i \in I} X_i,$$

and for instance call the projection of a measure μ on $\mathbb{R}^I$ onto the first component $\mathrm{proj}_{X_1}(\mu)$.
Several different notations will be used to refer to elements of $\mathbb{R}^I$. For any vector $\omega \in \mathbb{R}^I$, its entries are denoted with

$$\omega = (\omega_t)_{t \in I} = (\omega_1, \omega_2, \ldots).$$

Parts of the vector (path) ω will be referred to by

$$(\omega_t)_{t \in I_{\leq s}} = \omega_{\upharpoonright[0,s]}, \quad (\omega_t)_{t \in I_{>s}} = \omega_{\upharpoonright(s,T]}, \quad s \in I,$$

where $\omega_{\upharpoonright J}$ with $J \subseteq I$ stands for the restriction of ω onto $\mathbb{R}^J$.
If $\omega \in \mathbb{R}^I$, $s \in I$ and $\theta \in \mathbb{R}^{I_{>s}}$, we may use $\oplus$ to indicate the concatenation of the paths $\omega_{\upharpoonright[0,s]}$ and θ, such that

$$\omega_{\upharpoonright[0,s]} \oplus \theta := (\omega_1, \ldots, \omega_s, \omega_s + \theta_1, \omega_s + \theta_2, \ldots) \in \mathbb{R}^I.$$

In the following $Z = (Z_t)_{t \in I}$ will denote a distinguished stochastic process. If Z is assumed to have independent increments, i.e., for any $t_1, \ldots, t_n \in I$ with $t_1 < \ldots < t_n$ the increments $Z_{t_1}, Z_{t_2} - Z_{t_1}, \ldots, Z_{t_n} - Z_{t_{n-1}}$ are independent. It is convenient to define $(p_i)_{i \in I}$ via

$$Z_t = Z_0 + \sum_{i \leq t} p_i,$$

where Z_0 is the initial distribution of the stochastic process Z. The measure induced by the process starting in 0, $\tilde{Z}_t := \sum_{i \leq t} p_i$, on $\mathbb{R}^I$ is denote by P.
For signed measures ξ there exists a Hahn-Jordan decomposition,

$$\xi = \xi^+ - \xi^-,$$

where ξ^+ and ξ^- are the positive and negative parts of ξ, respectively.

4.2. The Different Problems

Let $(\Omega, \mathcal{G}, \mathbb{G} := (\mathcal{G}_t)_{t\in I}, \mathbb{P})$ be an abstract filtered probability space and $Z := (Z_t)_{t\in I}$ be the stochastic, real-valued and $\mathbb{G}$-adapted process of interest. Further, let ν denote a (discrete) probability measure on $(I, \mathcal{B}(I))$. We assume that the process Z is uniformly integrable, i.e.,

$$\forall \varepsilon > 0,\ \exists \delta > 0: \quad \int_E |Z_t| \mathrm{d}\mathbb{P} < \varepsilon$$

whenever $Z_t \in L^1(\mathbb{P})$ for all $t \in I$ and $\mathbb{P}(E) < \delta$. Furthermore, we denote with μ the probability measure induced onto the measurable space $(\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I))$ by the stochastic process Z via

$$\mu(B) := Z_\# \mathbb{P}(B) \quad \forall B \in \mathcal{B}(\mathbb{R}^I),$$

and call the probability triplet $(\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I), \mu)$ the path space of Z. The payoff function c is assumed to be real-valued and Borel-measurable on S with

$$S := \{(x,t) \mid x \in \mathbb{R}^I,\ t \in I\}.$$

The space S is adequate for our purposes since for a given time t and path $(Z_s(\omega))_{s\le t} =: x$ up to the time t, the function c returns the payoff $c(x,t)$. Note that the space is Polish as it is the direct sum of Polish spaces. For example, the topology induced on S by the metric $d: S \times S \to \mathbb{R}$ defined as

$$((x_i)_{i\le s}, s), (y_i)_{i\le t}, t)) \mapsto \max\left(|t-s|, \max_{i\le \min(s,t)} (|x_i - y_i|)\right),$$

causes (S,d) to be Polish. Further, there exists a surjective, open, continuous map r with

$$\begin{aligned} r:\ (\mathbb{R}^I \times I, \mathcal{B}(\mathbb{R}^I \times I)) &\to (S,d), \\ ((x_s)_{s\in I}, t) &\mapsto ((x_s)_{s\le t}, t), \end{aligned} \tag{4.2.1}$$

such that the topology on S is the final topology on S with respect to the map r. Note that the map r is Borel-measurable.

4.2.1. Adapted Random Probability Measures

Instead of restricting ourselves to $\mathbb{G}$-stopping times on $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$, we introduce a generalization of the notion of $\mathbb{G}$-stopping times. As seen in Chapter 3, we assume τ to be a $\mathbb{G}$-stopping time, then it can be naturally identified with a $\mathbb{G}$-adapted stochastic process $\gamma := (\gamma_t)_{t\in I}$ such that

$$\gamma_t(\omega) := \mathbb{1}_{\{\tau(\omega)\}}(t), \quad \omega \in \Omega, t \in I.$$

Thus, for a.e. ω the stochastic process γ defines a probability measure on I, which in turn tells us the probability of having already stopped at time t. This leads us to the following defintion of *adapted random probability measures* which is slightly attenuated from the Definition 3.1.1.

Definition 4.2.2 (Adapted Random Probability Measure).
For a real-valued, stochastic process $\gamma = (\gamma_t)_{t\in I}$ on $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$, we write $\gamma \in \mathcal{M}_I$, if

(a) $\gamma_t \geq 0$ a.s. for all $t \in I$,

(b) $\sum_{t\in I} \gamma_t = 1$, a.s.,

(c) γ is $\mathbb{G}$-adapted.

The space of all these adapted random probability measure is denoted with $\mathcal{M}_I$. Given a probability measure $\nu = (\nu_t)_{t\in I}$ on I, we say that the stochastic process γ above is in $\mathcal{M}_I^\nu$, if in addition

(d) $\mathbb{E}[\gamma_t] = \nu_t$ for all $t \in I$.

As explained at the beginning of this chapter, we want to maximize the expected payoff given a cost function $c\colon S \to \mathbb{R}$ and a stochastic process Z where now the maximization is taken over all adapted random probability measures which continue along a given probability measure.

Problem (OptStop$^\gamma$). Given a Borel-measurable payoff function $c : S \to \mathbb{R}$ and a probability measure ν on I, we seek to find a maximizer of

$$\gamma \mapsto \mathbb{E}\left[\sum_{t\in I} c((Z_s)_{s\le t}, t)\gamma_t\right], \quad \gamma \in \mathcal{M}_I^\nu.$$

Remark 4.2.3. Note that this is an enlargement of the standard optimal stopping problem

$$\tau \mapsto \mathbb{E}[c((Z_s)_{s\le\tau}, \tau)],$$

where τ is a $\mathbb{G}$-stopping time and $\mathbb{P}(\tau = t) = \nu_t$ for all $t \in I$. We denote the space of all $\mathbb{G}$-stopping times with $\mathcal{T}_I$ and its restriction to all stopping times τ such that $\mathcal{L}(\tau) = \nu$ with $\mathcal{T}_I^\nu$. Obviously it holds that

$$\sup_{\tau\in\mathcal{T}_I^\nu} \mathbb{E}[c((Z_s)_{s\le\tau}, \tau)] \le \sup_{\gamma\in\mathcal{M}_I^\nu} \mathbb{E}\left[\sum_{t\in I} c((Z_s)_{s\le t}, t)\gamma_t\right],$$

because $\mathcal{T}_I^\nu$ is embedded in $\mathcal{M}_I^\nu$.

4.2.2. Randomized Stopping Times and Couplings

Rather than working with an abstract filtered probability space, it is possible to work on the path space $(\mathbb{R}^I, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{t\in I}, \mu)$ of the stochastic process Z, where $\mathbb{F}$ is the natural filtration of Z and $\mathcal{F}$ the Borel-σ-algebra on $\mathbb{R}^I$. If $T < \infty$ the space $\mathbb{R}^I$ equipped with the product topology is a complete metric space, where the metric can be chosen as

$$\max_{t\in I} |x_t - y_t|.$$

If we admit $T = \infty$, i.e., an infinite time horizon, the path space with the product topology remains Polish as a countable product of Polish spaces. Furthermore, a possible metric which induces the product topology is

$$\rho(x,y): \mathbb{R}^I \times \mathbb{R}^I \to \mathbb{R},$$
$$(x,y) \mapsto \sum_{t\in I} \frac{1}{2^t} \frac{|x_t - y_t|}{1+|x_t - y_t|}.$$

As stated in the introduction of this chapter, instead using stopping times which stop at one point in time, it is possible to generalize this with so-called randomized stopping times. A randomized stopping time π tries to mimic stopping times by assigning almost every path ω a probability measure π_ω on I which again tells us the probability with which we stop at time t.

Definition 4.2.4 (Randomized Stopping Times)**.**
A probability measure π on $\mathbb{R}^I \times I$ is called randomized stopping time, if

(a) $\mathrm{proj}_{\mathbb{R}^I}(\pi) = \mu$,

(b) the mapping $\omega \mapsto \pi_\omega(t)$ is $\mathcal{F}_t$-measurable for all $t \in I$, where $(\pi_\omega)_{\omega\in\mathbb{R}^I}$ is a disintegration of π. Or equivalently, the with π associated process $A := (A_t)_{t\in I}$, where $A_t(\omega) := \sum_{s\le t} \pi_\omega(s)$ is $\mathcal{F}_t$-measurable.

Again, the space of all randomized stopping times on $\mathbb{R}^I \times I$ which satisfy (a) and (b) are denoted with $\mathrm{RST}(\mu)$. Given a probability measure ν on I, we are interested in random stopping times π such that

(c) $\mathrm{proj}_I(\pi) = \nu$.

The restriction of $\mathrm{RST}(\mu)$ to all probability measures which in addition satisfy (c) is denoted with $\mathrm{RST}(\mu,\nu)$.

Remark 4.2.5. The marginal of the random stopping time π is assumed to be distributed with the law of μ. This can be understood as that the probabilities of the paths are preserved. Since we are working on Polish spaces, the (unique) disintegration $(\pi_\omega)_{\omega\in\mathbb{R}^I}$ exists and assigns μ-almost every path ω a probability measure on I.

In the setting of optimal transport it is more convenient to work with so-called couplings which are product probability measures such that the marginals satisfy a certain law. For our case it is reasonable to consider all couplings on $\mathbb{R}^I \times I$ between μ and ν.

Definition 4.2.6 (Couplings)**.**
A coupling on $\mathbb{R}^I \times I$ with marginals μ and ν is a product probability measure π on $\mathbb{R}^I \times I$ such that

(i) $\mathrm{proj}_{\mathbb{R}^I}(\pi) = \mu$,

(ii) $\mathrm{proj}_I(\pi) = \nu$.

The space of all these product measures on $\mathbb{R}^I \times I$ is denoted with $\mathrm{Cpl}(\mu,\nu)$. The restriction of $\mathrm{Cpl}(\mu,\nu)$ to all couplings satisfying

(iii) $\int \mathbb{1}_{\{t\}}(s)(g - \mathbb{E}[g|\mathcal{F}_t])(\omega)\mathrm{d}\pi(\omega,s) = 0 \quad$ for all $g \in B_b(\mathbb{R}^I),\ t \in I$

is denoted with $\mathrm{Cpl}^{ad}(\mu,\nu)$.

Remark 4.2.7. For a coupling π the property (iii) corresponds to property (b) of randomized stopping times in the Definition 4.2.4. In fact, $\mathrm{Cpl}^{ad}(\mu,\nu)$ coincides with $\mathrm{RST}(\mu,\nu)$. This follows by Lemma 4.2.8 which is an adaptation of [9, Theorem 3.8], where also a proof for the time continuous case can be found.

Lemma 4.2.8. *Let $\pi \in \mathrm{Cpl}(\mu,\nu)$. Then the following are equivalent:*

(a) $\pi \in \mathrm{Cpl}^{ad}(\mu,\nu)$,

(b) Given a disintegration $(\pi_\omega)_{\omega\in\mathbb{R}^I}$ of π, the random variable $\pi_\omega(t)$ is $\mathcal{F}_t$- measurable for all $t \in I$.

Proof. To show the equivalence, we use a different characterization of measurability of integrable random variables, see e.g. in [21]:
An integrable random variable X on $\mathbb{R}^I$ is $\mathcal{F}_t$-measurable if and only if

$$\mathbb{E}[X(Y - \mathbb{E}[Y|\mathcal{F}_t])] = 0 \quad \forall\ Y \text{ integrable and Borel-measurable.}$$

Instead of working with all integrable random variables, we can restrict us to bounded, Borel-measurable random variables. Thus, by a monotone class argument and setting $X := \pi_\omega(t)$, this is equivalent to

$$\mathbb{E}\Big[\mathbb{1}_{\{t\}}(s)\big(g - \mathbb{E}[g|\mathcal{F}_t]\big)\Big] = \int_{\mathbb{R}^I} \pi_\omega(t)\big(g - \mathbb{E}[g|\mathcal{F}_t]\big)(\omega)\mu(\mathrm{d}\omega) = 0 \quad \forall g \in B_b(\mathbb{R}^I).$$ ❑

As already explained, we want to maximize the expected payoff given a cost function $c\colon S \to \mathbb{R}$ and the paths ω of a stochastic process Z where now the maximization is taken over all randomized stopping times which are in $\mathrm{RST}(\mu,\nu)$.

Problem (OptStop$^\pi$)**.** Let $\tilde{c}: \ S \to \mathbb{R}$ be Borel-measurable, then we can define the Borel-measurable function $c := \tilde{c} \circ r$, with r given by (4.2.1), by

$$\begin{aligned} c:\ & \mathbb{R}^I \times I \to \mathbb{R}, \\ & (\omega,t) \mapsto \tilde{c}((\omega)_{s\le t}, t). \end{aligned}$$

We want to find a maximizer of

$$\pi \mapsto \int_{\mathbb{R}^I\times I} c(\omega,t)\mathrm{d}\pi(\omega,t), \quad \pi \in \mathrm{RST}(\mu,\nu).$$

Remark 4.2.9. By Lemma 4.2.8 the problem OptStop$^\pi$ is equivalent to

$$\pi \mapsto \int_{\mathbb{R}^I\times I} c(\omega,t)\mathrm{d}\pi(\omega,t), \quad \pi \in \mathrm{Cpl}^{ad}(\mu,\nu).$$

4.2.3. Connection between the Different Views

The next theorem gives us the connection between the different ways of formalizing our considered problem.

Theorem 4.2.10. *If the filtration $\mathbb{G}$ of the abstract probability space $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ coincides with the natural filtration of the stochastic process Z, then there is a bijection between $\mathcal{M}_I$ and* $\mathrm{RST}(\mu)$. *Given a probability measure ν on I, then there is a bijection from $\mathcal{M}_I^\nu$ into* $\mathrm{RST}(\mu, \nu)$.

Proof. Since $\mathbb{G}$ coincides with the natural filtration of the stochastic process Z, there is a Borel-measurable functions h_γ for every $\gamma \in \mathcal{M}_I$ such that

$$\begin{aligned} &h_\gamma : S \to \mathbb{R}, \\ &\gamma_t(\bar{\omega}) = h_\gamma(Z_s(\bar{\omega})_{s \le t}, t), \quad \mathbb{P}\text{-a.e., } t \in I. \end{aligned}$$

We already know that the mapping $r : \mathbb{R}^I \times I \to S$ is Borel-measurable. Thus,

$$\Phi_\gamma := h_\gamma \circ r : \mathbb{R}^I \times I \to \mathbb{R}$$

is Borel-measurable, $\Phi_\gamma(\cdot, t)$ is $\mathcal{F}_t$-measurable and

$$\gamma_t(\bar{\omega}) = \Phi_\gamma(Z(\bar{\omega}), t), \quad \mathbb{P}\text{-a.e.}$$

Therefore, we deduce that for any $C \in \mathcal{G}_t$

$$\mathbb{E}[\mathbb{1}_C \gamma_t] = \mathbb{E}[\mathbb{1}_C \Phi_\gamma(Z, t)] = \mathbb{E}[\mathbb{1}_{\tilde{C}}(\omega) \Phi_\gamma(\omega, t)],$$

where $\tilde{C}$ is the with C associated set in $\mathcal{F}_t$. We define π such that $\pi(\mathrm{d}\omega, t) := \Phi_\gamma(\omega, t)\mu(\mathrm{d}\omega)$ which indeed defines a probability measure on $\mathbb{R}^I \times I$, and $\pi \in \mathrm{RST}(\mu)$. As a result, the map Ψ

$$\Psi : \mathcal{M}_I \to \mathrm{RST}(\mu), \ \gamma \mapsto \pi,$$

is well-defined and injective.
For any $\pi \in \mathrm{RST}(\mu)$, the map $\omega \mapsto \pi_\omega(t)$ is $\mathcal{F}_t$-measurable. Hence, $\bar{\omega} \mapsto \pi_{Z(\bar{\omega})}$ is $\mathcal{G}_t$-measurable. Therefore, we may define $\tilde{\gamma}_t(\bar{\omega}) = \pi_{Z(\bar{\omega})}(t)$ and conclude

$$\Psi((\tilde{\gamma}_t)_{t \in I}) = \pi(Z(\bar{\omega}), t), \quad \mathbb{P}\text{-a.e, } t \in I,$$

which proves the first part of the assertion. Using Lemma 4.2.8 the second part follows analogously. ❑

Remark 4.2.11. If the probability space $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ coincides with $(\mathbb{R}^I, \mathcal{F}, \mathbb{F}, \mu)$, then the bijection of Theorem 4.2.10 follows due to the relation of disintegration and product measure on Polish spaces.

Corollary 4.2.12. *Under the assumptions of Theorem 4.2.10, the optimization problems* $\textsc{OptStop}^\gamma$ *and* $\textsc{OptStop}^\pi$ *are equivalent.*

Proof. For any $\gamma \in \mathcal{M}_I^\nu$, we define a product measure on $\Omega \times I$ via $\tilde{\gamma}(d\bar{\omega}, t) = \gamma_t(\bar{\omega})\mathbb{P}(d\bar{\omega})$. Following the proof of Theorem 4.2.10 we see that $(Z, id)_\# \tilde{\gamma} = \pi$, where π is the associated product measure on $\mathbb{R}^I \times I$.

$$\int_\Omega \sum_{t\in I} \tilde{c}(Z_s(\bar{\omega})_{s\leq t}, t)\gamma_t(\bar{\omega})\,\mathrm{d}\mathbb{P}(\bar{\omega}) = \int_{\Omega\times I} \tilde{c}(Z_s(\bar{\omega})_{s\leq t}, t)\,\mathrm{d}\tilde{\gamma}(\bar{\omega}, t),$$
$$\int_{\mathbb{R}^I\times I} \tilde{c}(Z_s(\bar{\omega})_{s\leq t}, t)\,\mathrm{d}\pi(\bar{\omega}, t) = \int_{\mathbb{R}^I\times I} c(\omega, t)\,\mathrm{d}\pi(\omega, t).$$

Following the second part of the proof of Theorem 4.2.10, we may define for any $\pi \in \mathrm{Cpl}^{ad}(\mu, \nu)$ an adapted random probability measure $\gamma \in \mathcal{M}_I^\nu$ via

$$\gamma_t(\bar{\omega}) := \pi_{Z(\bar{\omega})}(t),$$

where π_ω is the disintegration of π. Then

$$\int_{\mathbb{R}^I\times I} c(\omega, t)\,\mathrm{d}\pi(\omega, t) = \int_{\mathbb{R}^I} \sum_{t\in I} \tilde{c}((\omega_s)_{s\leq t}, t)\pi_\omega(t)\,\mathrm{d}\mu(\omega) = \int_\Omega \sum_{t\in I} \tilde{c}(Z_s(\bar{\omega})_{s\leq t}, t)\gamma_t(\bar{\omega})\,\mathbb{P}(d\bar{\omega}).$$
❑

We have shown the connection between the two different approaches. From now on, however, we only deal with the second possibility of considering our problem, namely to formulate it as an optimal transport problem.

4.3. Existence of a Maximizer

The main statement given in Theorem 4.3.4 comprises that there exists an optimal randomized stopping time for our problem OptStop$^\pi$. To prove this, we consider first needed help statements and remember the notable *Prokhorov's* Theorem.

Proposition 4.3.1. *For every $b : \mathbb{R}^I \times I \to \mathbb{R}$, bounded and Borel-measurable, the functional*

$$F : \ \mathrm{Cpl}(\mu, \nu) \to \mathbb{R}, \ \pi \mapsto \int b(\omega, t)\mathrm{d}\pi(\omega, t) =: \pi(b),$$

is continuous w.r.t. weak topology on $\mathrm{Cpl}(\mu, \nu)$.

Proof. Given a sequence $\pi_n \in \mathrm{Cpl}(\mu, \nu)$, $n \in \mathbb{N}$, such that $\pi_n \rightharpoonup \pi$ as $n \to \infty$, we know by [13, Lemma 2.3] that for all $A \in \mathcal{B}(\mathbb{R}^I)$ and $t \in I$

$$\pi_n(A \times \{t\}) \to \pi(A \times \{t\}).$$

Since b is bounded and measurable, there exists a sequence of simple functions b_m such that $|b - b_m| < \frac{1}{m}$, μ-a.e. Therefore

$$\forall \varepsilon > 0, \forall n \in \mathbb{N}, \exists m_\varepsilon : |\pi(b - b_m)| < \varepsilon \text{ and } |\pi_n(b - b_m)| < \varepsilon, \quad m \geq m_\varepsilon,$$

and

$$\forall \varepsilon > 0, \forall m \in \mathbb{N}, \exists n_\varepsilon : |\pi_n(b_m) - \pi(b_m)| < \varepsilon, \quad n \geq n_\varepsilon.$$

We conclude

$$|\pi(b)-\pi_n(b)| \le |\pi(b-b_m)| + |\pi(b_m)-\pi_n(b_m)| + |\pi_n(b-b_m)| < 3\varepsilon, \quad n \ge n_\varepsilon,\ m \ge m_\varepsilon.$$

❑

Corollary 4.3.2. *For every $h\colon \mathbb{R}^I \times I \to \mathbb{R} \cup \{-\infty\}$, bounded from above and Borel-measurable, the functional*

$$H : \operatorname{Cpl}(\mu,\nu) \to \mathbb{R},\ \pi \mapsto \int h(\omega,t)\,\mathrm{d}\pi(\omega,t),$$

is upper semicontinuous w.r.t. the weak topology on $\operatorname{Cpl}(\mu,\nu)$.

Proof. We define $h_n := \max(h,-n)$, $n \in \mathbb{N}$, which are bounded, measurable and $h_n \searrow h$ pointwise as $n \to \infty$. By Proposition 4.3.1, we can define a sequence of continuous functionals

$$H_n : \operatorname{Cpl}(\mu,\nu) \to \mathbb{R},\ \pi \mapsto \pi(h_n), \quad \text{where } \inf_n H_n(\pi) = H(\pi).$$

Let $\pi_m \rightharpoonup \pi$ in $\operatorname{Cpl}(\mu,\nu)$ as $m \to \infty$, then

$$H(\pi) = \inf_n H_n(\pi) = \inf_n \limsup_m H_n(\pi_m) \ge \limsup_m \inf_n H_n(\pi_m) = \limsup_m H(\pi_m).$$

❑

For the sake of completeness, we want to state the notable *Prokhorov's* Theorem, see [73, Lemma 4.4].

Theorem 4.3.3 (Prokhorov)**.** *Let $\mathcal{P}(X)$ be the set of Borel probability measures on a topological space X. If X is a Polish space, then a set $P \subseteq \mathcal{P}(X)$ is precompact for the weak topology if and only if it is tight, i.e., for any $\varepsilon > 0$ there is a compact set K_ε such that $\pi(X \backslash K_\varepsilon) \le \varepsilon$ for all $\pi \in P$.*

To show the existence, we will proceed as follows:

(I) Show that the set, over which the supremum is taken, is compact.

(II) Show that the functional is upper semicontinuous.

Note that $\operatorname{Cpl}^{ad}(\mu,\nu)$ is non-empty, since the product measure $\mu \otimes \nu \in \operatorname{Cpl}^{ad}(\mu,\nu)$ and it holds for all $g \in B_b(\mathbb{R}^I)$ that

$$\begin{aligned}
&\int_{\mathbb{R}^I \times I} \mathbb{1}_{\{t\}}(s)(g - \mathbb{E}[g|\mathcal{F}_t])(\omega)\,\mathrm{d}\mu\otimes\nu(\omega,s) \\
&= \int_I \mathbb{1}_{\{t\}}(s)\,\mathrm{d}\nu(s) \int_{\mathbb{R}^I} (g - \mathbb{E}[g|\mathcal{F}_t])(\omega)\,\mathrm{d}\mu(\omega) = \nu(\{t\})\mathbb{E}_\mu[g - \mathbb{E}[g|\mathcal{F}_t]] = 0.
\end{aligned}$$

Thus, a maximizing sequence exists and the compactness provides a maximizer.

Theorem 4.3.4 (Existence of a Maximizer)**.** *Let $c : S \to \mathbb{R} \cup \{-\infty\}$ be bounded from above and Borel-measurable, then there exists a solution to* OptStop$^\pi$.

Proof. As a direct consequence of Prokhorov's Theorem, see Theorem 4.3.3, we get that $\mathrm{Cpl}(\mu,\nu)$ is relatively compact. By [13, Lemma 2.3], we obtain in addition that $\mathrm{Cpl}(\mu,\nu)$ is closed, hence compact in the weak topology. To see the compactness of $\mathrm{Cpl}^{ad}(\mu,\nu)$, we consider

$$b(\omega,s) := \mathbb{1}_{\{t\}}(s)(g - \mathbb{E}[g|\mathcal{F}_t])(\omega)$$

for a bounded and Borel-measurable function g on $\mathbb{R}^I$ and $t \in I$. Let $\pi_n \rightharpoonup \pi$ in $\mathrm{Cpl}(\mu,\nu)$ as $n \to \infty$ and $\pi_n \in \mathrm{Cpl}^{ad}(\mu,\nu),\ n \in \mathbb{N}$, by applying Proposition 4.3.1, we obtain

$$0 = \pi_n(b) \to \pi(b).$$

Since s and g were arbitrary, we conclude $\pi \in \mathrm{Cpl}^{ad}(\mu,\nu)$ signifying the compactness of $\mathrm{Cpl}^{ad}(\mu,\nu)$.

Now, choose a sequence $\pi_n \in \mathrm{Cpl}^{ad}(\mu,\nu)$ such that

$$\lim_n \pi_n(c) = \sup_{\tilde{\pi}\in\mathrm{Cpl}^{ad}(\mu,\nu)} \int_{\mathbb{R}^I\times I} c(\omega,s)\, d\tilde{\pi}(\omega,s) =: C.$$

Due to the compactness of $\mathrm{Cpl}^{ad}(\mu,\nu)$, we can extract a convergent subsequence $\pi_{n_k} \rightharpoonup \pi \in \mathrm{Cpl}^{ad}(\mu,\nu)$ such that π possesses the desired property by Corollary 4.3.2

$$C = \limsup_{n_k} \pi_{n_k}(c) \le \pi(c) \le C.$$

❏

4.4. Duality

In the theory of optimal transport is the Kantorovich duality a basic concept. The classical Monge-Kantorovich problem deals with the topic of minimizing the expected loss given a cost function $c:\ X\times Y \to \mathbb{R}\cup\{\infty\}$ when iterating over all couplings $\pi \in \mathrm{Cpl}(\mu,\nu)$, where μ and ν are probability measures on X and Y, respectively. Formally, we seek to find a coupling of (μ,ν) which minimizes the total cost function

$$\mathrm{Cpl}(\mu,\nu) \ni \pi \mapsto \int c\, \mathrm{d}\pi$$

among all possible couplings. The function c can be interpreted as the cost of moving mass from X, which is distributed according to μ, to Y which shall be distributed according to ν. The couplings $\pi \in \mathrm{Cpl}(\mu,\nu)$ are called transport plans; and a coupling which minimizes the expected loss is called an optimal transport plan.

Our considered problem OptStop^π is a maximization problem. However, we can switch from a minimization problem to a maximization problem, simply by multiplying the payoff function c with -1, and hence call it cost function, because it holds that

$$\inf_{\pi\in\mathrm{Cpl}(\mu,\nu)} \pi(c) = -\sup_{\pi\in\mathrm{Cpl}(\mu,\nu)} \pi(-c).$$

Therefore we want to state the usual Kantorovich duality Theorem, see [73, Theorem 5.10].

Theorem 4.4.1 (Kantorovich duality). *Let (X,μ) and (Y,ν) be two Polish probability spaces and let $c: X \times Y \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous cost function, such that*

$$\forall (x,y) \in X \times Y, \quad c(x,y) \geq a(x) + b(y)$$

for some real-valued upper semicontinuous functions $a \in L^1(\mu)$ and $b \in L^1(\nu)$. Then there is duality

$$\inf_{\pi \in \mathrm{Cpl}(\mu,\nu)} \pi(c) = \sup_{\substack{(f,g) \in C_b(X) \times C_b(Y) \\ f_1 + f_2 \leq c}} \left(\int_X f_1 \,\mathrm{d}\mu + \int_Y f_2 \,\mathrm{d}\nu \right).$$

We are interested into maximizing $\textsc{OptStop}^{\pi}$ which is a maximization problem over the space $\mathrm{Cpl}^{ad}(\mu,\nu)$. In fact, this space is a restriction of $\mathrm{Cpl}(\mu,\nu)$. We may define a space W as the linear span of

$$\left\{ w(\omega,s) := \mathbb{1}_{\{t\}}(s)(g - \mathbb{E}[g|\mathcal{F}_t])(\omega) \mid g \in C(\mathbb{R}^I) \cap L^1(\mu),\ t \in I \right\}. \tag{4.4.2}$$

Evidently, $\mathrm{Cpl}^{ad}(\mu,\nu)$ coincides with the restriction of $\mathrm{Cpl}(\mu,\nu)$ to all couplings π satisfying the additional linear constraints

$$\pi(w) = 0 \quad \text{for all } w \in W. \tag{4.4.3}$$

Thus the Kantorovich duality Theorem has to be extended to the case where linear constraints are posed to $\mathrm{Cpl}(\mu,\nu)$. A generalized version of this problem was shown by Zaev in [76]. Following the proof of [76, Theorem 2.1] and using Theorem 4.4.1 yields the following theorem:

Theorem 4.4.4 (Duality). *If the cost function c satisfies the assumptions of Theorem 4.4.1 with $X := \mathbb{R}^I$ and $Y := I$. Then there is duality*

$$\inf_{\pi \in \mathrm{Cpl}^{ad}(\mu,\nu)} \pi(c) = \sup_{\substack{(f_1,f_2,w) \in C_b(\mathbb{R}^I) \times C_b(I) \times W \\ f_1 + f_2 + w \leq c}} \left(\int_{\mathbb{R}^I} f_1 \,\mathrm{d}\mu + \sum_{t \in I} f_2(t)\nu(t) \right).$$

Proof. The inequality

$$\inf_{\pi \in \mathrm{Cpl}^{ad}(\mu,\nu)} \pi(c) \geq \sup_{f_1 + f_2 + w \leq c} \big(\mu(f_1) + \nu(f_2) \big)$$

follows immediately from

$$\inf_{\pi \in \mathrm{Cpl}^{ad}(\mu,\nu)} \pi(c) \geq \inf_{\pi \in \mathrm{Cpl}^{ad}(\mu,\nu)} \sup_{f_1 + f_2 + w \leq c} \big(\mu(f_1) + \nu(f_2) \big) = \sup_{f_1 + f_2 + w \leq c} \big(\mu(f_1) + \nu(f_2) \big).$$

For the reverse inequality we consider

$$\sup_{f_1 + f_2 + w \leq c} \big(\mu(f_1) + \nu(f_2) \big) = \sup_{w \in W} \sup_{f_1 + f_2 \leq c - w} \big(\mu(f_1) + \nu(f_2) \big).$$

Note that $W \subseteq C_b(\mathbb{R}^I \times I)$, thus, $c - w$ is again lower semicontinuous on $\mathbb{R}^I \times I$. We may choose $a_w(x) := a(x) - \|w\|_\infty$ and $b_w(x) := b(x) - \|w\|_\infty$ which in turn satisfy the assumption of Theorem 4.4.1. Thereby we obtain that

$$\sup_{f_1+f_2+w\leq c} \big(\mu(f_1) + \nu(f_2)\big) = \sup_{w\in W} \inf_{\pi\in \mathrm{Cpl}^{ad}(\mu,\nu)} \pi(c-w).$$

For any $\pi \notin \mathrm{Cpl}^{ad}(\mu,\nu)$ there exists a $w \in W$ such that $\pi(w) < 0$, and

$$\sup_{\alpha>0} \pi(c - \alpha w) = +\infty.$$

Since $\mathrm{Cpl}^{ad}(\mu,\nu)$ is not empty, we conclude that

$$\inf_{\pi\in \mathrm{Cpl}^{ad}(\mu,\nu)} \pi(c) = \sup_{\substack{(f_1,f_2,w)\in C_b(\mathbb{R}^I)\times C_b(I)\times W \\ f_1+f_2+w\leq c}} \big(\mu(f_1) + \nu(f_2)\big).$$

❑

4.5. Examples

In this section we investigate examples and determine optimal maximizers. At first we consider the two time steps. In this case we can reformulate appropriately the problem and then construct the optimal strategy. This can be used to generalize it on a finite time domain and to determine the optimal strategy for special cost function.

4.5.1. Example: Two Time Steps

First, we want to consider a stochastic process $(Z_t)_{t\in I}$ with independent increments on the time domain $I := \{1,2\}$ and maximize the following functional on $\mathrm{RST}(\mu,\nu)$

$$\pi \mapsto \int_{\mathbb{R}^2\times\{1,2\}} c(\omega,t)\, d\pi(\omega,t), \tag{4.5.1}$$

as in $\textsc{OptStop}^\pi$. Remember that the set $\mathrm{RST}(\mu,\nu)$ coincides with the set $\mathrm{Cpl}^{ad}(\mu,\nu)$. To facilitate notations, we will denote by $\pi(\cdot,t)$ for $t \in I$ and $\pi \in \mathrm{RST}(\mu,\nu)$ the (sub-)probability measure on $\mathbb{R}^I = X_1 \times X_2$ induced by

$$B \mapsto \pi(B,t) \quad B \in \mathcal{B}(\mathbb{R}^I).$$

Given a randomized stopping time π, we may consider the measure m on $X_1 = \mathbb{R}$ such that

$$m(A) := \pi(A \times \mathbb{R} \times \{1\}) = \mathrm{proj}_{X_1}(\pi(\cdot,1))(A) \quad \forall A \in \mathcal{B}(\mathbb{R}).$$

Remember that $\omega \mapsto \pi_\omega(t)$ is $\mathcal{F}_t$-adapted, in particular for $t = 1$, we find a Z_1-measurable function $h : X_1 \times X_2 \to \mathbb{R}$ such that

$$\pi_\omega(1) = h(\omega), \quad \mu\text{-a.e., } \omega \in \mathbb{R}^I, \tag{4.5.2}$$

and for $h = \tilde{h} \circ r$ analogously to (4.2.1)

$$h(\omega) = h(\tilde{\omega}) =: \tilde{h}(\omega_1) \quad \omega,\ \tilde{\omega} \in \mathbb{R}^I,\ \omega_1 = \tilde{\omega}_1.$$

Therefore, we deduce that

$$m(A) = \int \mathbb{1}_{A\times\mathbb{R}}(\omega)\,\pi_\omega(1)\,\mathrm{d}\mu(\omega) = \int \mathbb{1}_A(\omega_1)h(\omega_1)\,\mathrm{d}\,\mathrm{proj}_{X_1}(\mu)(\omega),$$

which implies $m \leq \mathrm{proj}_{X_1}(\mu)$ and let us define a measure n on $\mathbb{R}$ satisfying

$$n := \mathrm{proj}_{X_1}(\mu) - m.$$

Informally, the measures m and n describe with which probability a path is stopped at time 1 or continues to time 2, respectively. Remember that $Z_2 - Z_1 \sim p_2$. Using the measures m and n in the maximization problem (4.5.1) yields to

$$\begin{aligned}
\int_{\mathbb{R}^I\times I} c(\omega,t)\,\mathrm{d}\pi(\omega,t) &= \int_{\mathbb{R}^I} c(\omega,1)\,\mathrm{d}\pi(\omega,1) + \int_{\mathbb{R}^I} c(\omega,2)\,\mathrm{d}\pi(\omega,2)\\
&= \int_{X_1} \tilde{c}((\omega_1),1)\,\mathrm{d}m(\omega_1) + \int_{\mathbb{R}^I} c(\omega,2)(1-h(\omega))\,\mu(\mathrm{d}\omega)\\
&= \int_{X_1} \tilde{c}((\omega_1),1)\,\mathrm{d}m(\omega_1) + \int_{X_1} (1-\tilde{h}(\omega_1)) \int_{\mathbb{R}} c((\omega_1,\omega_1+z),2)\,\mathrm{d}p_2(z)\,\mathrm{d}\,\mathrm{proj}_{X_1}(\mu)(\omega_1)\\
&= \int_{X_1} \tilde{c}((\omega_1),1)\,\mathrm{d}m(\omega_1) + \int_{X_1}\int_{\mathbb{R}} c((\omega_1,\omega_1+z),2)\,\mathrm{d}p_2(z)\,\mathrm{d}n(\omega_1). \qquad (4.5.3)
\end{aligned}$$

This equality can be interpreted in the way that the first integral in (4.5.3) describes the expected payoff if the process is stopped at time 1, whereas the second integral describes the expected payoff at time 2. Note that

$$\int_{\mathbb{R}} c((x,x+z),2)\,\mathrm{d}p_2(z)$$

is the expected payoff when we stop at time 2, conditioned on $\omega_1 = x$.
For any path $\omega := (\omega_1,\omega_2) \in \mathbb{R}^I$ and the corresponding probability measure π_ω on I, we see that

$$(c(\omega,1)\pi_\omega(1) + c(\omega,2)\pi_\omega(t)) - c(\omega,1) = (c(\omega,2)-c(\omega,1))\pi_\omega(2).$$

Without loss of generality we may assume that

$$c(\omega,1) = 0. \qquad (4.5.4)$$

If $c(\omega,1) \neq 0$, than we might define $\bar{c}(\omega,t) := c(\omega,t) - c(\omega,1)$. Clearly, it holds that $\bar{c}(\omega,1) = 0$. Because subtracting a constant from the functional solving the problem given in (4.5.1) does not change the property of being a maximizer of it, we will instead maximize

$$\pi \mapsto \int_{\mathbb{R}^I\times I} \bar{c}(\omega,t)\,\mathrm{d}\pi(\omega,t) = \int_{\mathbb{R}^I\times I} c(\omega,t)\,\mathrm{d}\pi(\omega,t) - \int_{\mathbb{R}^I} c(\omega,1)\,\mathrm{d}\mu(\omega),$$

where the last equality holds due to $\mathrm{proj}_{\mathbb{R}^I}(\pi) = \mu$. Note that both maximization problems are equivalent, but $\bar{c}(\omega,1) = 0$.

Using the reformulation (4.5.3) and assumption (4.5.4) our problem given in (4.5.1) is reduced to the following maximization problem:
Among all measures n on $X_1 = \mathbb{R}$ satisfying

$$n \leq \mathrm{proj}_{X_1}(\mu) \quad \text{and} \quad n(\mathbb{R}) = n(X_1) = \nu(2)$$

find the maximizer of

$$n \mapsto \int_{X_1} \int_{\mathbb{R}} c((x, x+z), 2)\, \mathrm{d}p_2(z)\, \mathrm{d}n(x).$$

Define the function $k : X_1 \to \mathbb{R}$ by

$$k(x) := \int_{\mathbb{R}} c((x, x+z), 2)\, \mathrm{d}p_2(z), \tag{4.5.5}$$

where $Z_2 - Z_1 \sim p_2$. Then we get that $\int_{X_1} \int_{\mathbb{R}} c((x, x+z), 2)\, \mathrm{d}p_2(z) \mathrm{d}n(x) = n(k)$.
The considerations above lead to the following theorem:

Theorem 4.5.6. *Let $I = \{1, 2\}$, $\mathbb{R}^I = X_1 \times X_2$. Given an optimal $\pi^* \in \mathrm{RST}(\mu, \nu)$ such that*

$$\pi^*(c) = \sup_{\pi \in \mathrm{RST}(\mu,\nu)} \pi(c). \tag{4.5.7}$$

Then,

$$n_{\pi^*} := \mathrm{proj}_{X_1}(\mu - \pi^*(\cdot, 1))$$

is a measure on $X_1 = \mathbb{R}$ satisfying

$$n_{\pi^*} \leq \mathrm{proj}_{X_1}(\mu), \quad n_{\pi^*}(\mathbb{R}) = n_{\pi^*}(X_1) = \nu(2), \tag{4.5.8}$$

and maximizes under all measures satisfying (4.5.8)

$$n_{\pi^*}(k) = \sup_{n} n(k). \tag{4.5.9}$$

Vice versa, let n^* be a measure on $\mathbb{R}(\mu, \nu)$ satisfying (4.5.8) *and maximizing* (4.5.9).
Then,

$$\pi_{n^*}((\mathrm{d}x, \mathrm{d}y), t) = \begin{cases} (\mathrm{proj}_{X_1}(\mu) - n^*)(\mathrm{d}x) p_2(\mathrm{d}y) & \text{for } t = 1, \\ n^*(\mathrm{d}x) p_2(\mathrm{d}y) & \text{for } t = 2, \end{cases}$$

defines a RST which maximizes (4.5.7).

Remark 4.5.10. Note the overlaps of the definitions with regard to the arguments c and (ω, t) or ω_1.

Proof. First, we take a closer look at n_{π^*} and note that it satisfies (4.5.8) and, due to Lemma 4.2.8 and (4.5.2),

$$n_{\pi^*}(\mathrm{d}x) = \mathrm{proj}_{X_1}(\mu - \pi^*(\cdot, 1))(\mathrm{d}x) = (1 - \tilde{h}(x))\, \mathrm{proj}_{X_1}(\mu)(\mathrm{d}x),$$

implying

$$\pi^*(c) = \int_{X_1\times\mathbb{R}} c((\omega_1,\omega_1+z),2)\pi^*_{(\omega_1,\omega_1+z)}(2)\,\mathrm{d}\mu((\omega_1,\omega_1+z)) = \int_{\mathbb{R}} k(\omega_1)\,\mathrm{d}n_{\pi^*}(\omega_1). \quad (4.5.11)$$

Furthermore, π_{n^*} defines a measure on $\mathbb{R}^I \times I$ such that $\mathrm{proj}_I(\pi_{n^*}) = \nu$ and

$$\pi_{n^*}(c) = \int_{\mathbb{R}} k(\omega_1)\,\mathrm{d}n^*(\omega_1). \quad (4.5.12)$$

For any bounded, Borel-measurable function g on $\mathbb{R}^I$, the function

$$f(x) := \int_{\mathbb{R}} g(x,x+z)\,\mathrm{d}p_2(z)$$

is Borel-measurable and

$$f(Z_1) = \mathbb{E}[f(Z_1)|\mathcal{F}_1] = \int_{\mathbb{R}} \mathbb{E}[g(Z_1,Z_1+z)|\mathcal{F}_1]\,\mathrm{d}p_2(z) \quad \text{a.e.},$$

which particularly yields for $t\in I$

$$\int_{X_1}\int_{\mathbb{R}} (g-\mathbb{E}[g|\mathcal{F}_t])(\omega_1,\omega_1+z)\,\mathrm{d}p_2(z)\,\mathrm{d}\,\mathrm{proj}_{X_1}(\mu)(\omega_1) = 0,$$

implying that $\pi_{n^*}\in \mathrm{RST}(\mu,\nu)$. Due to (4.5.11) and (4.5.12) the optimality of n_{π^*} and π_{n^*}, respectively, follow, and thus the assertion. ❑

Remark 4.5.13. By Theorem 4.5.6, it becomes obvious that for two time steps, it is sufficient to consider a quantile q of k given in (4.5.5) such that q is maximal in $\mathbb{R}$ with the property $\nu(2) \le \mathrm{proj}_{X_1}(U_q)$, where $U_q := \{x\in X_1 \ : \ k(x)\ge q\}$. Then we define

$$n(\mathrm{d}x) := \begin{cases} \mathrm{proj}_{X_1}(\mu)(\mathrm{d}x) & k(x) > q, \\ \alpha\cdot \mathrm{proj}_{X_1}(\mu)(\mathrm{d}x) & k(x) = q, \\ 0 & \text{else}, \end{cases} \quad (4.5.14)$$

where $\alpha\in[0,1]$ is chosen such that $n(\mathbb{R}) = \nu(2)$. It is apparent that n maximizes (4.5.9), since for any other measure $\tilde{n}$ satisfying (4.5.8), it holds

$$n(k) - \tilde{n}(k) = \int_{\mathbb{R}} k(x)(n-\tilde{n})(\mathrm{d}x) \ge \int_{\mathbb{R}} q - k(x)(\tilde{n}-n)^+(\mathrm{d}x) \ge 0.$$

The second part of Theorem 4.5.6 yields the maximizer π^*_n. As already mentioned, this method can be interpreted as not-stopping those paths, which have a higher expected payoff in the next turn. Instead of interpreting it this way, we can view it as stopping all paths with a lower expected payoff – indeed, this would result in the same quantiles.
For a special class of cost functions, we will be able to extend the idea of using quantiles of the corresponding k for each time step inductively, to construct an optimizer for an arbitrary amount of time steps.

4.5.2. Example: Symmetric Random Walk

For $I := \{1, \dots, T\}$, $T \in \mathbb{N}$, we consider a symmetric random walk $(Z_t)_{t \in I}$ on $\mathbb{Z}$ starting at 0, where the increments $Z_s - Z_{s-1}$ are independent and uniformly distributed on $\{-1, 1\}$. Let the payoff function c be

$$c : \mathbb{R}^I \times I \to \mathbb{R} : (\omega, t) \mapsto t \cdot \omega_t,$$

which is indeed $\mathcal{F}_t$-adapted and bounded, if the time horizon is finite.

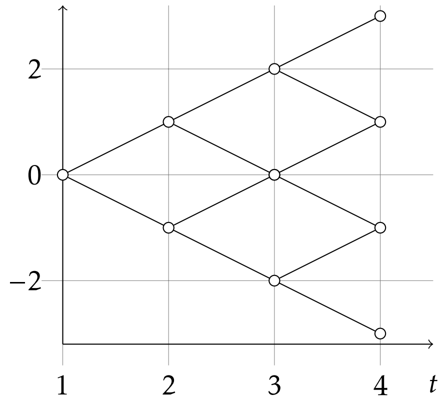


Figure 4.1.: Paths of the process $(Z_t)_{t \in I}$ starting at 0 for $I := \{1, 2, 3, 4\}$.

We want to find an optimal $\pi \in \mathrm{RST}$. Note that for $\omega, \eta \in \mathbb{R}^I$ and $t \in I$

$$\begin{aligned} \int \tilde{c}\Big((\omega_1, \dots, \omega_{t-1}, \omega_{t-1} + z), t\Big) \mathrm{d}p_t(z) &\geq \int \tilde{c}\Big((\eta_1, \dots, \eta_{t-1}, \eta_{t-1} + z), t\Big) \mathrm{d}p_t(z) \\ \iff \int t \cdot (\omega_{t-1} + z) \mathrm{d}p_t(z) &\geq \int t \cdot (\eta_{t-1} + z) \mathrm{d}p_t(z) \\ \iff \omega_{t-1} &\geq \eta_{t-1}. \end{aligned} \tag{4.5.15}$$

By using Theorem 4.5.6 we are able to solve the problem restricted to two time steps. By skillful projection of the path space, we may consider a case with exactly two time steps. For the marginal of π to satisfy the constraint ν, we have to define $\pi(\omega, 1) = \nu(1)$. For every step of our recursion $i \geq 2$ we consider the two dimensional space $X_{t_i} \times X_{t_{i+1}}$. Note that for $I := \{1, \dots, T\}$, $T \in \mathbb{N}$, we consider $X_t \times X_{t+1}$, and the recursion step i corresponds to the time step t. Therefore we define measures $\mu_i := \mathrm{proj}_{X_t \times X_{t+1}}(\mu - \sum_{s<t} \pi(\cdot, s))$ and ν_i such that

$$\nu_i(1) = \nu(t), \quad \nu_i(2) = \sum_{s \in I_{>t}} \nu(s).$$

Then we have to solve the problem of finding the maximizer of

$$n \mapsto \int_{X_t} \int_{\mathbb{R}} c((x, x+z), t+1) \, \mathrm{d}p_{t+1}(z) \, \mathrm{d}n(x),$$

among all measures n_i on $X_t = \mathbb{R}$ satisfying

$$n_i \leq \mathrm{proj}_{X_t}(\mu_i) \quad \text{and} \quad n_i(\mathbb{R}) = \nu_i(2),$$

which can be done according to Remark 4.5.13, particularly (4.5.14). Thus, by applying Theorem 4.5.6 recursively, we obtain a randomized stopping time $\pi^i \in \mathrm{RST}(\mu_i, \nu_i)$, $t_i \in I$, which can be naturally merged into $\pi \in \mathrm{RST}(\mu, \nu)$ via

$$\pi(\omega, t) := (\mu(\omega) - \sum_{s<t} \pi(\omega, s)) \cdot \pi^t_{(\omega_t, \omega_{t+1})}(t).$$

Optimality can be shown in the following way: Starting with another arbitrary $\xi \in \mathrm{RST}(\mu, \nu)$, we assume that there exists a minimal t such that $\xi(\cdot, t) \neq \pi(\cdot, t)$. Therefore, there exist $\zeta, \eta \in \mathbb{R}^I$ such that

$$\alpha_1 := \xi_\eta(t) - \pi_\eta(t) > 0, \quad \alpha_2 := \pi_\zeta(t) - \xi_\zeta(t) > 0$$

and let $\alpha := \min(\alpha_1, \alpha_2)$. The overall mass of all paths $\omega \in \mathbb{R}^I$, such that the initial segments $\omega_{\restriction[0,t]}$ coincide with that of η or ζ, is 2^{-t}. Hence, it is possible to swap $2^{-t}\alpha$ mass from $\{\omega \in \mathbb{R}^I : \omega_{\restriction[0,t]} = \eta_{\restriction[0,t]}\}$ to $\{\omega \in \mathbb{R}^I : \omega_{\restriction[0,t]} = \zeta_{\restriction[0,t]}\}$. A new measure is gradually defined by

$$\tilde{\xi}(\omega, s) = \xi(\omega, s), \quad \omega \in \mathbb{R}^I, s < t$$

and

$$\tilde{\xi}_\omega(t) := \begin{cases} \xi_\omega(t) - \alpha & \forall \omega \in \mathbb{R}^I \text{ such that } \omega_{\restriction[0,t]} = \eta_{\restriction[0,t]}, \\ \xi_\omega(t) + \alpha & \forall \omega \in \mathbb{R}^I \text{ such that } \omega_{\restriction[0,t]} = \zeta_{\restriction[0,t]}, \\ \xi_\omega(t) & \text{otherwise}. \end{cases}$$

So far, $\nu(s) = \mathrm{proj}_I(\tilde{\xi})$ still holds as long as $s \leq t$. To fully preserve the marginals, i.e., $\tilde{\xi} \in \mathrm{RST}(\mu, \nu)$, mass is carefully added to the remaining paths $\omega \in \mathbb{R}^I$, where $\omega_{\restriction[0,t]} \in \{\eta_{\restriction[0,t]}, \zeta_{\restriction[0,t]}\}$. Let $\omega \in \mathbb{R}^I$ such that $\omega_{\restriction[0,t]} = \eta_{\restriction[0,t]}$, then there exists $\theta \in \mathbb{R}^I$ such that $\theta_{\restriction[0,t]} = \zeta_{\restriction[0,t]}$ and $(\theta - \omega)(r) = 0$ for all $r > t$. We may set

$$\tilde{\xi}_\omega(s) := \xi_\omega(s) + \frac{\alpha}{1 - \sum_{r \leq t} \xi_\theta(r)} \cdot \xi_\theta(s), \quad \tilde{\xi}_\theta(s) := \xi_\theta(s) - \frac{\alpha}{1 - \sum_{r \leq t} \xi_\theta(r)} \cdot \xi_\theta(s),$$

which yields the correct marginals for $\tilde{\xi}$. Clearly, $\tilde{\xi} \in \mathrm{RST}(\mu, \nu)$ and due to Theorem 4.5.6 together with equation (4.5.15) it holds that $\eta_t \geq \zeta_t$, yielding

$$\tilde{\xi}(c) - \xi(c) = \int_{\{\omega \in \mathbb{R}^I : \omega_{\restriction[0,t]} = \eta_{\restriction[0,t]}\} \times I_{>t}} (\eta_t - \zeta_t)(s - t)(\tilde{\xi} - \xi)^+(\mathrm{d}\omega, s) \geq 0. \tag{4.5.16}$$

Hence, continuing this construction recursively leads in a finite amount of steps to

$$\tilde{\xi}(\omega, s) = \pi(\omega, s), \quad s \leq t, \ \omega \in \mathbb{R}^I.$$

By equation (4.5.16), $\pi(c)$ is an upper bound for the payoff of any constructed $\tilde{\xi}$, and as a consequence an upper bound for $\xi(c)$.

Remark 4.5.17. The method shown in Example 4.5.2 can be used to show optimality for the introduced "greedy" strategy π for a larger class of optimization problems. In the setting of Example 4.5.2, we cannot expect uniqueness of the optimizer π, since for any $\eta, \zeta \in \mathbb{R}^I$, $\eta_t = \zeta_t$ and $t \in I$ such that

$$\pi_\eta(t) > 0, \quad \pi_\zeta(t) < 1,$$

it is possible to swap some mass analogously as described above, creating a new randomized stopping time, but preserving marginals and payoff.

4.5.3. Example: Generalized Setting

Let $(Z_t)_{t\in I}$ be the stochastic process with independent increments and c be the payoff function of the form

$$c(\omega,t) = f(t)\cdot\omega_t,$$

where $f : I \to \mathbb{R}_+$ is monotone increasing. Motivated by Theorem 4.5.6 and Example 4.5.2, we want to show that a greedy algorithm is optimal here, cf. Theorem 4.5.27. Analogous to (4.5.15), we obtain for the expected payoff

$$\begin{aligned} &\int \tilde{c}((\omega_1,\dots,\omega_t,\omega_t+z),t+1)\mathrm{d}p_{t+1}(z) \geq \int \tilde{c}((\eta_1,\dots,\eta_t,\eta_t+z),t+1)\mathrm{d}p_{t+1}(z) \\ \iff{}& f(t+1)\omega_t + \int f(t+1)z\,\mathrm{d}p_{t+1}(z) \geq f(t+1)\eta_t + \int f(t+1)z\,\mathrm{d}p_{t+1}(z) \qquad (4.5.18) \\ \iff{}& \omega_t \geq \eta_t. \end{aligned}$$

We construct a randomized stopping time π by defining a quantile q_t for any $t\in I$ such that

$$q_t := \inf\Big\{q\in\mathbb{R} : \ \mathrm{proj}_{X_t}\Big(\mu - \sum_{s<t}\pi(\cdot,s)\Big)((-\infty,q]) \geq \nu_t\Big\},$$

where $\pi(\omega,t)$ can be defined as

$$\pi(\omega,t) := \mathrm{proj}_{X_t}\Big(\mu - \sum_{s<t}\pi(\cdot,s)\Big)_{\restriction(-\infty,q_t]}(\omega_t), \qquad (4.5.19)$$

if the quantile q_t is exact, otherwise it can be defined similar to Remark 4.5.13.
As in Example 4.5.2, we will show optimality of this strategy by transforming any randomized stopping time iteratively into the proposed one, without lowering the payoff, cf. Theorem 4.5.27. But, before we can show optimality we need some preparations to conduct the swapping of mass.

Lemma 4.5.20. *Let m, n be finite measures on $[0,1]$ such that $m([0,1]) = n([0,1])$. Then there exists a Borel-measurable map $U = (U_1,U_2)$ from $[0,1]\times[0,1]$ onto $[0,1]\times[0,1]$ such that*

$$m([0,x)) + u\cdot m(\{x\}) = n([0,U_1(x,u))) + U_2(x,u)\cdot n(\{U_1(x,u)\}), \quad x,u\in[0,1].$$

In addition, the map U_1 is surjective onto the support of n denoted by $\mathrm{supp}(n)$.

Proof. We can easily extend the measures m and n to measures M and N on $[0,1]\times[0,1]$ by defining them via

$$M(\mathrm{d}x,\mathrm{d}u) = m(\mathrm{d}x), \quad N(\mathrm{d}y,\mathrm{d}v) = n(\mathrm{d}y).$$

For a given pair $(x,u)\in[0,1]\times[0,1]$ we may define the first component of U as

$$U_1(x,u) := \inf\Big\{y\in[0,1] : \ M([0,x)\times[0,1]) + M(\{x\}\times[0,u]) \leq n([0,y])\Big\} \in \mathrm{supp}(n). \quad (4.5.21)$$

The corresponding second component of U can be defined as follows

$$U_2(x,u) := \inf\Big\{v \in [0,1] : \ M([0,x)\times[0,1]) + M(\{x\}\times[0,u]) \\ = N([0,U_1(x,u))\times[0,1]) + N(\{U_1(x,u)\}\times[0,v])\Big\}. \tag{4.5.22}$$

By construction, $U(x,u) = (U_1(x,u), U_2(x,u))$ is well-defined and Borel-measurable. Further,

$$\begin{aligned} m([0,x)) + u\cdot m(\{x\}) &= M([0,x)\times[0,1]) + M(\{x\}\times[0,u]) \\ &= N([0,U_1(x,u))\times[0,1]) + N(\{U_1(x)\}\times[0,U_2(x,u)]) \\ &= n([0,U_1(x,u))) + U_2(x,u)\cdot n(\{U_1(x,u)\}), \end{aligned}$$

where the second equality follows from (4.5.22).
Assume that there exists $(x,u) \in [0,1]\times[0,1]$ with $U_1(x,u) =: y \notin \mathrm{supp}(n)$, then there exists a $\delta > 0$ such that $n([y-\delta, y+\delta]) = 0$, and hence

$$M([0,x)\times[0,1]) + M(\{x\}\times[0,u]) \le n([0,y-\delta]),$$

which contradicts the definition of $U_1(x,u)$, see (4.5.21). Furthermore, for any $y \in \mathrm{supp}(n)$ there exists $(x,u) \in [0,1]\times[0,1]$ with

$$m([0,x)) \le n([0,y]) \le m([0,x]),$$

$$M([0,x)\times[0,1]) + M(\{x\}\times[0,u]) = n([0,y]).$$

For $z \in \mathrm{supp}(n)$, $z < y$ implies $n([0,z]) < n([0,y])$, which yields $U_1(x,u) = y$ and surjectivity. ❑

Lemma 4.5.23. *Under the assumptions of Lemma 4.5.20, we may define a map* $V_y : [0,1] \to [0,1]$ *for a fixed* $y \in \mathrm{supp}(n)$ *by*

$$V_y(x) := \begin{cases} 1 & x \in \mathrm{int}(\gamma_y),\ n(\{y\}) \neq 0, \\ \sup_{(x,u)\in U_1^{-1}(\{y\})} u - \inf_{(x,v)\in U_1^{-1}(\{y\})} v & x \in \partial\gamma_y,\ n(\{y\}) \neq 0, \\ 0 & \textit{else}, \end{cases}$$

where $\gamma_y := \{x : \exists u \in [0,1] \text{ s.t. } U_1(x,u) = y\}$, $\mathrm{int}(\gamma_y)$ *denotes the interior of the set* γ_y, $\partial\gamma_y$ *the boundary of this set and* $U_1^{-1}(\cdot)$ *is the preimage of* U_1. *The set* γ_y *is a closed interval and the map* V_y *is well-defined. Furthermore, the maps* $y \mapsto \inf_{(x,u)\in U_1^{-1}(\{y\})} x =: x_l(y)$ *and* $(x,y) \mapsto V_y(x)$ *are Borel-measurable.*

Proof. Let $y \in \mathrm{supp}(n)$. Note that for any $(x_1,u_1),(x_2,u_2) \in U_1^{-1}(\{y\})$ holds

$$(x_1,u_1) \le (x,u) \le (x_2,u_2) \implies (x,u) \in U_1^{-1}(\{y\}),$$

where $\le$ refers to the lexicographical order. Hence,

$$\gamma_y = \Big\{x \in [0,1] \mid \exists u \in [0,1] \text{ s.t. } (x,u) \in U_1^{-1}(\{y\})\Big\}.$$

Especially, γ_y is an interval with left and right boundary points x_l and x_r. When $n(\{y\}) \neq 0$, for any point $x \in (x_l, x_r)$ follows $V_y(x) = 1$. Further, γ_y contains its boundary points, since $(x_l, 1), (x_r, 0) \in U_1^{-1}(\{y\})$. If $y_1, y_2 \in \operatorname{supp}(n)$, $y_1 < y_2$ with

$$\gamma_{y_1} = [a_l, a_r],\ \gamma_{y_2} = [b_l, b_r],$$

implies that $a_r \leq b_l$. Therefore, the map $y \mapsto x_l(y)$ is monotonously increasing, and hence Borel-measurable. As a matter of fact there can only be a countable amount of point masses of n, which shows measurability of $(x, y) \mapsto V_y(x)$. ❑

To simplify notation, we set

$$\vec{x} := (x, 0, \ldots, 0) \in \mathbb{R}^I$$

for any starting point $x \in \mathbb{R}$. Given the starting distribution $\operatorname{proj}_{X_1}(\mu) = \frac{1}{2}(\delta_x + \delta_y)$ for $x \neq y$ and $\xi \in \operatorname{RST}(\mu, \nu)$ we can construct, by virtue of Z's independent increments, another probability measure $\tilde{\xi} \in \operatorname{RST}(\mu, \nu)$ such that

$$\tilde{\xi}_{\omega+\vec{x}} = \xi_{\omega+\vec{y}}, \quad \tilde{\xi}_{\omega+\vec{y}} = \xi_{\omega+\vec{x}}, \quad \omega \in \mathbb{R}^I.$$

If the starting distributions are arbitrary (sub-)probability measures, we can construct another randomized stopping time by following the idea of "swapping branches", see the following lemma.

Lemma 4.5.24. *Under the assumptions of Lemma 4.5.20, let m and n be the starting distributions of the (sub-)probability measures μ and $\tilde{\mu}$ associated with Z, i.e.,*

$$m = \operatorname{proj}_{X_1}(\mu), \quad n = \operatorname{proj}_{X_1}(\tilde{\mu}).$$

Then, for fixed $y \in \operatorname{supp}(n)$ the measure

$$m_y(\mathrm{d}x) = \begin{cases} \delta_{x_l(y)}(\mathrm{d}x) & n(\{y\}) = 0, \\ \frac{V_y(x) m(\mathrm{d}x)}{n(\{y\})} & \textit{else}, \end{cases}$$

is well-defined.
Furthermore, for every $\xi \in \operatorname{RST}(\mu, \nu)$ there exists a $\tilde{\xi} \in \operatorname{RST}(\tilde{\mu}, \nu)$ with disintegration $(\tilde{\xi}_\omega)_{\omega \in \mathbb{R}^I}$ such that for $t \in I$ and $\omega \in \mathbb{R}^I$

$$\tilde{\xi}_\omega(t) := \begin{cases} \int \xi_{\omega+\vec{x}-\vec{\omega}_1}(t) m_{\omega_1}(\mathrm{d}x) & \omega_1 \in \operatorname{supp}(n), \\ \mathbb{1}_{\{1\}}(t) & \textit{else}. \end{cases} \tag{4.5.25}$$

Proof. According to Lemma 4.5.23 the measure m_y is well-defined and $y \mapsto m_y$ measurable. Following the proof of Lemma 4.5.20, there exist $u_l, u_r \in [0,1]$ such that

$$m([0, x_l)) + u_l \cdot m(\{x_l\}) = n([0, y)),$$
$$m([0, x_r)) + u_r \cdot m(\{x_r\}) = n([0, y]),$$

where $[x_l, x_r] = \gamma_y$. We will only discuss the case that $x_l < x_r$, since the other case can be dealt in similar fashion. If $n(\{y\}) > 0$, it holds that $u_l = 1 - V_y(x_l)$ and $u_r = V_y(x_r)$.

$$m(V_y) = m([x_l, x_r)) + u_r \cdot m(\{x_r\}) - (1 - u_l) \cdot m(\{x_l\}) = n(\{y\}).$$

Thus m_y is a probability measure on $[0,1]$. As a composition of measurable functions, $\omega \mapsto \tilde{\xi}_\omega$ is measurable. By construction $\tilde{\xi}$ is $\mathbb{F}$-adapted, therefore it remains to establish the marginal properties. Let $\omega \in \mathbb{R}^I$ and $\omega_1 \in \operatorname{supp}(n)$.

$$\sum_{t\in I} \tilde{\xi}_\omega(t) \overset{(4.5.25)}{=} \int \sum_{t\in I} \tilde{\xi}_{\omega+\vec{x}-\vec{\omega}_1} m_{\omega_1}(\mathrm{d}x) = 1,$$

which implies

$$\operatorname{proj}_{\mathbb{R}^I}(\tilde{\xi})(\mathrm{d}\omega) = \tilde{\mu}(\mathrm{d}\omega).$$

Then

$$\begin{aligned} \operatorname{proj}_I(\tilde{\xi})(t) &= \int_{\mathbb{R}^I} \tilde{\xi}(\omega,t)\tilde{\mu}(\mathrm{d}\omega) = \int \tilde{\xi}_{\theta+\vec{y}}(t)\mathbb{P}(\mathrm{d}\theta)n(\mathrm{d}y) \\ &= \int\int \xi_{\theta+\vec{x}}(t) m_y(\mathrm{d}x)n(\mathrm{d}y)\mathbb{P}(\mathrm{d}\theta). \end{aligned}$$

To prove $\operatorname{proj}_I(\tilde{\xi}) = \nu(t)$ it is sufficient to show that for any interval $A := [a,b] \subseteq [0,1]$

$$m(A) = \int_{[0,1]^2} \mathbb{1}_A(x) m_y(\mathrm{d}x)n(\mathrm{d}y), \tag{4.5.26}$$

since the assertion follows then by the monotone class theorem. With Lemma 4.5.20 we obtain for $U(a,0) =: (y_1, v_1)$ and $U(b,1) =: (y_2, v_2)$. As above, we assume that $y_1 < y_2$ which yields

$$\begin{aligned} m([0,a)) &= n([0,y_1)) + v_1 \cdot n(\{y_1\}), \\ m([0,b]) &= n([0,y_2)) + v_2 \cdot n(\{y_2\}), \end{aligned}$$

cf. (4.5.21) and (4.5.22). And hence

$$\begin{aligned} &m([a,b]) = (1-v_1)\cdot n(\{y_1\}) + n((y_1,y_2)) + v_2 \cdot n(\{y_2\}), \\ &m([0,a)) \le n([0,y]) \le m([0,b]) \quad \forall y \in (y_1,y_2) \cap \operatorname{supp}(n). \end{aligned}$$

For any $y \in (y_1,y_2) \cap \operatorname{supp}(n)$, $x \in \gamma_y$ and $v \in [0,1]$, we note that

$$m([0,a]) \le n([0,y_1]) \le m([0,x)) + v\cdot m(\{x\}) \le n([0,y_2)) \le m([0,b]),$$

which implies that $\gamma_y \subseteq A$ and $m_y(\gamma_y \cap A) = 1$, and thus

$$n((y_1,y_2)) = \int_{(y_1,y_2)} \int_{[0,1]} \mathbb{1}_A(x) m_y(\mathrm{d}x)n(\mathrm{d}y).$$

In the case that $n(\{y_1,y_2\}) = 0$, the assertion follows. If y_1 or y_2 are point masses, we have to show that $(1-v_1) = m_{y_1}(A)$ and $v_2 = m_{y_2}(A)$, respectively. We only consider the case that y_l is a point mass, since the other cases (y_3 or $y_1 = y_2$ is a point mass) follow analogously. We know that $\gamma_{y_1} \cap A = [a,c] \subseteq [a,b]$ and

$$\begin{aligned} m([0,a)) &= n([0,y_1)) + v_1 \cdot n(\{y_1\}), \\ n([0,y_1]) &= m([0,c)) + u\cdot m(\{c\}). \end{aligned}$$

Based on these two equations, the assertion follows:

$$1 - v_1 = \frac{1}{n(\{y_1\})}(m([a,c]) + u \cdot m(\{c\})) = m_{y_1}(A).$$

Analogous considerations can be made for ν_2. ❑

Theorem 4.5.27. *The greedy strategy π is optimal, i.e.,*

$$\pi(c) \geq \xi(c), \quad \forall \xi \in \mathrm{RST}(\mu, \nu).$$

If f is strictly increasing, π is the unique optimizer in the following sense:

Any $\tilde{\pi} \in \mathrm{RST}(\mu,\nu)$ with $\tilde{\pi}(c) = \pi(c)$ satisfies for all $t \in I$, $t < T$

$$\mathrm{proj}_{X_t \times X_{t+1}}(\pi(\cdot,t) - \tilde{\pi}(\cdot,t)) = 0, \quad \mathrm{proj}_{X_t \times X_{t+1}}(\mu)\text{-}a.e.$$

Proof. Let $t \in I$ be fixed and assume that $\pi(\cdot,s) = \xi(\cdot,s)$, μ-a.s., for all $s < t$. Hence, there exists a $\mathcal{F}_t$-measurable set $B \subseteq \mathbb{R}^I$ with full measure such that $\pi_\omega(s) = \xi_\omega(s)$ holds pointwise for all $\omega \in B$, $s < t$.

$$A^+ := \{\omega : (\pi - \xi)_\omega(t) > 0\} \cap B, \quad A^- := \{\omega : (\xi - \pi)_\omega(t) > 0\} \cap B.$$

Define $A := A^+ \cup A^-$. In view of the (quantile) structure of π, cf. (4.5.19), it follows that

$$\omega_t \leq \eta_t \quad \forall \omega \in A^+, \ \forall \eta \in A^-. \tag{4.5.28}$$

Let the finite measures ϕ^+ and ϕ^- on $A \times I$ be given by

$$\phi^+ := (\pi - \xi)^+_{\restriction A \times I}, \quad \phi^- := (\xi - \pi)^+_{\restriction A \times I}.$$

Note that $\mathrm{proj}_{\mathbb{R}^I}(\phi^+) = \mathrm{proj}_{\mathbb{R}^I}(\phi^-)$, which implies

$$0 < \phi^+(A^+, t) =: \alpha \leq \phi^+(A^+ \times I_{\geq t}) = \phi^-(A^+ \times I_{>t}) =: \beta.$$

The second marginal property of π and ξ, i.e., $\mathrm{proj}_I(\pi)(t) = \nu(t) = \mathrm{proj}_I(\xi(t))$, yields

$$\alpha = \phi^+(A^+, t) = \phi^-(A^-, t).$$

We may define two $\mathbb{F}$-adapted measures ψ and χ via

$$\psi(\mathrm{d}\omega, s) := \frac{\alpha}{\beta} \cdot \begin{cases} \phi^-(\mathrm{d}\omega, s) & (\omega, s) \in A^+ \times I, \\ 0 & \text{else,} \end{cases}$$

$$\chi(\mathrm{d}\omega, s) := \begin{cases} \phi^-(\mathrm{d}\omega, s) & (\omega, s) \in A^- \times \{t\}, \\ 0 & \text{else.} \end{cases}$$

Due the scaling factor of $\frac{\alpha}{\beta}$, we obtain $\psi(A^+ \times I) = \chi(A^- \times I)$. Therefore, we can define m and n as the starting distributions of $\bar{\psi} := \mathrm{proj}_{\mathbb{R}^{I_{\geq t}}}(\psi)$ and $\bar{\chi} := \mathrm{proj}_{\mathbb{R}^{I_{\geq t}}}(\chi)$ such that

$$m := \mathrm{proj}_{X_t}(\bar{\psi}), \quad n := \mathrm{proj}_{X_t}(\bar{\chi}).$$

Let $C \subseteq A$ be a $\mathcal{F}_s$-measurable set for $s \geq t$ and C' its projection onto $\mathbb{R}^{I_{\geq t}}$, then

$$\psi(C) = \bar{\psi}(C'), \quad \chi(C) = \bar{\chi}(C').$$

Again by a monotone class argument it follows that

$$\psi(c) = \bar{\psi}(c'), \quad \chi(c) = \bar{\chi}(c'),$$

where c' is the natural restriction of c onto $\mathbb{R}^{I_{\geq t}} \times I_{\geq t}$. Applying Lemma 4.5.24 to m, n, $\bar{\psi}$ and $\bar{\chi}$ results in two measures $\tilde{\psi}$ and $\tilde{\chi}$, which preserve the marginals of $\bar{\psi}$ and $\bar{\chi}$, respectively. In the final step, we want to extend $\tilde{\chi}$ and $\tilde{\psi}$ to measures $\hat{\psi}$ and $\hat{\chi}$ on $\mathbb{R}^I \times I$.

$$\hat{\psi}(\mathrm{d}\omega, s) := \begin{cases} \tilde{\psi}_{\omega_{\upharpoonright [t,T]}}(s)\,\mathrm{proj}_{\mathbb{R}^I}(\psi)(\mathrm{d}\omega) & \omega \in A^+,\ s \geq t, \\ 0 & \text{else.} \end{cases}$$

$$\hat{\chi}(\mathrm{d}\omega, s) := \begin{cases} \tilde{\chi}_{\omega_{\upharpoonright [t,T]}}(s)\,\mathrm{proj}_{\mathbb{R}^I}(\chi)(\mathrm{d}\omega) & \omega \in A^-,\ s \geq t, \\ 0 & \text{else.} \end{cases}$$

For $s \geq t$ we obtain

$$\mathrm{proj}_I(\psi + \chi)(s) = \mathrm{proj}_{I_{\geq t}}(\bar{\psi} + \bar{\chi})(s) = \mathrm{proj}_{I_{\geq t}}(\hat{\psi} + \hat{\chi})(s) = \mathrm{proj}_I(\tilde{\psi} + \tilde{\chi})(s).$$

Then

$$\mathrm{proj}_{\mathbb{R}^I}(\psi + \chi) = \mathrm{proj}_{\mathbb{R}^{I_{\geq t}}}(\tilde{\psi} + \tilde{\chi}) = \mathrm{proj}_{\mathbb{R}^{I_{\geq t}}}(\tilde{\psi} + \tilde{\chi}) = \mathrm{proj}_{\mathbb{R}^I}(\hat{\psi} + \hat{\chi})$$

holds μ-a.e. Hence, we are able to define $\tilde{\xi} \in \mathrm{RST}(\mu, \nu)$ via

$$\tilde{\xi} := \xi - \psi - \chi + \hat{\psi} + \hat{\chi}.$$

$$\begin{aligned}
&\psi(c) + \chi(c) = \bar{\psi}(c') + \bar{\chi}(c') \\
&= \int c'(\theta + \vec{y}, 1)\bar{\psi}_{\theta+\vec{y}}(1)\mathbb{P}(\mathrm{d}\theta)n(\mathrm{d}y) + \int c'(\theta + \vec{x}, s))\bar{\chi}_{\theta+\vec{x}}(s)\mathbb{P}(\mathrm{d}\theta)m(\mathrm{d}x) \\
&\leq \int c'(\theta + \vec{y}, s)\bar{\chi}_{\theta+\vec{x}}(s)\mathbb{P}(\mathrm{d}\theta)m_y(\mathrm{d}x)n(\mathrm{d}y) + \int c'(\theta + \vec{x}, 1)\bar{\psi}_{\theta+\vec{y}}(1)\mathbb{P}(\mathrm{d}\theta)n_x(\mathrm{d}y)m(\mathrm{d}x) \\
&= \tilde{\psi}(c') + \tilde{\chi}(c') = \hat{\psi}(c) + \hat{\chi}(c),
\end{aligned}$$

where the inequality holds due to (4.5.28), which implies

$$x \geq y \quad \forall x \in \mathrm{supp}(m),\ y \in \mathrm{supp}(n),$$

together with the explicit form of c' and (4.5.26). We see that we can exchange ξ with $\tilde{\xi}$ without lowering the payoff. Further, $\tilde{\xi}(\omega, s) = \pi(\omega, s)$ ω-a.e. for all $s \leq t$. By continuing this procedure iteratively over t, we can transform the initial ξ into π without lowering the payoff. Hence, π is optimal. If f is monotonously strictly increasing, this inequality holds strictly if and only if

$$\mathrm{proj}_{X_t \times X_{t+1}}(\pi(\cdot, t) - \xi(\cdot, t)) \neq 0, \quad \mathrm{proj}_{X_t \times X_{t+1}}(\mu) - \text{a.e.}$$

Thus, π is the unique optimizer in the sense described above. ❑

4.6. Monotonicity Principle in an Example

To test if a randomized stopping time is a possible candidate for optimality in problem OptStop$^\pi$, different monotonicity criteria were developed. In this context, the so-called c-cyclical monotonicity as in [73] deserves a special mention, which is in fact a geometric property of the support of an optimal transport plan. In the initial form, the monotonicity was shown only for couplings which do not have to satisfy additional adaptivity constraints. Zaev introduced (c, W)-cyclical monotonicity in [76, Theorem 3.6], which enhances the notion with constraints, denoted by W. In our considerations, randomized stopping times are couplings satisfying additional linear constraints given through (4.4.2) and (4.4.3). Thus, Zaev's monotonicity principle can be applied naturally. Contrary to the classical c-monotonicity, the (c, W)-monotonicity of a support of a randomized stopping time is a necessary optimality condition, but in general not sufficient. In an independent work, Beiglböck and Griessler found a closely related monotonicity principle which includes the result [76, Theorem 3.6] as a special case, see [11, Theorem 1.4].
Inspired by the classical c-monotonicity which shows that optimality is an attribute of the support of a coupling, other different monotonicity principles have been developed in the area of martingale optimal transport problems, cf. [9] and Chapter 7. The approach of the previous section for showing optimality is strongly inspired by the latter chapter of this thesis which deals with time-continuous distribution-constrained optimal stopping problems where the underlying stochastic process is a Brownian motion. Analogously, it is possible to find a monotonicity principle for the time-discrete case. Again, we assume that $(Z_t)_{t\in I}$ is a stochastic process in discrete time with independent increments.

Definition 4.6.1. The set RST^t_κ of randomized stopping times (of a stochastic process Z with initial distribution κ) is defined as the set of all Z-adapted probability measures π on $\mathbb{R}^{I_{\geq t}} \times I_{\geq t}$ such that $Z \sim \mathrm{proj}_{\mathbb{R}^{I_{\geq t}}}(\pi)$.

Definition 4.6.2 (Concatenation)**.**
For every $t \in I$ we have an operation $\odot$ of concatenation, which is a map into $\mathbb{R}^{I_{\geq t}}$ and is defined for $(\omega, s) \in \mathbb{R}^{I_{\geq t}} \times I_{\geq t}$ and $\theta \in \mathbb{R}^{I_{\geq s}}$ with $\theta(s) = 0$ by

$$((\omega, s) \odot \theta)(r) = \begin{cases} \omega(r) & t \leq r \leq s, \\ \omega(s) + \theta(r) & r > s. \end{cases} \tag{4.6.3}$$

Definition 4.6.4 (Conditional randomized stopping times)**.**
For $\pi \in \mathrm{RST}(\mu, \nu)$ and $(\omega, t) \in S$, we define $\pi^{(\omega,t)} \in \mathrm{RST}^t$ by defining a disintegration $(\pi^{(\omega,t)}_\theta)_{\theta\in\mathbb{R}^{I_{\geq t}}}$ with respect to $\tilde{Z}$ as

$$\pi^{(\omega,t)}_\theta := \begin{cases} \frac{1}{1-\pi_{(\omega,t)}(I_{\leq t})}(\pi_{(\omega,t)\odot\theta})\restriction_{I_{\geq t}} & \text{for } \pi_{(\omega,t)}(I_{\leq t}) < 1, \\ \delta_t & \text{for } \pi_{(\omega,t)}(I_{\leq t}) = 1, \end{cases}$$

where δ_t is the Dirac measure concentrated at t and $\theta \in \mathbb{R}^{I_{\geq t}}$ with $\theta_1 = 0$.

Definition 4.6.5 (Relative Stop-Go pairs)**.**
For $\xi \in \mathrm{RST}(\mu,\nu)$ define $\mathrm{SG}^{\xi} \subseteq (\mathbb{R}^I \times I) \times (\mathbb{R}^I \times I)$ as the set of all pairs $(\omega,t),(\eta,t) \in \mathbb{R}^I \times I$ such that there exist $\tilde{\xi}_1 \in \mathrm{RST}^t_{\delta_{\omega(t)}}$ and $\tilde{\xi}_2 \in \mathrm{RST}^t_{\delta_{\eta(t)}}$ such that

- $\mathrm{proj}_{\{t,\dots,T\}}(\xi^{(\omega,t)} + \xi^{(\eta,t)}) = \mathrm{proj}_{\{t,\dots,T\}}(\tilde{\xi}_1 + \tilde{\xi}_2)$,
- $\xi^{(\omega,t)}(c) + \xi^{(\eta,t)}(c) < \xi_1(c) + \xi_2(c)$.

Theorem 4.6.6 (Monotonicity Principle)**.**
Assume that π is a solution of $\mathrm{O{\scriptsize PT}S{\scriptsize TOP}}^{\pi}$*, then there is a measurable, $\mathbb{F}$-adapted set $\Gamma \subseteq \mathbb{R}^I \times I$ such that*

$$\pi(\Gamma) = 1$$

and

$$SG \cap (\Gamma^{<} \times \Gamma) = \emptyset,$$

where $\Gamma^{<} := \{(\omega,s) \in \mathbb{R}^I \times I \; : \; (\omega,t) \in \Gamma \text{ for some } t > s\}$.

Equipped with this general result, we can easily show optimality of the greedy strategy introduced in the last section. For this class of payoff functions in particular, it can be shown that monotonicity is already a sufficient condition for being an optimizer.

Corollary 4.6.7. *Let the payoff function c be given as*

$$c(\omega,t) = f(t)\omega_t \quad \omega \in \mathbb{R}^I,\; t \in I,$$

where $f : I \to \mathbb{R}^+$ is monotonously increasing. Then the greedy strategy $\pi \in \mathrm{RST}(\mu,\nu)$ is a maximizer of $\mathrm{O{\scriptsize PT}S{\scriptsize TOP}}^{\pi}$*. If $\xi \in \mathrm{RST}(\mu,\nu)$ satisfies the assertions of Theorem 4.6.6, then*

$$\mathrm{proj}_{X_t \times X_{t+1}}(\xi(\cdot,t)) = \mathrm{proj}_{X_t \times X_{t+1}}(\pi(\cdot,t)), \quad \mathrm{proj}_{X_t \times X_t + 1}(\mu)\text{-}a.e.,\; t \in I.$$

Proof. From the construction of π via quantiles, see Example 4.5.3, there exists $A_t := (-\infty, a_t] \subseteq \mathbb{R},\; t \in I$ such that $M_t = \prod_{s<t}[a_s,\infty) \times A_t \times \prod_{r>t}\mathbb{R}$, each a_t is minimal with

$$\mu\left(\sum_{s \le t} M_s\right) \ge \sum_{s \le t} \nu(s).$$

Let $\xi \in \mathrm{RST}(\mu,\nu)$ be optimal, then we denote the set of Theorem 4.6.6 with Γ. Assume that

$$\mathrm{proj}_{X_s \times X_{s+1}}(\xi(\cdot,s) - \pi(\cdot,s)) = 0, \quad \mathrm{proj}_{X_s \times X_{s+1}}(\mu)\text{-a.s.},\; s < t < T,$$
$$\mathrm{proj}_{X_t \times X_{t+1}}(\xi(\cdot,t) - \pi(\cdot,t)) \neq 0\, quad\, \mathrm{proj}_{X_t \times X_{t+1}}(\mu)\text{-a.e.}$$

Then there exists $\omega \in \Gamma$ such that

$$\omega_s \in (a_s,\infty), \quad s < t, \quad \omega(t) \in [a_t,\infty) \text{ and } \xi_\omega(t) > 0.$$

Since ξ has to preserve the marginals, there exists $\eta \in M_t \cap \Gamma$ such that $\xi_\eta(s) > 0$ for an $s > t$, which yields $\eta_t < \omega_t$. We want to show that $((\eta,t),(\omega,t)) \in \mathrm{SG}^{\xi}$ which would lead to a contradiction. Therefore, we construct $\xi^1 \in \mathrm{RST}^t_{\delta_{\omega_t}}$ and $\xi^2 \in \mathrm{RST}^t_{\delta_{\eta_t}}$ by defining two disintegrations

$$\xi^1_\theta(s) := \xi^{(\omega,t)}_\theta(t) \cdot \xi^{(\eta,t)}_\theta(s) + \begin{cases} 0 & s = t, \\ \xi^{(\omega,t)}_\theta(s) & s > t. \end{cases}$$
$$\xi^2_\theta(s) := (1 - \xi^{(\omega,t)}_\theta(t)) \cdot \xi^{(\eta,t)}_\theta(s) + \begin{cases} \xi^{(\omega,t)}_\theta(t) & s = t, \\ 0 & s > t. \end{cases}$$

Computing the payoff yields

$$\begin{aligned}
&\xi^1(c) - \xi^{(\omega,t)}(c) + \xi^2(c) - \xi^{(\eta,t)}(c) \\
&= -\int f(t)\omega_t \xi^{(\omega,t)}_\theta(t) \cdot \left(1 - \xi^{(\eta,t)}_\theta(t)\right) d\mathbb{P}(\theta) + \sum_{t>s} \int f(s)(\omega_t + \theta_s)\xi^{(\omega,t)}_\theta(t) \cdot \xi^{(\eta,t)}_\theta(s)\, d\mathbb{P}(\theta) \\
&\quad + \int f(t)\eta_t \xi^{(\omega,t)}_\theta(t) \cdot \left(1 - \xi^{(\eta,t)}_\theta(t)\right) d\mathbb{P}(\theta) - \sum_{t>s} \int f(s)(\eta_t + \theta_s)\xi^{(\omega,t)}_\theta(t) \cdot \xi^{(\eta,t)}_\theta(s)\, d\mathbb{P}(\theta) \\
&= \sum_{s>t} \int (f(s) - f(t))(\omega_t - \eta_t)\xi^{(\omega,t)}_\theta(t)\xi^{(\eta,t)}(\mathrm{d}\omega, s) > 0.
\end{aligned}$$

Therefore, $((\eta,t),(\omega,t)) \in \mathrm{SG}^\xi \cap (\Gamma^< \times \Gamma)$ which is a contradiction. ❑

Corollary 4.6.8. *Under the assumption of Corollary 4.6.7, the support of the greedy strategy $\pi \in \mathrm{RST}(\mu,\nu)$ introduced in Example 4.5.3 satisfies the assertions of Theorem 4.6.6.*

Proof. Let $((\eta,t),(\omega,t)) \in \Gamma^< \times \Gamma$ and $\kappa := \frac{1}{2}(\delta_{\eta_t} + \delta_{\omega_t})$, then $\frac{1}{2}(\pi^{(\eta,t)} + \pi^{(\omega,t)}) =: \tilde{\pi} \in \mathrm{RST}^t_\kappa$ can be viewed as a greedy strategy to the auxiliary problem:

$$\text{Maximize } \xi \mapsto \xi(c) \text{ under } \xi \in \mathrm{RST}^t_\kappa \text{ s.t. } \mathrm{proj}_{I_{\geq t}}(\xi - \tilde{\pi}) = 0.$$

By applying Corollary 4.6.7 we obtain optimality of $\tilde{\pi}$, and

$$\mathrm{SG}^\pi \cap (\Gamma^< \times \Gamma) = \emptyset.$$

❑

4.7. Conclusion and Outlook

In Chapter 4.2, the optimization problems OptStop^{γ} and OptStop^{π} were formally introduced and a link between them was made. Based on the theory of optimal transport, the existence of optimizers is shown in Chapter 4.3. In addition, a Kantorovich-type duality theorem was developed, inspired by recent work of Zaev [76]. Chapter 4.5 deals with the optimization of a class of payoff functions. For this, an optimal strategy was found in Theorem 4.5.27. Finally, in Chapter 4.6 different geometric optimality criteria – so-called monotonicity principles – are formulated, which have their roots in the theory of optimal transport. It is shown how they can be adapted for OptStop^{π} and are applied to show optimality of the strategy introduced in Section 4.5.3.

We close with an outlook on future research possibilities. Considerations were made on more general monotonicity principles, for example on the monotonicity principle for Gozlan-type problem. These considerations have not been completed yet. Furthermore, efforts were made by Källblad [38] in the area of time-continuous distribution-constrained optimal stopping problems to reformulate the problem using so-called measure-valued martingales. In this current work [38], the following optimal stopping problem with a constraint placed on the distribution of the stopping time is considered: given a probability measure μ on $(0,\infty)$ (that corresponds to our measure ν on I) and a filtered probability space supporting a Brownian motion $(B_t)_{t\geq 0}$, we aim at finding

$$\sup_{\tau\in\mathcal{T}^{\mu}} \mathbb{E}[c(B_\cdot,\tau)], \tag{4.7.1}$$

where $\mathcal{T}^{\mu}$ is the set of stopping times with distribution μ and c is a given measurable cost function satisfying $c(\omega,t)=c(\omega_{\cdot\wedge t},t)$. This is a special case of our problem OptStop^{γ}.
Each stopping time τ in $\mathcal{T}^{\mu}$ is identified with the measure-valued martingale $(\xi_t)_{t\geq 0}$ defined as its conditional distribution given the current information:

$$\xi_t = \mathcal{L}(\tau \mid \mathcal{F}_t); \tag{4.7.2}$$

any such process will satisfy the initial condition $\xi_0=\mu$ along with a martingale property and a certain adaptedness condition corresponding to the stopping-time property of τ. When reformulating the optimal stopping problem as an optimization problem over such measure-valued martingales (MVMs), the distribution-constraint is then incorporated as an initial condition which allows the problem to be addressed as a stochastic control problem; the main result establishs that the dynamic programming principle holds for this problem.It should be outlined here how this could apply. If you transfer these consideration to the enlarged problem OptStop^{γ}, there arise interesting connections. Analogously we get the following for our problem OptStop^{γ} and notation: given a probability measure ν on I and a filtered probability space supporting a Brownian motion $(B_t)_{t\in I}$, we aim at finding

$$\sup_{\gamma\in\mathcal{M}_I^{\nu}} \mathbb{E}[c(B_\cdot,\gamma)], \tag{4.7.3}$$

where $\mathcal{M}_I^{\nu}$ is the set of adapted random probability measures with distribution ν and c is a given measurable cost function satisfying $c(\omega,t)=c(\omega_{\cdot\wedge t},t)$. Each adapted random probability measure γ in $\mathcal{M}_I^\nu$ is identified with the measure-valued martingale $(\xi_t)_{t\in I}$ defined as its conditional distribution given the current information:

$$\xi_t(A) = \mathbb{E}\Big[\sum_{u\in A}\gamma_u \;\Big|\; \mathcal{F}_t\Big]; \quad \forall t\in I, A\in\mathcal{B}(I), \tag{4.7.4}$$

any such process will satisfy the initial condition $\xi_0=\nu$ along with a martingale property and a certain adaptedness condition corresponding to the property of γ.
The process $(\xi_t)_{t\in I}$ defined in (4.7.4) is a measure-valued martingale, because it satisfies the following definition with $\mathcal{P}$ denotes the set of probability measures on I with finite first moment, which can easily checked:

Definition 4.7.5 (MVM)**.** Cf. [38, Definition 3.1]:
Given a filtered probability space supporting an adapted process $(\xi_t)_{t\in I}$ with $\xi_t\in\mathcal{P}$, we say that

- the process ξ is a measure-valued martingale (MVM) if $\xi_.(A)$ is a martingale, for any $A\in\mathcal{B}(I)$;
- a MVM is continuous if $t\mapsto\xi_t$ is continuous in the topology induced by $\mathcal{W}_1$, the first Wasserstein metric, for almost all $\omega\in\Omega$;
- a MVM is adapted if $\xi_t(I_{\le s})=\xi_u(I_{\le s})$ a.s. , for all $s\le t\le u$.

By Section 3.2 we know that $\textsc{OptStop}^\gamma$ coincides with $\textsc{OptStop}^\pi$. We will focus on these considerations. Let $T:=\sup(I)$. Each randomized stopping time π in $\mathrm{RST}(\mu,\nu)$ is identified with the measure-valued martingale $(\xi_t)_{t\in I}$ defined as its conditional distribution given the current information:

$$\xi_{T,\omega}=\pi_\omega; \quad \forall\omega\in\mathcal{B}(\mathbb{R}^I), \tag{4.7.6}$$
$$\xi_t=\mathbb{E}[\pi\mid\mathcal{F}_t]; \quad \text{a.s. } \forall t\in I, \tag{4.7.7}$$

any such process will satisfy the initial condition $\xi_0=\nu$ along with a martingale property and a certain adaptedness condition corresponding to the property of π.

Let $\mathrm{MVM}(\nu)$ denote the set of adapted measure-valued martingales ξ with $\xi_0=\nu$. Then we have to show that our original problem $\textsc{OptStop}^\pi$ (cf. [38, Problem 3.2]), indeed, admits the following equivalent formulation:

Problem ($\textsc{OptStop}^\xi$)**.** On the given space $(\Omega=\mathbb{R}^I,\mathcal{F},\mathbb{F},\mu)$, consider the problem of maximizing

$$\mathbb{E}\Big[\sum_I c(B_{\cdot\wedge s},s)\,dA_s^\xi\Big] \quad \text{with } A_t^\xi:=\xi_t(I_{\le t}), \quad \text{over } \xi\in\mathrm{MVM}(\nu). \tag{4.7.8}$$

Finally, this would be a further approach or formulation for our considered problem. Furthermore, it would be interesting in which form [38, Corollary 3.9] is applicable. In this case, the Relation to Bayraktar and Miller [8] should also be considered.

II.

Adapted Dependence in Continuous Time

5
The Problem

In this chapter we introduce our problem of study in continuous time, cf. [33, Section 7.1]. The distribution-constrained optimization problem, which we consider, is a modified version of an optimal stopping problem. We deal with financial and actuarial products, whose payoffs taking value during a certain time interval, are determined by an stochastic process. The time point of the payoff is modeled by a stopping time. This stopping time follows a given distribution and can depend on the underlying process. Our target is to deduce the estimation of the worst-case situation, that means, the supremum of the expected payoff over all stopping times satisfying the given marginals. We are interested in sufficient conditions such that there exists a maximizer. As in discrete time, the problem is denoted by OptStop^{τ}, because it does not change significantly. In discrete time we have used adapted random probability measure to extend the problem or alternatively formulate the problem of several withdrawals within the predefined time interval. Now, we have to exchange the adapted random probability measures by stochastic transition kernels to formulate OptStop^{γ}. Afterwards some general results are presented for OptStop^{γ} in Chapter 6. We can formulate these problems again as optimal transport problems and prove the existence of an optimal strategy for these problems by using the theory of optimal transport, see Chapter 7. This chapter is the main part and the results of it are already published ([10]). But first of all and similar to discrete time, we introduce the notational conventions and necessary assumptions for this part and look at the distribution-constrained optimal stopping problem OptStop^{τ}.

Notation 5.0.1. Throughout this part, we consider a continuous time setting and stick to the following notation.

(a) Let $I \subseteq \mathbb{R}$ (or $I \subseteq [0,\infty)$) denote a continuous time interval.

(b) For $t \in I$ we define the set $I_{<t} = (-\infty, t) \cap I$ of all times before t, the set $I_{\leq t} = (-\infty, t] \cap I$ of all times up to t, the set $I_{\geq t} = [t, \infty) \cap I$ of all times from t on, and the set $I_{>t} = (t, \infty) \cap I$ of all times after t.

(c) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in I}$.

(d) $\mathcal{T}_I$ denotes the set of all stopping times $\tau : \Omega \to I$. For a given probability distribution ν on I, let $\mathcal{T}_I^{\nu}$ be the set of all I-valued stopping times with distribution ν, i.e., $L(\tau) = \nu$.

Let $Z = (Z_t)_{t \in I}$ be the adapted process with $Z \in L^1(\mathbb{P})$, i.e., $\mathbb{E}[|Z_t|] < \infty$ for all $t \in I$, which will be needed from time to time. Furthermore, in continuous time we will often need to assume that the process Z is càdlàg or that it is at least right-continuous.
We now assume that the expected values we are interested in exist and are finite. Using stopping times in continuous time does not significantly change the formulation of our problem. Since the problems are similar, we will use the same notation as in discrete time. Then the value of a classical optimal stopping problem, which we will denote by $V_{\mathcal{T}}(Z)$, is given by

$$V_{\mathcal{T}}(Z) := \sup_{\tau \in \mathcal{T}_I} \mathbb{E}[Z_\tau].$$

Moreover, the distribution-constrained optimal stopping problem is given in the following way.

Problem (OptStop$^\tau$)**.** Consider a real-valued and $\mathbb{F}$-adapted stochastic process $Z = (Z_t)_{t \in I}$ such that $\mathbb{E}[Z_\tau^+]$ or $\mathbb{E}[Z_\tau^-]$ is finite for every $\tau \in \mathcal{T}_I^\nu$. Find sufficient conditions such that among all stopping times $\tau \in \mathcal{T}_I^\nu$ there exists a maximizer τ^* solving

$$\mathbb{E}[Z_{\tau^*}] = \sup_{\tau \in \mathcal{T}_I^\nu} \mathbb{E}[Z_\tau] =: V_{\mathcal{T}}^\nu(Z).$$

If $\mathcal{T}_I^\nu = \emptyset$, we set $V_{\mathcal{T}}^\nu(Z) = -\infty$.

Then the connection between the standard and distribution-constrained optimal stopping problem is given for a process Z by

$$V_{\mathcal{T}}^\nu(Z) := \sup_{\tau \in \mathcal{T}_I^\nu} \mathbb{E}[Z_\tau] \leq \sup_{\tau \in \mathcal{T}_I} \mathbb{E}[Z_\tau] =: V_{\mathcal{T}}(Z). \tag{5.0.2}$$

Note that we set $\sup_{\tau \in \mathcal{T}_I^\nu} \mathbb{E}[Z_\tau] = -\infty$ in the case $\mathcal{T}_I^\nu = \emptyset$.

6

Adapted Random Probability Measure

In this chapter we want to introduce our distribution-constrained optimization problem OptStop^{γ} in continuous time similar to the one in discrete time. For this we have to exchange the adapted random probability measures by stochastic transition kernels. Some preliminaries to stochastic transition kernels are given in Section A.5 at the beginning of this work. Similar to discrete time, the formulation and some results are stated in Section 6.1. In Section 6.2 we give some results for a special class of processes.

6.1. The Problem

For the similar definition of $V^{\nu}_{\mathcal{M}}(Z)$ and formulation of OptStop^{γ} we have to exchange the adapted random probability measures by stochastic transition kernels, which are given in the following definition.

Definition 6.1.1. Cf. [33, Definition 7.6]: For a fixed probability measure ν on $\mathcal{B}_I$ we say that a stochastic transition kernel $\Gamma : \Omega \times \mathcal{B}_I \to [0,1]$ is in $\mathcal{M}^{\nu}_I$ if for all $t \in I$

(a) $\Omega \ni \omega \mapsto \Gamma(\omega, I_{\leq t})$ is $\mathcal{F}_t$-measurable,

(b) $\mathbb{E}[\Gamma(\cdot, I_{\leq t})] = \nu(I_{\leq t})$.

For a $(\mathcal{F} \otimes \mathcal{B}_I)$-measurable process $Z : \Omega \times I \to \mathbb{R}$ and $\Gamma \in \mathcal{M}^{\nu}_I$ with

$$\mathbb{P}\left(\left\{ \int_I Z^-_t \,\Gamma(\cdot, dt) < \infty \right\} \cup \left\{ \int_I Z^+_t \,\Gamma(\cdot, dt) < \infty \right\}\right) = 1$$

we define

$$Z_\Gamma := \int_I Z_t \,\Gamma(\cdot, dt). \tag{6.1.2}$$

For an adapted process Z with $\mathbb{E}[\int_I Z^-_t \Gamma(\cdot, dt)] < \infty$ or $\mathbb{E}[\int_I Z^+_t \Gamma(\cdot, dt)] < \infty$ for all $\Gamma \in \mathcal{M}^{\nu}_I$, we are now interested in the value

$$V^{\nu}_{\mathcal{M}}(Z) := \sup_{\Gamma \in \mathcal{M}^{\nu}_I} \mathbb{E}[Z_\Gamma]. \tag{6.1.3}$$

Remark 6.1.4. See [33, Remark 7.7]: For the definition of Z_Γ it is sufficient to assume that Z is a $(\mathcal{F} \otimes \mathcal{B}_I)$-measurable process. Furthermore it is necessary to assume that the process Z is adapted to find a reasonable optimal stochastic transition kernel Γ^* satisfying

$$\mathbb{E}[Z_{\Gamma^*}] = \sup_{\Gamma \in \mathcal{M}_I^\nu} E[Z_\Gamma].$$

Then there exists a modification of Z, which is progressively measurable (see Definition A.2.6).

Using the preliminaries in Section A.5 we get the following:

Remark 6.1.5. Note that $(\Omega, \mathcal{F})$ and $(I, \mathcal{B}_I)$ are measurable spaces. In addition, $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$ and Γ is a stochastic transition kernel from $\Omega \times \mathcal{B}_I$ to I.

(a) Applying Lemma A.5.11 for $f : \Omega \times I \to [0, \infty]$ with $f(\omega, t) = Z_t(\omega)$ we know that the map

$$\omega \mapsto \int Z_t(\omega)\Gamma(\omega, dt)$$

is well-defined and $\mathcal{F}$-measurable, because Z is $(\mathcal{F} \otimes \mathcal{B}_I)$-measurable.

(b) Using Corollary A.5.14 we have that $\mathbb{P} \otimes \Gamma$ is a probability measure on $(\Omega \times I, \mathcal{F} \otimes \mathcal{B}_I)$ and uniquely determined by

$$\mathbb{P} \otimes \Gamma(A_1 \times A_2) = \int_{A_1} \Gamma(\omega, A_2)\mathbb{P}(d\omega) \quad \text{for all } A_1 \in \mathcal{F}, A_2 \in \mathcal{B}_I.$$

(c) Let Z be non-negative or in $L^1(\mathbb{P} \otimes \Gamma)$. Then the thoughts above and the Fubini for transition kernels, see Theorem A.5.18, allow the given a.s. definition of Z_Γ in Definition 6.1.1, because

$$\begin{aligned}\mathbb{E}[Z_\Gamma] &= \int_{\Omega \times I} Z_t(\omega)\, \mathbb{P} \otimes \Gamma(d\omega, dt) = \int_{\Omega \times I} Z_t(\omega)\, d(\mathbb{P} \otimes \Gamma)(\omega, t) \\ &\overset{\text{Theorem A.5.18}}{=} \int_\Omega \left(\int_I Z_t(\omega)\Gamma(\omega, dt) \right) \mathbb{P}(d\omega) = \mathbb{E}\left[\int_I Z_t \Gamma(\cdot, dt) \right].\end{aligned}$$

Note that the expectations are with respect to different measures.

(d) Note that for all $\Gamma \in \mathcal{M}_I^\nu$ the marginal/projection from $\mathbb{P} \otimes \Gamma$ to I is also a probability measure, because of the definition of Γ. It holds that

$$\text{proj}_I(\mathbb{P} \otimes \Gamma) = \nu \quad \text{for all } \Gamma \in \mathcal{M}_I^\nu,$$

such that ν can be rewritten as

$$\nu(B) = \mathbb{E}[\Gamma(\cdot, B)] = \int_\Omega \Gamma(\omega, B)\, d\mathbb{P}(\omega) \quad \text{for all } B \in \mathcal{B}_I.$$

Then the distribution-constrained optimization problem $\textsc{OptStop}^{\gamma}$ is given in the following way:

Problem ($\textsc{OptStop}^{\gamma}$)**.** Consider a real-valued and $\mathbb{F}$-adapted stochastic process $Z = (Z_t)_{t\in I}$. Find sufficient conditions such that:

(a) For every $\Gamma \in \mathcal{M}_I^{\nu}$, the integral $\int_I Z_t \Gamma(\cdot, dt)$ defining Z_Γ, is $\mathbb{P}$-a.s. absolutely convergent in $\overline{\mathbb{R}}$ satisfying $\mathbb{E}[\int_I Z_t^- \Gamma(\cdot, dt)] < \infty$ or $\mathbb{E}[\int_I Z_t^+ \Gamma(\cdot, dt)] < \infty$.

(b) There exists a maximizer $\Gamma^* \in \mathcal{M}_I^{\nu}$ solving

$$\mathbb{E}[Z_{\Gamma^*}] = \sup_{\Gamma \in \mathcal{M}_I^{\nu}} \mathbb{E}[Z_\Gamma]. \tag{6.1.6}$$

As in discrete time, we have that $\mathcal{T}_I^{\nu}$ can be embedded in $\mathcal{M}_I^{\nu}$ and $\mathcal{M}_I^{\nu}$ is not empty.

Example 6.1.7. Given an $\mathbb{F}$-stopping time $\tau \in \mathcal{T}_I^{\nu}$, it can be naturally identified with the stochastic transition kernel $\Gamma \in \mathcal{M}_I^{\nu}$ defined by

$$\Gamma(\omega, t) = \mathbb{1}_{\{\tau(\omega)\}}(t), \quad \omega \in \Omega, \ t \in I. \tag{6.1.8}$$

Example 6.1.9. The set $\mathcal{M}_I^{\nu}$ is never empty, because it contains the stochastic transition kernel Γ defined by $\Gamma(\cdot, t) = \nu_t \mathbb{1}_\Omega$ for all $t \in I$.

Therefore, the introduced distribution-constrained optimization problem $\textsc{OptStop}^{\gamma}$ is an enlargement of the problem $\textsc{OptStop}^{\tau}$ and it holds obviously that $V_{\mathcal{T}}^{\nu}(Z) \le V_{\mathcal{M}}^{\nu}(Z)$. In [33, Chapter 7] and [33, Chapter 8], there are already discussed some elementary general results and also bounds for special cases. Similar to discrete time the value of the classical optimal stopping problem, which we have denoted by $V_{\mathcal{T}}(Z)$, is an upper bound for $V_{\mathcal{T}}^{\nu}(Z)$ and also for $V_{\mathcal{M}}^{\nu}(Z)$, which is proven with [33, Lemma 8.17]. Furthermore it holds again that $V_{\text{ind}}^{\nu}(Z)$ gives us a lower bound. For every $\Gamma \in \mathcal{M}_I^{\nu}$ which is independent of Z and satisfies $\mathbb{E}[\int_I Z_t^- \Gamma(\cdot, dt)] < \infty$ or $\mathbb{E}[\int_I Z_t^+ \Gamma(\cdot, dt)] < \infty$ we get by using [33, Lemma 7.8] that

$$V_{\text{ind}}^{\nu}(Z) = \int_I \mathbb{E}[Z_t]\, \nu(dt).$$

Then we have for independent Z and $\tau \in \mathcal{T}_I^{\nu} \neq \emptyset$ or $\Gamma \in \mathcal{M}_I^{\nu}$ satisfying $\mathbb{E}[\int_I Z_t^- \Gamma(\cdot, dt)] < \infty$ or $\mathbb{E}[\int_I Z_t^+ \Gamma(\cdot, dt)] < \infty$

$$V_{\text{ind}}^{\nu}(Z) \le V_{\mathcal{T}}^{\nu}(Z) \le V_{\mathcal{M}}^{\nu}(Z).$$

In discrete time we have seen that we can also embed the original set $\mathcal{M}_I^{\nu}$ into a set $\tilde{\mathcal{T}}_I^{\nu}$ corresponding to an enlarged filtration, if the underlying process Z retains its original measurability. We also get this in continuous time by [33, Theorem 7.12], which is given in the following way:

Theorem 6.1.10. *Let $I \subseteq [0,\infty)$ and Γ be an adapted stochastic transition kernel with respect to the filtration $(\mathcal{F}_t)_{t\in I}$. By extending the probability space if necessary, we may assume w.l.o.g. that there exists a random variable U, uniformly distributed on $[0,1]$ and independent of $\mathcal{F}_\infty := \sigma(\bigcup_{t\in I} \mathcal{F}_t)$. Then*

$$\tau(\omega) := \inf\{t \in I \mid U(\omega) \le \Gamma(\omega, I_{\le t})\}, \quad \omega \in \Omega,$$

satisfies $\{\tau \le t\} = \{U \le \Gamma(\cdot, I_{\le t})\}$ *for every* $t \in I$*, hence* τ *is a stopping time with respect to the filtration* $\tilde{\mathcal{F}} = (\tilde{\mathcal{F}}_t)_{t\in I}$ *defined by* $\tilde{\mathcal{F}}_t := \sigma(\mathcal{F}_t \cup \sigma(U))$ *for* $t \in I$ *and satisfies* $\mathbb{P}(\tau \le t|\mathcal{F}_t) \overset{a.s.}{=} \Gamma(\cdot, I_{\le t})$ *for all* $t \in I$*. Let* $Z : \Omega \times I \to \mathbb{R}$ *be an* $(\mathcal{F}_\infty \otimes \mathcal{B}_I)$*-measurable process such that* $\mathbb{E}[Z_\tau^-] < \infty$*. Then*

$$\mathbb{E}[Z_\tau|\mathcal{F}_\infty] \overset{a.s.}{=} Z_\Gamma \quad \text{and} \quad \mathbb{E}[Z_\tau] = \mathbb{E}[Z_\Gamma].$$

In [33, Section 8.3], there is given a discrete approximation, which will allow us to transfer results from discrete time to continuous time. We want to specify this one again here and its application given by [33, Lemma 8.13].

Proposition 6.1.11. *See [33, Proposition 8.10]:*
Given a continuous time interval $I \subseteq [0,\infty)$ *with* $0 \in I$*. Let* $0 = t_0^{(n)} < t_1^{(n)} < \ldots < t_{m_n}^{(n)}$ *be a partition of the time interval* I*, such that the length of the corresponding subintervals tends to zero as* $n \to \infty$ *and that* $t_{m_n}^{(n)} \to \sup(I)$ *for* $n \to \infty$ *in case* $\sup(I) \in I$ *or* $t_{m_n}^{(n)} \to \infty)$ *for* $n \to \infty$ *in case* $I = [0,\infty)$*. Given a stochastic transition kernel* Γ*, for a fixed* $n \in \mathbb{N}$ *define a discrete adapted random probability measure* γ^n *by*

$$\gamma^n_{t_k^{(n)}} = \begin{cases} \Gamma(\cdot,\{0\}) & \text{if } k = 0, \\ \Gamma(\cdot,[0,t_k^{(n)}]) - \Gamma(\cdot,[0,t_{k-1}^{(n)}]) & \text{if } k = 1,\ldots,m_n - 1, \\ \Gamma(\cdot,I) - \Gamma(\cdot,[0,t_{m_n-1}^{(n)}]) & \text{if } k = m_n. \end{cases}$$

Define a sequence of stochastic transition kernels $(\Gamma^n)_{n\in\mathbb{N}}$ *by*

$$\Gamma^n = \sum_{k=0}^{m_n} \gamma^n_{t_k^{(n)}} \delta_{t_k^{(n)}},$$

where $\delta_{t_k^{(n)}}$ *denotes the Dirac measure. Then for every right-continuous process* $Z = (Z_t)_{t\in I}$ *with* $\mathbb{E}[\sup_{t\in I}|Z_t|] < \infty$

$$\lim_{n\to\infty} \int_I Z_t \Gamma^n(dt) = \int_I Z_t \Gamma(dt) \quad \text{pointwise on } \Omega \text{ and in } L^1.$$

Then it holds that:

Lemma 6.1.12. *See [33, Lemma 8.13]:*
For a right-continuous process $Z = (Z_t)_{t\in I}$ *with* $\mathbb{E}[\sup_{t\in I}|Z_t|] < \infty$ *and a discrete approximation as defined in Proposition 6.1.11 we have*

$$\mathbb{E}\bigg[\int_I Z_t \Gamma(dt)\bigg] = \mathbb{E}\bigg[\lim_{n\to\infty}\int_I Z_t \Gamma^n(dt)\bigg] = \lim_{n\to\infty} \mathbb{E}\bigg[\int_I Z_t \Gamma^n(dt)\bigg].$$

6.2. Product of a Martingale and a Deterministic Function

In this section a special class of processes are considered, for which it is possible to find an optimal strategy and the extremal value resulting from it, analogously to Section 3.5.2 of Part I where this class is considered in discrete time. Again we consider a continuous time interval $I \subseteq \mathbb{R}$ and adapted stochastic processes $Z = (Z_t)_{t \in I}$ with $Z \in L^1(\mathbb{P})$ and we assume that the process Z is càdlàg or that it is at least right-continuous. Let M be a right-continuous martingale, ν be a given distribution on I and the support of ν is defined as

$$J = \mathrm{supp}(\nu) := \{t \in I \mid \nu_t > 0\}.$$

The considered, adapted process Z is given in the form

$$Z_t = f(t) M_t, \quad t \in I, \tag{6.2.1}$$

where f is in

$$\mathcal{F}_\nu(M) := \Big\{ f : I \to \mathbb{R} \,\Big|\, f \text{ non-decreasing function,}$$
$$(\omega, t) \mapsto f(t) M_t(\omega) \text{ is } (\mathbb{P} \otimes \Gamma)\text{-integrable for all } \Gamma \in \mathcal{M}_I^\nu \Big\}.$$

Using the stochastic transition kernels to describe our problem, we are interested in

$$V_{\mathcal{M}}^\nu(Z) := \sup_{\Gamma \in \mathcal{M}_I^\nu} \mathbb{E}[Z_\Gamma] = \sup_{\Gamma \in \mathcal{M}_I^\nu} \mathbb{E}\bigg[\int_I Z_t \Gamma(\cdot, dt) \bigg].$$

The main theorem of this section will give us a characterization of an optimal strategy for the problem OptStop$^\gamma$. Before we formulate the main theorem we want to take a look at $\mathcal{F}_\nu(M)$ and some specific properties.

Lemma 6.2.2 ($\mathcal{F}_\nu(M)$)**.**

(a) For every $f \in \mathcal{F}_\nu(M)$ and $\Gamma \in \mathcal{M}_I^\nu$ the Lebesgue integral defining Z_Γ is almost surely finite and $\mathbb{E}[|Z_\Gamma|] < \infty$.

(b) If $\mathcal{F}_\nu(M)$ includes the identically one function, then

(i) M_Γ is almost surely finite and $\mathbb{E}[|M_\Gamma|] < \infty$ for every $\Gamma \in \mathcal{M}_I^\nu$,

(ii) all bounded non-decreasing functions $f : I \to \mathbb{R}$ are in $\mathcal{F}_\nu(M)$, especially all constant functions.

Proof. (a) For every $f \in \mathcal{F}_\nu(M)$ we have that $|f(t)M_t(\omega)|$ is $(\mathbb{P} \otimes \Gamma)$-integrable for all $\Gamma \in \mathcal{M}_I^\nu$ such that

$$\begin{aligned} \infty > \mathbb{E}[|f(t)M_t(\omega)|] &= \int_{\Omega \times I} |f(t)M_t(\omega)| \, d(\mathbb{P} \otimes \Gamma)(\omega, t) \\ &\overset{\text{Theorem A.5.18}}{=} \int_\Omega \bigg(\int_I |Z_t(\omega)| \Gamma(\omega, dt) \bigg) \mathbb{P}(d\omega) \geq \int_\Omega \bigg| \underbrace{\int_I Z_t(\omega) \Gamma(\omega, dt)}_{= Z_\Gamma} \bigg| \mathbb{P}(d\omega) \\ &= \mathbb{E}[|Z_\Gamma|]. \end{aligned}$$

This implies that Z_Γ is almost surely finite and $\mathbb{E}[|Z_\Gamma|] < \infty$.

(b) (i) If $\mathcal{F}_\nu(M)$ includes the identically one function, we have that $|M_t(\omega)|$ is $(\mathbb{P}\otimes\Gamma)$-integrable for all $\Gamma \in \mathcal{M}_I^\nu$ such that M_Γ is almost surely finite and $\mathbb{E}[|M_\Gamma|] < \infty$, cf. the thoughts above.

(ii) $\mathcal{F}_\nu(M)$ includes the identically one function which means that $|M_t(\omega)|$ is $(\mathbb{P}\otimes\Gamma)$-integrable for all $\Gamma \in \mathcal{M}_I^\nu$. Moreover, for every bounded non-decreasing functions $f : I \to \mathbb{R}$ it holds that $\sup_{t\in I} |f(t)| < \infty$. Using these statements we get that $|f(t)M_t(\omega)| \le \sup_{t\in I} |f(t)||M_t(\omega)|$ is $(\mathbb{P}\otimes\Gamma)$-integrable for all $\Gamma \in \mathcal{M}_I^\nu$. Thus $f \in \mathcal{F}_\nu(M)$. Especially for every constant function $f \equiv c$, $c \in \mathbb{R}$ we have that the equality $|f(t)M_t(\omega)| = |c||M_t(\omega)|$. ❑

Remark 6.2.3. Note that:

(a) If $\mathcal{F}_\nu(M)$ includes the identically one function, then $M \in L^1(\mathbb{P}\otimes\Gamma)$. $M \in L^1(\mathbb{P}\otimes\Gamma)$ is equivalent to the condition $\mathbb{E}[\int_I |M_t|\Gamma(\cdot,dt)] < \infty$, which implies that $\int_I |M_t|\Gamma(\cdot,dt) < \infty$, $\mathbb{P}$-a.s.

(b) For every $f \in \mathcal{F}_\nu(M)$ the process Z given by (6.2.1) is in $L^1(\mathbb{P}\otimes\Gamma)$.

Definition 6.2.4. Fix $\Gamma \in \mathcal{M}_I^\nu$. For every progressively measurable martingale M with $M \in L^1(\mathbb{P}\otimes\Gamma)$, we define $\Gamma_M : \Omega \times \mathcal{B}_I \to \mathbb{R}$ by

$$\Gamma_M(\omega,B) := \begin{cases} \int_B M_t(\omega)\Gamma(\omega,dt), & \text{if } \int_I |M_t(\omega)|\Gamma(\omega,dt) < \infty, \\ 0, & \text{otherwise} \end{cases}$$

for $\omega \in \Omega$, $B \in \mathcal{B}_I$.

Lemma 6.2.5. Γ_M *given as in Definition 6.2.4 is a (signed) transition kernel.*

Proof. We have to show that Γ_M satisfies the condition of Definition A.5.10.

(a) At first, we have to show that $\omega \mapsto \Gamma_M(\omega, I_{\le t})$ is $\mathcal{F}_t$-measurable. We decompose M into its positive part $M_t^+ := \max\{M_t, 0\}$ and its negative part $M_t^- := \max\{-M_t, 0\}$, such that $M_t := M_t^+ - M_t^-$ for every $t \in I$. Due to the progressively measurability of M we know that the restriction of $M^\pm$ to $I_{\le t} \times \Omega$ is $(\mathcal{B}_{I_{\le t}} \otimes \mathcal{F}_t)$-measurable for every $t \in I$. Using Lemma A.5.11 for $f(\omega,t) := M_t^\pm(\omega)$ it follows that

$$\omega \mapsto \int_{I_{\le t}} M_s^\pm(\omega)\Gamma(\omega,ds) \tag{6.2.6}$$

is $\mathcal{F}_t$-measurable and well-defined.

(b) Furthermore we have to show that $B \mapsto \Gamma_M(\omega,B)$ is a signed measure on $(I,\mathcal{B}_I)$ for every $\omega \in \Omega$. Combine the knowledge that $B \mapsto \Gamma(\omega,B)$ is a probability measure on $(I,\mathcal{B}_I)$ for every $\omega \in \Omega$ with $\mathbb{E}[\int |M_t|\Gamma(\cdot,dt)] < \infty$, see Remark 6.2.3, implies that $B \mapsto \int_B M_t(\omega)\Gamma(\omega,dt)$ is a signed measure. ❑

Remark 6.2.7. Using Corollary A.5.14 we have that $\mathbb{P} \otimes \Gamma_M$ is a signed measure on $(\Omega \times I, \mathcal{F} \otimes \mathcal{B}_I)$ and uniquely determined by

$$\mathbb{P} \otimes \Gamma_M(A_1 \times A_2) = \int_{A_1} \Gamma_M(\omega, A_2)\mathbb{P}(d\omega) \quad \text{for all } A_1 \in \mathcal{F}, A_2 \in \mathcal{B}_I.$$

Due to Γ_M depends on M and Γ, it is obvious that $\mathbb{P} \otimes \Gamma_M$ is absolutely continuous with respect to $\mathbb{P} \otimes \Gamma$. Under the assumption that a measure λ is σ-finite, the Radon-Nikodym theorem characterizes the absolute continuity of η with respect to λ with the existence of a function $f \in L^1(\lambda)$ such that $\eta = f\lambda$, i.e., such that

$$\eta(A) = \int_A f \, d\lambda \quad \text{for every } A \in \mathcal{B}_I.$$

Some authors use the name "Differentiation of measures" for the decomposition above and the density f is sometimes denoted by $\frac{d\eta}{d\lambda}$ which we will use. The absolute continuity of η with respect to λ is denoted by $\eta \ll \lambda$.

Furthermore, using the definitions given above we have for all $A_1 \in \mathcal{F}$ and $A_2 \in \mathcal{B}_I$ that

$$\begin{aligned}(\mathbb{P} \otimes \Gamma_M)(A_1, A_2) &= \int_{A_1} \Gamma_M(\omega, A_2)\,\mathbb{P}(d\omega) = \int_{A_1}\int_{A_2} M_t(\omega)\,\Gamma(\omega, dt)\mathbb{P}(d\omega) \\ &= \int_{A_1 \times A_2} M_t(\omega)\, d(\mathbb{P} \otimes \Gamma)(\omega, t)\end{aligned}$$

such that

$$\frac{d(\mathbb{P} \otimes \Gamma_M)}{d(\mathbb{P} \otimes \Gamma)}(\omega, t) = M_t(\omega).$$

Remark 6.2.8. In addition, we define

$$\mu_{\Gamma,M}(B) := \mathbb{E}[\Gamma_M(\cdot, B)] = \mathbb{E}\left[\int_B M_t \Gamma(\cdot, dt)\right] \quad \text{for all } B \in \mathcal{B}_I.$$

Because $\mathcal{F}_\nu(M)$ contains the identically one function, we get that M_Γ is well-defined and $\mathbb{P}\otimes\Gamma$-integrable, cf. Lemma 6.2.2. Remembering that $\mathbb{E}[\int |M_t|\Gamma(\cdot, dt)] < \infty$, if $M \in L^1(\mathbb{P}\otimes\Gamma)$. Then $\mu_{\Gamma,M}$ is a signed measure and a projection from $\mathbb{P} \otimes \Gamma_M$ to I. In a nutshell, we will consider the probability measure μ_Γ and the signed measure $\mu_{\Gamma,M}$ which are given by

$$\mu_\Gamma(B) = \mathbb{E}[\Gamma(\cdot, B)] = \int_\Omega \Gamma(\cdot, B)\, d\mathbb{P} = \nu(B) \quad \text{and} \tag{6.2.9}$$

$$\mu_{\Gamma,M}(B) = \mathbb{E}[\Gamma_M(\cdot, B)] = \int_\Omega \int_B M_t \Gamma(\cdot, dt)\, d\mathbb{P} \quad \text{for all } B \in \mathcal{B}_I \tag{6.2.10}$$

and we know that

$$\text{proj}_I(\mathbb{P} \otimes \Gamma) = \mu_\Gamma = \nu \quad \text{and} \quad \text{proj}_I(\mathbb{P} \otimes \Gamma_M) = \mu_{\Gamma,M}$$

and $\mu_{\Gamma,M}$ is absolutely continuous with respect to $\mu_\Gamma = \nu$, i.e., if $\nu(A) = 0$ for any $A \in \mathcal{B}_I$, then $\mu_{\Gamma,M}(A) = 0$.

Now, we will formulate the main theorem of this subsection.

Theorem 6.2.11. *Given a continuous time interval $I \subseteq \mathbb{R}$, a probability distribution ν on I and a right-continuous martingale M. Assume that $\mathcal{F}_\nu(M)$ contains the identically one function. Then for a stochastic transition kernel $\Gamma^* \in \mathcal{M}_I^\nu$ the following properties are equivalent:*

(a) *Γ^* is optimal for all processes $(Z_t)_{t\in I}$ given as*

$$Z_t = f(t)M_t, \quad t \in I, \tag{6.2.12}$$

with $f \in \mathcal{F}_\nu(M)$ and f is bounded.

(b) *Γ^* satisfies $\mathbb{E}[M_{\Gamma^*}] = \mathbb{E}[M_\Gamma]$ and*

$$\mathbb{E}\left[\int_{I_{>s}} M_t \Gamma^*(\cdot, dt)\right] \geq \mathbb{E}\left[\int_{I_{>s}} M_t \Gamma(\cdot, dt)\right] \tag{6.2.13}$$

for all $s \in I$ and all $\Gamma \in \mathcal{M}_I^\nu$.

(c) *Γ^* is optimal for all processes $(Z_t)_{t\in I}$ given as*

$$Z_t = f(t)M_t, \quad t \in I,$$

with $f \in \mathcal{F}_\nu(M)$.

Proof of Theorem 6.2.11.

1. (a) implies (b):
   Note that based on the claimed conditions we have that Z_Γ and M_Γ are well-defined and in $L^1(\mathbb{P} \otimes \Gamma)$ for all $\Gamma \in \mathcal{M}_I^\nu$, see Lemma 6.2.2. A stochastic transition kernel $\Gamma^* \in \mathcal{M}_I^\nu$ is optimal for $(Z_t)_{t\in I}$, if $\mathbb{E}[Z_{\Gamma^*}] \geq \mathbb{E}[Z_\Gamma]$ for all $\Gamma \in \mathcal{M}_I^\nu$. For every $(Z_t)_{t\in I}$ given in the form as in condition (a) of Theorem 6.2.11, the optimality of Γ^* implies

   $$\mathbb{E}[Z_{\Gamma^*}] = \mathbb{E}\left[\int_I f(t)M_t \Gamma^*(\cdot, dt)\right] \geq \mathbb{E}[Z_\Gamma] = \mathbb{E}\left[\int_I f(t)M_t \Gamma(\cdot, dt)\right], \quad \Gamma \in \mathcal{M}_I^\nu. \tag{6.2.14}$$

   For a fixed $s \in I$ the function $f_s : I \to \mathbb{R}$ defined as

   $$f_s(t) := \mathbb{1}_{I_{>s}}(t), \quad t \in I, \tag{6.2.15}$$

   is a special non-decreasing deterministic function and bounded by 1. We have that

   $$f_s(t)M_t = \begin{cases} 0 & \text{for } t \in I_{\leq s}, \\ M_t & \text{for } t \in I_{>s}, \end{cases}$$

   and $\int_I |f_s(t)M_t|\Gamma(\omega, dt) = \int_{I_{>s}} |M_t|\Gamma(\omega, dt) \leq \int_I |M_t|\Gamma(\omega, dt)$ for almost all $\omega \in \Omega$. Due to the identically one function being in $\mathcal{F}_\nu(M)$, we get that $f_s \in \mathcal{F}_\nu(M)$. The inequality (6.2.14) holds for every non-decreasing deterministic function $f \in \mathcal{F}_\nu(M)$, particularly for f_s, $s \in I$. For $s \in I$ and for every $\Gamma \in \mathcal{M}_I^\nu$ we have

   $$\mathbb{E}\left[\int_I f_s(t)M_t \Gamma(\cdot, dt)\right] = \mathbb{E}\left[\int_{I_{>s}} M_t \Gamma(\cdot, dt)\right], \tag{6.2.16}$$

so that we get for every $s \in I$ and every $\Gamma \in \mathcal{M}_I^\nu$ with the special choice f_s that

$$\mathbb{E}\left[\int_{I_{>s}} M_t \Gamma^*(\cdot, dt)\right] \geq \mathbb{E}\left[\int_{I_{>s}} M_t \Gamma(\cdot, dt)\right].$$

Now, we want show that $\mathbb{E}[M_\Gamma]$ is the same real number for all $\Gamma \in \mathcal{M}_I^\nu$. Applying the inequation (6.2.14) for the identically one function, we get immediately that

$$\mathbb{E}[M_{\Gamma^*}] \geq \mathbb{E}[M_\Gamma] \quad \text{for all } \Gamma \in \mathcal{M}_I^\nu. \tag{6.2.17}$$

If $\mathcal{F}_\nu(M)$ includes the identically one function, then we know that all constant functions are in $\mathcal{F}_\nu(M)$, see Lemma 6.2.2. Therefore we can also apply the inequality (6.2.14) for the function which is identically minus one, i.e., $g(t) = -1$, for all $t \in I$. Note that g is bounded by one. We get for every $\Gamma \in \mathcal{M}_I^\nu$ that

$$\mathbb{E}[M_{\Gamma^*}] \leq \mathbb{E}[M_\Gamma] \quad \text{for all } \Gamma \in \mathcal{M}_I^\nu. \tag{6.2.18}$$

Putting the inequations (6.2.17) and (6.2.18) together we get the assertion.

2. (b) implies (a) and (b) implies (c):
Using Remark 6.2.8 and Definition 6.2.4, we define for all $B \in \mathcal{B}_I$ and for every $\Gamma \in \mathcal{M}_I^\nu$

$$\mu_{\Gamma,M}(B) := \mathbb{E}[\Gamma_M(\cdot, B)] = \mathbb{E}\left[\int_B M_t \Gamma(\cdot, dt)\right].$$

- Because $\mathcal{F}_\nu(M)$ contains the identically one function, we get that M_Γ is well-defined and $\mathbb{P} \otimes \Gamma$-integrable, cf. Lemma 6.2.2.
Remember that $\mathbb{E}[\int |M_t| \Gamma(\cdot, dt)] < \infty$, if $M \in L^1(\mathbb{P} \otimes \Gamma)$. Then $\mu_{\Gamma,M}$ is a signed measure with finite total variation, because of

$$|\mu_{\Gamma,M}|(I) = \mathbb{E}\left[\int_I |M_t| \Gamma(\cdot, dt)\right] < \infty.$$

- Due to $\mathbb{E}[M_\Gamma]$ is the same real number for all $\Gamma \in \mathcal{M}_I^\nu$, we have for every $\Gamma, \tilde{\Gamma} \in \mathcal{M}_I^\nu$ that $\mu_{\Gamma,M}(I) = \mathbb{E}[\int_I M_t \Gamma(\cdot, dt)] = \mathbb{E}[M_\Gamma] = \mathbb{E}[M_{\tilde{\Gamma}}] = \mu_{\tilde{\Gamma},M}(I)$.

Now, we assume that $\Gamma^* \in \mathcal{M}_I^\nu$ is optimal and satisfies inequation (6.2.13), i.e.,

$$\mathbb{E}\left[\int_{I_{>s}} M_t \Gamma^*(\cdot, dt)\right] \geq \mathbb{E}\left[\int_{I_{>s}} M_t \Gamma(\cdot, dt)\right]$$

for all $s \in I$ and $\Gamma \in \mathcal{M}_I^\nu$. Using equation (6.2.10) we have

$$\mu_{\Gamma^*,M}(I_{>s}) \geq \mu_{\Gamma,M}(I_{>s}) \quad \text{for all } s \in I,$$

this means $\mu_{\Gamma^*,M}$ dominates $\mu_{\Gamma,M}$ in first order, cf. Definition A.1.11. In addition, due to Lemma A.1.18 we have equivalently

$$\int_I f(t)\, d\mu_{\Gamma^*,M}(t) \geq \int_I f(t)\, d\mu_{\Gamma,M}(t), \tag{6.2.19}$$

for all functions f such that the integrals exist. Because of the claimed conditions of Theorem 6.2.11 we know that the integrals in equation (6.2.19) exist for <u>all</u> $f \in \mathcal{F}_\nu(M)$ and Z_Γ is almost surely finite and $\mathbb{E}[|Z_\Gamma|] < \infty$, see Lemma 6.2.2. Therefore by using Remark 6.2.7 we have that

$$\mathbb{E}[Z_\Gamma] = \int_I f(t)\, d\mu_{\Gamma,M}(t), \tag{6.2.20}$$

because

$$\begin{aligned}
\mathbb{E}[Z_\Gamma] &= \int_{\Omega\times I} Z_t(\omega)\, d(\mathbb{P}\otimes\Gamma)(\omega,t) = \int_{\Omega\times I} f(t) M_t(\omega)\, d(\mathbb{P}\otimes\Gamma)(\omega,t) \\
&= \int_{\Omega\times I} f(t) \frac{d(\mathbb{P}\otimes\Gamma_M)}{d(\mathbb{P}\otimes\Gamma)}(\omega,t)\, d(\mathbb{P}\otimes\Gamma)(\omega,t) = \int_{\Omega\times I} f(t)\, d(\mathbb{P}\otimes\Gamma_M)(\omega,t) \\
&= \int_I f(t) \int_\Omega d(\mathbb{P}\otimes\Gamma_M)(\omega,t) = \int_I f(t)\, d\mu_{\Gamma,M}(t).
\end{aligned}$$

Finally, with equation (6.2.20) and inequation (6.2.19) it follows for every $(Z_t)_{t\in I}$ given in the form as in condition (a) of Lemma 6.2.11 that

$$\mathbb{E}[Z_{\Gamma^*}] \geq \mathbb{E}[Z_\Gamma] \quad \text{for all } \Gamma \in \mathcal{M}_I^\nu$$

such that we get the assertion.

3. (c) implies (b):
   It follows immediately using (a) implies (b). ❑

7

Randomized Stopping Times

In this chapter we consider the distribution-constrained stopping problems from a mass transport perspective. For this we reformulate the problem OptStop$^\tau$ in terms of optimal transport into OptStop$^\tau$. Using the methods and techniques of optimal transport theory we obtain the existence of optimal stopping times of a Brownian motion with given marginals. However, the cost process must be at least measurable and appropriately adapted. Certain continuity assurances then guarantee the existence of solutions to the considered problem. The main result is Theorem 7.1.10. To prove it, we have to introduce randomized stopping times and the problem OptStop$^\pi$. Furthermore, ideas and concepts from the optimal transport (and its martingale variant) are adapted to obtain a geometric description of the optimal strategy. A fundamental idea in optimal transport is that the optimality of a transport plan is reflected by the geometry of its support set which can be characterized using the notion of *c-cyclical monotonicity*. The relevance of this concept for the theory of optimal transport has been fully recognized by Gangbo and McCann [29], based on earlier work of Knott and Smith [42] and Rüschendorf [67, 68] among others. Inspired by these ideas, the literature on martingale optimal transport has developed a 'monotonicity principle' which allows to characterize martingale transport plans through geometric properties of their support sets, cf. [14, 76, 11, 9, 32, 15]. This martingale optimal transport problem arises naturally in robust finance; papers to investigate such problems include [35, 12, 28, 22, 17, 30, 55], and this topic is commonly referred to as martingale optimal transport. In mathematical finance, transport techniques complement the Skorokhod embedding approach (see [56, 34] for an overview) to model-independent/robust finance.
The methods work for a large class of cost processes and it is shown that for many cost processes a solution is given by the first hitting time of a barrier in a suitable phase space. As a by-product we recover classical solutions of the inverse first passage time problem / Shiryaev's problem. The results of this section of the thesis are already in [10] published.

In this chapter, let $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space and $(B_t)_{t\geq 0}$ be a Brownian motion started[1] in 0 on some filtered probability space $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions and let ν be a measure on $(0,\infty)$. The main contribution of this chapter is to establish a monotonicity principle which is applicable to distribution-constrained

[1] We note that the results presented in this section remain valid for Brownian motions started according to a general law λ at the cost of slightly more tedious moment conditions in the formulation of Corollaries 7.4.1 and 7.4.9.

optimal stopping problems OptStop$^{\tau}$. The transport approach turns out to be remarkably powerful, in particular we will find that questions as raised in Problems OptStop$^{\psi(B_t,t)}$ and OptStop$^{B_t^*}$ can be addressed using a relatively intuitive set of arguments. We emphasize that the solutions to the constrained optimal stopping problems provided in Corollaries 7.4.1 and 7.4.9 represent particular applications of the abstract results obtained below. Figure 7.1 presents graphical depictions of stopping rules of several further solutions of constrained optimal stopping problems (together with the respective optimality properties). These stopping rules can be derived – under suitable moment conditions – using arguments very similar to those required for Corollaries 7.4.1 and 7.4.9.

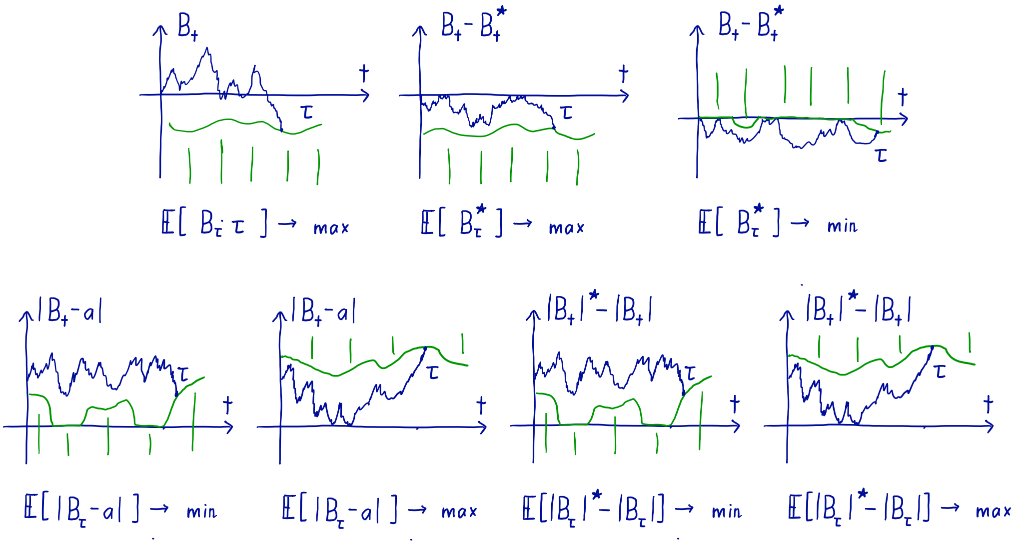


Figure 7.1.: Solutions to constrained optimal stopping problems.

The remainder of this chapter is organized as follows: in Section 7.1 we reformulate our considered problem in terms of optimal transport and we specify the necessary assumptions, notations and definitions. Furthermore the main results are given in Theorem 7.1.10 and Theorem 7.1.18. The existence of an optimizer of OptStop$^{\tau}$ is shown in Section 7.2. For this we introduce randomized stopping times and the corresponding problem OptStop$^{\pi}$. We show that Problem OptStop$^{\pi}$ has a solution and therefore OptStop$^{\tau}$. Section 7.3 deals with the monotonicity principle and is devoted to the proof of Theorem 7.1.18. Finally, in Section 7.4 examples are investigated and optimal maximizers are determined.

7.1. The Problem and Main Results

Assumption 7.1.1. Let $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space and $(B_t)_{t\geq 0}$ be an adapted process which has continuous paths on $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\geq 0})$, such that B can be regarded as a measurable map from Ω to $C(\mathbb{R}_+)$, the space of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}$. The cost function c will always be a measurable map $C(\mathbb{R}_+) \times \mathbb{R}_+ \to \mathbb{R}$. Our given probability measure on $\mathbb{R}_+$ will be denoted by ν.

Here we formulate our main optimal stopping problem in terms of minimization, following the usual convention in the optimal transport literature (which is also used in the closely related paper [9]). Clearly, a sign change transforms this into a maximization problem and in our applications we will in fact turn to this latter version when resulting formulations appear more natural, similar to Remark 3.1.11. We trust that this will not cause confusion. Then the problem we consider can be stated as follows.

Problem (OptStop$^\tau$). Among all stopping times $\tau \sim \nu$ find the minimizer of

$$\tau \mapsto \mathbb{E}[c(B, \tau)].$$

Remark 7.1.2. Bayraktar and Miller [8] consider the same optimization problem that we treat here. However their setup and methods are rather distinct from the ones used here: they assume that the target distribution is given by finitely many atoms and that the target functional depends solely on the terminal value of Brownian motion. Following the measure valued martingale approach of Cox and Källblad [20], [8] address the constrained optimal stopping problem using a Bellman perspective.

Remark 7.1.3. The problem to construct a stopping time τ of Brownian motion such that the law of τ matches a given distribution on the real line was proposed by Shiryaev in his Banach Center lectures in the 1970's, it has since been called Shiryaev's problem or inverse first passage problem. Dudley and Gutmann [24] provide an abstract measure-theoretic construction. An early barrier-type solution to the inverse first passage problem was given by Anulova [4]. She constructs a symmetric two-sided barrier (corresponding to the case $a = 0$ in the sixth picture of Figure 7.1). Anulova discretises the measure μ and concludes through approximation arguments. The solution to the inverse first passage problem given in Corollary 7.4.1 was derived by Chen, Cheng, and Chadam, and Saunders [18] based on a variational inequality which describes the corresponding barrier. Notably, this is predated by a (formal) PDE description of such barriers given by Avellaneda and Zhu [5] in the context of credit risk modeling. Ekström and Janson [27] relate this solution to an optimal stopping problem and provide an integral equation for the barrier. Analytic solutions to the inverse first passage problem are known only in a few cases ([16, 45, 58, 69, 1, 2]). An interesting connection between the inverse first passage problem and Skorokhod's problem is provided by Jaimungal, Kreinin, and Valov [36].

Throughout we will also make the following assumptions without further mention:

Assumption 7.1.4.

(a) Let c be measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted, where $(\mathcal{F}_t^0)_{t\geq 0}$ is the filtration on $C(\mathbb{R}_+)$ generated by the canonical process $(\omega \mapsto \omega(t))_{t\in\mathbb{R}_+}$.

(b) There is a $\mathbb{G}_0$-measurable random variable U which is uniformly distributed on $[0,1]$ and independent of the process $(B_t)_{t\geq 0}$.

(c) There is a probability measure λ such that $(B_t)_{t\geq 0}$ is a Brownian motion with initial law λ, i.e., $B_0 \sim \lambda$.

(d) The problem is well-posed in the sense that $\mathbb{E}[c(B,\tau)]$ is defined and $> -\infty$ for all stopping times $\tau \sim \nu$ and that $\mathbb{E}[c(B,\tau)] < \infty$ for at least one such stopping time.

(e) $\int t^{p_0}\, d\nu(t) < \infty$, where $p_0 \geq 0$ is some constant that we fix here and that can be chosen when applying the results from this section.

Remark 7.1.5. A note on language: The adjective "adapted" is usually applied to processes whose time argument is written in subscript form. For any filtered measurable space $\tilde{\Omega}$ and any function $f : \tilde{\Omega}\times\mathbb{R}_+ \to \mathbb{R}$ (or possibly $f : \tilde{\Omega}\times\mathbb{R}_+ \to [-\infty,\infty]$) we will interchangeably think of f simply as a function or as the process $Y_t(\omega) := f(\omega,t)$. And so f being adapted means the same thing as $(Y_t)_{t\in\mathbb{R}_+}$ being adapted. Similarly for a subset Γ of $\tilde{\Omega}\times\mathbb{R}_+$ we may also think of Γ as its indicator function or as the process $Y_t(\omega) := \mathbb{1}_\Gamma(\omega,t)$ and will also say that the set Γ is adapted. Note that Γ is the common notation in optimal transport and does not stand in connection with the transition kernels from the previous section. We trust this will not cause confusion from now on.

Remark 7.1.6. With that in mind, Assumption 7.1.4.(a) should seem like an obvious thing to ask for from the cost function. Also, knowing about the existence of optional projections, it should be clear no later than (7.2.4) that Assumption 7.1.4.(a) does not pose a real restriction on the class of problems we are treating.

Remark 7.1.7. The role of Assumption 7.1.4.(b) should become clearer soon. We would like to note at this point though that often enough our results, put together, will imply that the solution of Problem OptStop$^\tau$ for a space $(\Omega,\mathcal{G},(\mathcal{G}_t)_{t\geq 0},\mathbb{P})$ which satisfies Assumption 7.1.4.(b) is essentially the same as the solution of the problem for a space which may not satisfy said assumption, and we will find that we can describe this solution in detail. This can be seen executed in the proofs of the corollaries stated in the Section 7.4.

Remark 7.1.8. The methods in this chapter work not just for Brownian motion but for a class of processes which is conceptually bigger, but then turns out to not include much beyond Brownian motion – namely for any space-homogeneous but possibly time-inhomogeneous Markov process with continuous paths which has the strong Markov property. (Here space-homogeneous means that starting the process at location x and then moving its paths to start at location y results in a version of the process started at y.) If the reader so wishes, you may think of B as a process from this slightly larger class of processes. Care was taken not to reference any properties of Brownian motion beyond those stated here. In particular, our results apply to multi-dimensional Brownian motion.

Remark 7.1.9. The constant p_0 in Assumption 7.1.4.(e) will (implicitly) appear in the statement of Theorem 7.1.18, one of the main results. It is role is to ensure that $\mathbb{E}[\varphi(B,\tau)]$ will be finite for some (class of) function(s) φ and any solution τ of OptStop$^\tau$. (The choice $\varphi(B,\tau) = \tau^{p_0}$ is somewhat arbitrary here.)

The main results are Theorem 7.1.10 and Theorem 7.1.18. We give two versions of Theorem 7.1.10. Version A is easier to state and may feel more natural, but we will need Version B (which is more general and has essentially the same proof as Version A) in the proof of the corollaries in the Section 7.4.

Theorem 7.1.10.

Version A. *Assume that the cost function c is bounded from below and lower semicontinuous when we equip $C(\mathbb{R}_+)$ with the topology of uniform convergence on compacts. Then the Problem* OptStop$^\tau$ *has a solution.*

Version B. *Assume that the cost function c is lower semicontinuous when we equip $C(\mathbb{R}_+) \times \mathbb{R}_+$ with the product topology of two Polish topologies which generate the right sigma-algebras on $C(\mathbb{R}_+)$ and $\mathbb{R}_+$ respectively and that the set $\{c_-(B,\tau) : \tau \sim \nu,\ \tau$ is a stopping time$\}$ is uniformly integrable, where $c_- := -c \vee 0$ denotes the negative part of c. Then the Problem* OptStop$^\tau$ *has a solution.*

To state Theorem 7.1.18 we need a few more definitions.

Remark 7.1.11. We will find it convenient to talk about processes that do not start at time 0 but instead at some time $t > 0$. Similarly we will consider stopping times taking values in $[t,\infty)$. These will be defined on the space $C([t,\infty))$ equipped with the filtration $(\mathcal{F}_s^t)_{s\geq t}$, generated by the canonical process $(\omega \mapsto \omega(s))_{s\geq t}$ again. We refer to the distribution of Brownian motion started at time t and location x by $\mathbb{W}_x^t$. This is a measure on $C([t,\infty))$. For a probability measure κ on $\mathbb{R}$ we write $\mathbb{W}_\kappa^t$ for the distribution of Brownian motion started at time t with initial law κ.

Definition 7.1.12 (Concatenation)**.**
For every $t \in \mathbb{R}_+$ we have an operation $\odot$ of concatenation, which is a map into $C([t,\infty))$ and is defined for $(\omega,s) \in C([t,\infty)) \times [t,\infty)$ and $\theta \in C([s,\infty))$ with $\theta(s) = 0$ by

$$((\omega,s)\odot\theta)(r) = \begin{cases} \omega(r) & t \leq r \leq s \\ \omega(s) + \theta(r) & r > s \end{cases} . \tag{7.1.13}$$

Definition 7.1.14 (Stop-Go pairs)**.**
The set of Stop-Go pairs $\mathrm{SG} \subseteq (C(\mathbb{R}_+)\times\mathbb{R}_+)\times(C(\mathbb{R}_+)\times\mathbb{R}_+)$ is defined as the set of all pairs $((\omega,t),(\eta,t))$ (note that the time components have to match) such that

$$c(\omega,t) + \int c((\eta,t)\odot\theta,\sigma(\theta))\,d\mathbb{W}_0^t(\theta) < c(\eta,t) + \int c((\omega,t)\odot\theta,\sigma(\theta))\,d\mathbb{W}_0^t(\theta) \tag{7.1.15}$$

for all $(\mathcal{F}_s^t)_{s\geq t}$-stopping times σ for which $\mathbb{W}_0^t(\sigma = t) < 1$, $\mathbb{W}_0^t(\sigma = \infty) = 0$, $\int \sigma^{p_0}\,d\mathbb{W}_0^t < \infty$ and for which both sides in (7.1.15) are defined and finite.

A hopefully intuitive way of putting the definition of Stop-Go pairs into words is the following: $((\omega,s),(\eta,t))$ form a Stop-Go pair if and only if, irrespective of how we might stop after time t (i.e., which stopping rule σ we might use after time t), stopping ω at time t and letting η go on is better – i.e., has lower cost – than stopping η and letting ω go on. As hinted at earlier, the definition of Stop-Go pairs depends on the parameter p_0 from Assumption 7.1.4.(e). A larger p_0 means that we are asking for more in Assumption 7.1.4.(e) and implies that we get a larger set SG, as we are quantifying over fewer stopping times σ in the definition of SG. This in turn implies that the conclusion of Theorem 7.1.18 below will be stronger.

Definition 7.1.16 (Initial Segments)**.**
For a set $\Gamma \subseteq C(\mathbb{R}_+)\times\mathbb{R}_+$ define the set $\Gamma^< \subseteq C(\mathbb{R}_+)\times\mathbb{R}_+$ by

$$\Gamma^< = \{(\omega,s) : (\omega,t) \in \Gamma \text{ for some } t > s\} . \tag{7.1.17}$$

Theorem 7.1.18 (Monotonicity Principle)**.**
Assume that τ solves OptStop$^\tau$*. Then there is a measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted set $\Gamma \subseteq C(\mathbb{R}_+)\times\mathbb{R}_+$ such that*

$$\mathbb{P}[((B_t)_{t\geq 0}, \tau) \in \Gamma] = 1$$

and

$$\mathrm{SG} \cap \left(\Gamma^< \times \Gamma\right) = \emptyset\,. \tag{7.1.19}$$

The following lemma should give a first hint about how the Monotonicity Principle can be applied.

Lemma 7.1.20. *Let τ be a solution of* OptStop$^\tau$ *and assume that the cost function c is such that there exists a measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted process $(Y_t)_{t\geq 0}$ such that*

$$Y_t(\omega) < Y_t(\eta) \implies ((\omega, t), (\eta, t)) \in \mathrm{SG}\,. \tag{7.1.21}$$

Define the barriers $\check{\mathcal{R}}, \hat{\mathcal{R}} \subseteq \mathbb{R}\times\mathbb{R}_+$ by

$$\check{\mathcal{R}} = \bigcup_{(\omega,t)\in\Gamma} (-\infty, Y_t(\omega)] \times \{t\}\,,$$
$$\hat{\mathcal{R}} = \bigcup_{(\omega,t)\in\Gamma} (-\infty, Y_t(\omega)) \times \{t\}\,,$$

where Γ is a set with the properties in Theorem 7.1.18. Define the functions $\check{\tau}$ and $\hat{\tau}$ on $C(\mathbb{R}_+)$ by

$$\check{\tau}(\tilde{\omega}) = \inf\left\{t \in \mathbb{R}_+ : (Y_t(B(\tilde{\omega})), t) \in \check{\mathcal{R}}\right\}\,,$$
$$\hat{\tau}(\tilde{\omega}) = \inf\left\{t \in \mathbb{R}_+ : (Y_t(B(\tilde{\omega})), t) \in \hat{\mathcal{R}}\right\}.$$

Then

$$\check{\tau} \leq \tau \leq \hat{\tau}\,, \quad \mathbb{P}\text{-}a.s. \tag{7.1.22}$$

When applying this Lemma to show that some optimal stopping problem has a barrier-type solution as symbolized for example by the pictures in Figure 7.1 the process $Y_t(B)$ is of course with what we are labelling the vertical axes in the pictures. So for the first picture $Y_t(\omega) = \omega(t)$, for the second one $Y_t(\omega) = \omega(t) - \sup_{s\leq t}\omega(s)$, for the third $Y_t(\omega) = -(\omega(t) - \sup_{s\leq t}\omega(s))$ (the sign is flipped relative to the labelling in the picture because in this picture the barrier is drawn "up" instead of "down"), etc.
Notice that, contrary to customs, when we draw the barriers $\check{\mathcal{R}} \setminus \hat{\mathcal{R}}$ in the pictures coordinate is the in Figure 7.1 the first coordinate is the vertical axis and the second horizontal axis. This is to make cross-referencing and comparison with [9] easier, we follow their convention of always having time as the second coordinate but still in the pictures it seems more natural to put the independent variable on the horizontal axis.
Note that a priori $\check{\tau}$ and $\hat{\tau}$ do not need to be stopping times or even measurable, as we do not know much about the sets $\check{\mathcal{R}}$ and $\hat{\mathcal{R}}$.
Using the properties of a concrete process $(Y_t)_{t\geq 0}$ we will be able to show in the proofs of Corollaries 7.4.1 and 7.4.9 that $\check{\tau} = \hat{\tau}$ a.s. (this should not be surprising as for each time t the barriers $\check{\mathcal{R}}$ and $\hat{\mathcal{R}}$ differ by at most a single point) and therefore that the optimizer τ is the hitting time of a barrier.

Proof of Lemma 7.1.20. Let $\tilde{\omega} \in \Omega$ s.t. $(B(\tilde{\omega}), \tau(\tilde{\omega})) \in \Gamma$. By assumption this holds for $\mathbb{P}$-a.a. $\tilde{\omega}$. Then $\left(Y_{\tau(\tilde{\omega})}(B(\tilde{\omega})), \tau(\tilde{\omega})\right) \in \underset{\vee}{\mathcal{R}}$ and therefore $\underset{\vee}{\tau}(\tilde{\omega}) \le \tau(\tilde{\omega})$.
Next, we show that $\hat{\tau}(\tilde{\omega}) \ge \tau(\tilde{\omega})$. Assume that $(Y_t(B(\tilde{\omega})), t) \in \hat{\mathcal{R}}$. We want to show that $t \ge \tau(\tilde{\omega})$. By the definition of $\hat{\mathcal{R}}$ we find that there is $\eta \in C(\mathbb{R}_+)$ with $(\eta, t) \in \Gamma$ and $Y_t(B(\tilde{\omega})) < Y_t(\eta)$, so by (7.1.21) we know $((B(\tilde{\omega}), t), (\eta, t)) \in \mathrm{SG}$. Assuming, if possible, $t < \tau(\tilde{\omega})$ we get according to Definition 7.1.16 that $(B(\tilde{\omega}), t) \in \Gamma^<$. Therefore we have that $((B(\tilde{\omega}), t), (\eta, t)) \in \mathrm{SG} \cap (\Gamma^< \times \Gamma)$, but this is a contradiction to $\mathrm{SG} \cap (\Gamma^< \times \Gamma) = \emptyset$, so we must have $t \ge \tau(\tilde{\omega})$. ❑

Remark 7.1.23 (Duality). Problem $\mathrm{OptStop}^\tau$ is an infinite-dimensional linear programming problem and hence, one would expect that a corresponding dual problem can be formulated. Indeed, assuming that c is lower semicontinuous and bounded from below, the value of the optimization problem equals

$$\sup_{M,\psi} \mathbb{E}[M_0] + \int \psi \, d\mu,$$

where the supremum is taken over bounded $\mathbb{G}$-martingales $M = (M_t)_{t \ge 0}$ and bounded continuous functions $\psi : \mathbb{R}_+ \to \mathbb{R}$ satisfying (up to evanescence)

$$M_t + \psi(t) \le c(B, t)\,.$$

This can be established in complete analogy to the duality result given in [9] and we do not elaborate.

Remark 7.1.24. The reader interested in the time-discrete case is referred to Part I. There we also use a different approach and view on the problem, and the existence of an optimal strategy is proven through ideas and concepts from functional analysis, see Chapter 3, and from the theory of optimal transport, see Chapter 4.

7.2. Existence of an Optimizer

The proof of existence of solutions to the problem $\mathrm{OptStop}^\tau$ crucially depends on thinking of stopping times as the joint distribution of the process to be stopped and the stopping time. We introduce some concepts to make this precise and give a proof of Theorem 7.1.10 at the end of this section.

Lemma 7.2.1. *Let $G : C([t,\infty)) \to \mathbb{R}$, and $s \ge t$. The function*

$$\omega \mapsto \int G((\omega, s) \odot \theta) \, d\mathbb{W}_0^s(\theta)$$

is a version of the conditional expectation $\mathbb{E}_{\mathbb{W}_\lambda^t}[G|\mathcal{F}_s^t]$ (for any initial distribution λ). Henceforth, by $\mathbb{E}[G|\mathcal{F}_s^t]$ we will mean this function.
If $G \in C_b(C([t,\infty)))$, then $\mathbb{E}[G|\mathcal{F}_s^t] \in C_b(C([t,\infty)))$.

Proof. Obvious. ❑

Here we use $C_b(X)$ to denote the set of continuous bounded functions from a topological space X to $\mathbb{R}$. The last sentence of the lemma is, of course, true for any topology on $C([t,\infty))$ for which the map $\omega \mapsto \omega \odot \theta$ is continuous for all θ, but we will only need it for the topology of uniform convergence on compacts.[2]
Given spaces X and Y we will denote the projection from $X \times Y$ to X by $\operatorname{proj} X$ (and similarly for Y). For a measurable map $F : X \to Y$ between measure spaces and a measure ν on X we denote the pushforward of ν under F by $F_*(\nu) := D \mapsto \nu(F^{-1}[D])$.

Definition 7.2.2 (RST)**.**
The set RST^t_κ of *randomized stopping times* (of Brownian motion started at time t with initial distribution κ) is defined as the set of all subprobability measures π on $C([t,\infty)) \times [t,\infty)$ such that $(\operatorname{proj} C([t,\infty)))_*(\pi) \leq \mathbb{W}^t_\kappa$ and that

$$\int F(r)\Big(G(\omega) - \mathbb{E}[G|\mathcal{F}^t_s](\omega)\Big)\, d\pi(\omega, r) = 0 \tag{7.2.3}$$

for all $s > t$, all continuous bounded $G : C([t,\infty)) \to \mathbb{R}$ and all continuous bounded $F : [t,\infty) \to \mathbb{R}$ supported on $[t,s]$.

In this definition the topology on $C([t,\infty))$ is that of uniform convergence on compacts and the topology on $[t,\infty)$ is the usual order-induced topology.
Given a distribution μ on $C([t,\infty))$ we write

$$\mathsf{RST}^t_\kappa(\mu) := \left\{\pi \in \mathsf{RST}^t_\kappa : (\operatorname{proj}_{[t,\infty)})_*(\pi) = \mu\right\} .$$

We write $\mathsf{RST}^t_\kappa(\mathcal{P})$ for the set of all $\pi \in \mathsf{RST}^t_\kappa$ with mass 1 and call these the *finite* randomized stopping times.
In any of these, if we drop the superscript t then we will mean time $t = 0$, while, if we drop the subscript κ, then we mean that the initial distribution $\kappa = \delta_0$, i.e., the Brownian motion to be stopped is started deterministically in 0.

To explain the qualifier finite it may help to imagine that for a non-finite randomized stopping time of mass $\alpha < 1$, the mass $1-\alpha$ which is missing is placed along $C([t,\infty)) \times \{\infty\}$.

The following (7.2.4) from [9] shows that the problem $\textsc{OptStop}^\tau$ is equivalent to the following optimization problem $\textsc{OptStop}^\pi$ in the sense that a solution of one can be translated into a solution of the other and vice versa. This of course also implies that the values of the two problems are equal, thereby showing that the concrete space $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\geq 0}, \mathbb{P})$ has no bearing on this value, as long as Assumptions 7.1.1 and 7.1.4 are satisfied.
The definition we have given for a randomized stopping time is only the most convenient (for our purposes) of a number of possible equivalent definitions. Although Lemma 7.2.4 below should provide some intuition on what a randomized stopping time is, the reader may still wish to refer to [9, Theorem 3.8] for the other possible ways of defining randomized stopping times. The first step in connecting condition (7.2.3), which is one of the equivalent conditions listen in said theorem, to the others, is to notice that (7.2.3) can be rewritten as

$$\int \left(\int F(r)\, d\pi_\omega(r)\right)(G(\omega) - \mathbb{E}[G|\mathcal{F}^t_s](\omega))\, d\mathbb{W}^t_\kappa(\omega) = 0,$$

[2] And that choice is rather arbitrary itself, as close reading will reveal.

where π_ω is a disintegration of π with respect to $\mathbb{W}^t_\kappa$. This says that the function $\omega \mapsto \int F(r)\,d\pi_\omega(r)$ is orthogonal to $G - \mathbb{E}[G|\mathcal{F}^t_s]$ for all bounded continuous G, i.e., that it is a.s. $\mathcal{F}^t_s$-measurable whenever F is supported on $[t,s]$. A limit argument then shows that $\omega \mapsto \pi_\omega([t,s])$ is a.s. $\mathcal{F}^t_s$-measurable. Again, we refer the reader to [9] for a more detailed exposition.

Problem (OptStop$^\pi$). Among all randomized stopping times $\pi \in \mathsf{RST}_\lambda(\nu)$ find the minimizer of

$$\pi^* \mapsto \int c\,d\pi^* \,.$$

Lemma 7.2.4. *See [9, Lemma 3.11]:*
Let τ be an a.s. finite $(\mathcal{G}_t)_{t\geq 0}$-stopping time and consider

$$\begin{aligned}\Phi &: \Omega \to C(\mathbb{R}_+)\times[0,\infty],\\ \Phi(\omega) &:= ((B_t(\omega))_{t\geq 0}, \tau(\omega))\,.\end{aligned}$$

Then $\pi := \Phi_*(\mathbb{P})_{\restriction C(\mathbb{R}_+)\times\mathbb{R}_+}$ is a finite randomized stopping time, i.e., $\pi \in \mathsf{RST}_\lambda(\mathcal{P})$, and for any measurable process $F: C(\mathbb{R}_+)\times\mathbb{R}_+ \to \mathbb{R}$ we have

$$\int F\,d\pi = \mathbb{E}[F\cdot \mathbb{1}_{C(\mathbb{R}_+)\times\mathbb{R}_+}\circ\Phi] = \mathbb{E}[F(B,\tau)\cdot\mathbb{1}_{\mathbb{R}_+}(\tau)]\,. \tag{7.2.5}$$

For any $\pi \in \mathsf{RST}_\lambda(\mathcal{P})$, we can find an a.s. finite $(\mathcal{G}_t)_{t\geq 0}$-stopping time τ such that $\pi = \Phi_*(\mathbb{P})$ and (7.2.5) holds.
π is a finite randomized stopping time if and only if τ is a.s. finite.

Proof of Theorem 7.1.10. We prove Version B of the theorem. Version A is a special case. We show that Problem OptStop$^\pi$ has a solution. To this end we show that the set $\mathsf{RST}_\lambda(\nu)$ is compact (in the weak topology). From the fact that c is lower semicontinuous and bounded from below in an appropriate sense we then deduce by the Portmanteau theorem that the map

$$\begin{aligned}\hat{c} &: \mathsf{RST}_\lambda(\nu) \to (-\infty,\infty],\\ \hat{c}(\zeta) &:= \int c\,d\zeta\end{aligned}$$

is lower semicontinuous and therefore that the infimum $\inf_{\zeta\in\mathsf{RST}_\lambda(\nu)}\hat{c}(\zeta)$ is attained.
Now for the details: On each of the spaces $C(\mathbb{R}_+)$ and $\mathbb{R}_+$ we are dealing with two topologies, one coming from the (7.2.2) of randomized stopping times (to wit, the topology of uniform convergence on compacts on the space $C(\mathbb{R}_+)$ and the usual order-induced topology on $\mathbb{R}_+$) and one coming from the assumptions in the statement of this theorem. We can equip each of these spaces with the smallest topology which contains the two topologies in question. These are again Polish topologies and they still generate the standard sigma-algebras on the respective spaces. For the remainder of this proof all topological notions are to be understood relative to these topologies. So the topology on $C(\mathbb{R}_+)\times\mathbb{R}_+$ is the product topology of these two topologies, and the weak topology on the space of measures on $C(\mathbb{R}_+)\times\mathbb{R}_+$ is to be understood relative to this product topology, etc. The cost function c of course remains lower semicontinuous and by (7.2.1) the functions $(\omega,r)\mapsto F(r)\big(G(\omega)-\mathbb{E}[G|\mathcal{F}^0_s]\big)$ appearing in (7.2.2) are continuous.

Note that for $\pi \in \mathsf{RST}_\lambda(\nu)$ as ν has mass 1, so must π and $(\mathrm{proj}_{C(\mathbb{R}_+)})_*(\pi)$, which together with $(\mathrm{proj}_{C(\mathbb{R}_+)})_*(\pi) \leq \mathbb{W}^0_\lambda$ implies $(\mathrm{proj}_{C(\mathbb{R}_+)})_*(\pi) = \mathbb{W}^0_\lambda$. So we deduce

$$\mathsf{RST}_\lambda(\nu) = \left\{\pi \in \mathrm{Cpl}(\mathbb{W}^0_\lambda, \nu) : \int F(s)\Big(G - \mathbb{E}[G|\mathbb{F}^0_t]\Big)(\omega)\, d\pi(\omega, s) = 0 \ \ \forall (t, F, G) \in \star\right\},$$

where

$$\pi \in \mathrm{Cpl}(\mathbb{W}^0_\lambda, \nu) \iff (\mathrm{proj}_{C(\mathbb{R}_+)})_*(\pi) = \mathbb{W}^0_\lambda \text{ and } (\mathrm{proj}_{\mathbb{R}_+})_*(\pi) = \nu\,,$$

$(t, F, G) \in \star \iff$ $t > 0$, $F\colon \mathbb{R}_+ \to \mathbb{R}$ is bounded and continuous in the order-induced topologies, and 0 outside $[0, t]$, $G\colon C(\mathbb{R}_+) \to \mathbb{R}$ is bounded and continuous as a function from the topology of uniform convergence on compacts.

The set $\mathrm{Cpl}(\mathbb{W}^0_\lambda, \nu)$ is compact by Prokhorov's Theorem and the fact that pushforwards are continuous maps between measure spaces. The set $\mathrm{Cpl}(\mathbb{W}^0_\lambda, \nu)$ is closed because pushforwards are continuous maps. We show that it is also tight, so that Prokhorov's Theorem implies that it is compact. Let $\varepsilon > 0$ and choose compact sets $K_1 \subseteq C(\mathbb{R}_+)$, $K_2 \subseteq \mathbb{R}_+$ such that $\mathbb{W}^0_\lambda(K_1^c) < \frac{\varepsilon}{2}$ and $\nu(K_2^c) < \frac{\varepsilon}{2}$, then for all $\pi \in \mathrm{Cpl}(\mathbb{W}^0_\lambda, \nu)$ we have $\pi((K_1 \times K_2)^c) \leq \pi(K_1^c \times \mathbb{R}_+) + \pi(C(\mathbb{R}_+) \times K_2^c) < \varepsilon$. It remains to show that $\mathsf{RST}_\lambda(\nu)$ is a non-empty closed subset. It is non-empty because the product measure $\mathbb{W}^0_\lambda \otimes \nu \in \mathsf{RST}_\lambda(\nu)$. It is closed because, as noted, the function $(\omega, s) \mapsto F(s)\Big(G - \mathbb{E}[G|\mathcal{F}^0_t]\Big)(\omega)$ is continuous for all $(t, F, G) \in \star$.
Now, we show that $\hat{c}$ is lower semicontinuous. The functions $c^N := c \vee -N$ are each bounded from below and lower semicontinuous. By the Portmanteau theorem the maps $\hat{c}^N := \zeta \mapsto \int c^N\, d\zeta$ are lower semicontinuous. On $\mathsf{RST}_\lambda(\nu)$ they converge uniformly to $\hat{c}$ because

$$\sup_\zeta \left|\hat{c}(\zeta) - \hat{c}^N(\zeta)\right| \leq \sup_\zeta \int \left|c - c^N\right| d\zeta \leq \sup_{\zeta \in \mathsf{RST}_\lambda(\nu)} \int c_- \cdot 1_{\{c_- \geq N\}}\, d\zeta\,,$$

which converges to 0 as N goes to ∞ by the uniform integrability assumption. As a uniform limit of lower semicontinuous functions is again lower semicontinuous, we see that $\hat{c}$ is lower semicontinuous. ❑

7.3. Geometry of the Optimizer

This section is devoted to the proof of Theorem 7.1.18. The proof closely mimics that of Theorem 1.3/Theorem 5.7 in [9]. For the benefit of those readers already familiar with said paper we will first describe the changes required to the proofs there to make them work in our situation and then – for the sake of a more self-contained presentation – indulge in reiterating the main arguments and only citing results from [9] that we can use verbatim.

Sketch of differences in the proof of Theorem 7.1.18 relative to [9, Theorem 5.7].
Again the strategy is to show that for a larger set $\widehat{\mathrm{SG}}^{\xi} \supseteq \mathrm{SG}$ we can find a set $\Gamma \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$ such that $\widehat{\mathrm{SG}}^{\xi} \cap (\Gamma^< \times \Gamma) = \emptyset$. The definition of $\widehat{\mathrm{SG}}^{\xi}$ must of course be adapted analoguously to the changes required to the definition of SG.
Apart from that the only real changes are to [9, Theorem 5.8]. Whereas previously it was essential that the randomized stopping time $\xi^{r(\omega,s)}$ is also a valid randomized stopping time of the Markov process in question when started at a different time but the same location $\omega(s)$, we now need that $\xi^{r(\omega,s)}$ will also be a randomized stopping time of our Markov process when started at the same time s but in a different place. Of course, when we are talking about Brownian motion both are true, but this difference is the reason why in the case of the Skorokhod embedding the right class of processes. To generalize the argument to that of Feller processes, we do not need in our setup that our processes to be time-homogeneous but to be space-homogeneous. That we are able to plant this „bush" $\xi^{r(\omega,s)}$ in another location is what guarantees that the measure ξ_1^{π} defined in the proof of Theorem 5.8 of [9] is again a randomized stopping time.
Whereas in the Skorokhod case the task is to show that the new better randomized stopping time ξ^{π} *embeds* the same distribution as ξ we now have to show that the randomized stopping time we construct *has* the same distribution as ξ. The argument works along the same lines though – instead of using that $((\omega,s),(\eta,t)) \in \widehat{\mathrm{SG}}^{\xi}$ implies $\omega(s) = \eta(t)$ we now use that $((\omega,s),(\eta,t)) \in \widehat{\mathrm{SG}}^{\xi}$ implies $s = t$. ❑

We now present the argument in more detail. As may be clear by now, what we will show is that if $\xi \in \mathsf{RST}_{\lambda}(\nu)$ is a solution of $\textsc{OptStop}^{\pi}$, then there is a measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted set $\Gamma \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$ such that $\mathrm{SG} \cap (\Gamma^< \times \Gamma) = \emptyset$. Using Lemma 7.2.4 this implies Theorem 7.1.18.

We need to make some preparations. To align the notation with [9] and to make some technical steps easier it is useful to have another characterization of measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted processes and sets. To this end define

Definition 7.3.1.

$$S := \bigcup_{t\in\mathbb{R}_+} C([0,t]) \times \{t\},$$
$$r : C(\mathbb{R}_+) \times \mathbb{R}_+ \to S,$$
$$r(\omega,t) := \left(\omega_{\restriction[0,t]}, t\right).$$

r has many right inverses. A simple one is

$$r' : S \to C(\mathbb{R}_+) \times \mathbb{R}_+$$
$$r'(f,s) := \left(t \mapsto \begin{cases} f(t) & \text{for } t \leq s \\ f(s) & \text{for } t > s \end{cases}, \; s\right).$$

We endow S with the sigma algebra generated by r'.

[9, Theorem 3.2], which is a direct consequence of [21, Theorem IV. 97], asserts that a process X is measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted if and only if X factors as $X = X' \circ r$ for a measurable function $X' : S \to \mathbb{R}$. So a set $D \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$ is measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted if and only if $D = r^{-1}[D']$ for some measurable $D' \subseteq S$.
Note that $r(\omega, t) = r(\omega', t')$ implies $(\omega, t) \odot \theta = (\omega', t') \odot \theta$ and therefore

$$\mathrm{SG} = (r \times r)^{-1}[\mathrm{SG}']$$

for a set $\mathrm{SG}' \subseteq S \times S$ which is described by an expression almost identical to that in Definition 7.1.14. Namely we can overload $\odot$ to also be the name for the operation whose first operand is an element of S, such that $(\omega, t) \odot \theta = r(\omega, t) \odot \theta$ and note that as c is measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted we can write $c = c' \circ r$ and thus get a cost function c' which is defined on S.
Given an optimal $\xi \in \mathsf{RST}_\lambda(\nu)$ we may therefore rephrase our task as having to find a measurable set $\Gamma \subseteq S$ such that $r_*(\xi)$ is concentrated on Γ and that $\mathrm{SG}' \cap (\Gamma^< \times \Gamma) = \emptyset$, where $\Gamma^< := \left\{(g_{\restriction[0,s]}, s) : (g, t) \in \Gamma, s < t\right\}$.
Note that for $\Gamma \subseteq S$ although $\left(r^{-1}[\Gamma]\right)^<$ is not equal to $r^{-1}[\Gamma^<]$ we still have

$$\mathrm{SG} \cap \left(r^{-1}\left[\Gamma^<\right] \times r^{-1}[\Gamma]\right) = \emptyset \iff \mathrm{SG} \cap \left((r^{-1}[\Gamma])^< \times r^{-1}[\Gamma]\right) = \emptyset.$$

One of the main ingredients of the proof of [9, Theorem 1.3] and of our Theorem 7.1.18 is a procedure whereby we accumulate many infinitesimal changes to a given randomized stopping time ξ to build a new stopping time ξ^π. The guiding intuition for the authors is to picture these changes as replacing certain „branches“ of the stopping time ξ by different branches. Some of these branches will actually enter the statement of a somewhat stronger theorem (Theorem 7.3.11 below), so we begin by describing them. Our way to get a handle on „branches“ – i.e., infinitesimal parts of a randomized stopping time – is to describe them through a disintegration (w.r.t. $\mathbb{W}_\lambda^0$) of the randomized stopping time. We need the following statement from [9] which should also serve to provide more intuition on the nature of randomized stopping times.

Lemma 7.3.2. *See [9, Theorem 3.8]:*
Let ξ be a measure on $C(\mathbb{R}_+) \times \mathbb{R}_+$. Then $\xi \in \mathsf{RST}_\lambda$ if and only if there is a disintegration $(\xi_\omega)_{\omega\in C(\mathbb{R}_+)}$ of ξ w.r.t. $\mathbb{W}_\lambda^0$ such that $(\omega, t) \mapsto \xi_\omega([0, t])$ is measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted and maps into $[0, 1]$.

Using Lemma 7.3.2 above let us fix for the rest of this section both $\xi \in \mathsf{RST}_\lambda(\nu)$ and a disintegration $(\xi_\omega)_{\omega\in C(\mathbb{R}_+)}$ with the properties above. Both Definition 7.3.3 below and Theorem 7.3.11 implicitly depend on this particular disintegration and we emphasize that whenever we write ξ_ω in the following we are always referring to the same fixed disintegration with the properties given in Lemma 7.3.2. Note that the measurability properties of $(\xi_\omega)_{\omega\in C(\mathbb{R}_+)}$ imply that for any $I \subseteq [0, s]$ we can determine $\xi_\omega(I)$ from $\omega_{\restriction[0,s]}$ alone. For $(f, s) \in S$ we will again overload notation and use $\xi_{(f,s)}$ to refer to the measure on $[0, s]$ which is equal to $(\xi_\omega)_{\restriction[0,s]}$ for any $\omega \in C(\mathbb{R}_+)$ such that $r(\omega, s) = (f, s)$.

Definition 7.3.3 (Conditional randomized stopping time)**.**
For $(f,s) \in S$, we define a new randomized stopping time $\xi^{(f,s)} \in \mathsf{RST}^s$ by setting

$$\xi^{(f,s)}_\omega := \begin{cases} \frac{1}{1-\xi_{(f,s)}([0,s])}\left(\xi_{(f,s)\odot\omega}\right)_{\restriction(s,\infty)} & \text{for } \xi_{(f,s)}([0,s]) < 1\,, \\ \delta_s & \text{for } \xi_{(f,s)}([0,s]) = 1\,, \end{cases} \tag{7.3.4}$$

$$\int F\,d\xi^{(f,s)} := \int\int F(\omega,t)\,d\xi^{(f,s)}_\omega(t)d\,\mathbb{W}^s_0(\omega)$$

for all bounded measurable $F : C([s,\infty))\times[s,\infty) \to \mathbb{R}$, i.e., $\left(\xi^{(f,s)}_\omega\right)_{\omega\in C([s,\infty))}$ is the disintegration of $\xi^{(f,s)}$ w.r.t. $\mathbb{W}^s_0$.

Here δ_s is the Dirac measure concentrated at s. Really, the definition in the case where $\xi_{(f,s)}([0,s]) = 1$ is somewhat arbitrary – it is more a convenience to avoid partially defined functions. What we will use is that $\left(1-\xi_{(f,s)}([0,s])\right)\xi^{(f,s)}_\omega = \left(\xi_{(f,s)\odot\omega}\right)_{\restriction(s,\infty)}$.

Definition 7.3.5 (Relative Stop-Go pairs)**.**
The set SG^ξ consists of all $((f,t),(g,t)) \in S\times S$ (again the times have to match) such that either

$$c'(f,t)+\int c((g,t)\odot\theta,u)\,d\xi^{(f,t)}(\theta,u) < c'(g,t)+\int c((f,t)\odot\theta,u)\,d\xi^{(f,t)}(\theta,u) \tag{7.3.6}$$

or any one of

(a) $\xi^{(f,t)}(C(\mathbb{R}_+)\times\mathbb{R}_+) < 1$ or $\int s^{p_0}\,d\xi^{(f,t)}(\theta,s) = \infty\,,$

(b) the integral on the right hand side equals $\infty\,,$

(c) either of the integrals is not defined

holds. We also define

$$\widehat{\mathrm{SG}}^\xi := \mathrm{SG}^\xi \cup \left\{(f,s)\in S : \xi_{(f,s)}([0,s]) = 1\right\}\times S\,. \tag{7.3.7}$$

Lemma 7.3.9 below says that the numbered cases above are exceptional in an appropriate sense and one may consider them a technical detail. Note that when we say $((f,t),(g,t)) \in \mathrm{SG}^\xi$ we are implicitly saying that $\xi_{(f,t)}([0,t]) < 1$.

Note that the sets SG^ξ and $\widehat{\mathrm{SG}}^\xi$ are measurable (in contrast to SG, which may be more complicated).

Definition 7.3.8. We call a measurable set $F \subseteq S$ evanescent if $r^{-1}[F]$ is evanescent, that is, if $\mathbb{W}^0_\lambda(\mathrm{proj}_{C(\mathbb{R}_+)}[r^{-1}[F]]) = 0$.

Lemma 7.3.9. *See [9, Lemma 5.2]:*
Let $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \to \mathbb{R}$ be some measurable function for which $\int F\, d\xi \in \mathbb{R}$. Then the following sets are evanescent.

- $\left\{(f,s) \in S : \xi^{(f,s)}(C(\mathbb{R}_+) \times \mathbb{R}_+) < 1\right\}$
- $\left\{(f,s) \in S : \int F((f,s) \odot \theta, u)\, d\xi^{(f,s)}(\theta, u) \notin \mathbb{R}\right\}$

Proof. See [9]. ❑

Lemma 7.3.10. *See [9, Lemma 5.4]:*

$$\mathrm{SG}' \subseteq \widehat{\mathrm{SG}}^{\xi}.$$

Proof. Can be found in [9]. Note that they fix $p_0 = 1$. ❑

Theorem 7.3.11. *Assume that ξ is a solution of* OptStop$^{\pi}$*. Then there is a measurable set $\Gamma \subseteq S$ such that $r_*(\xi)(\Gamma) = 1$ and*

$$\widehat{\mathrm{SG}}^{\xi} \cap \left(\Gamma^{<} \times \Gamma\right) = \emptyset. \tag{7.3.12}$$

Our argument follows [9, Theorem 5.7]. We also need the following two auxilliary propositions, which in turn require some definitions.

Definition 7.3.13. Let v be a probability measure on some measure space Y. The set $\mathsf{JOIN}_{\lambda}(v)$ is the set of all subprobability measures π on $(C(\mathbb{R}_+) \times \mathbb{R}_+) \times Y$ such that

$$\begin{aligned}&(\operatorname{proj} Y)_*(\pi) \le v \quad \text{and}\\ &(\operatorname{proj} C(\mathbb{R}_+) \times \mathbb{R}_+)_*(\pi_{\restriction C(\mathbb{R}_+)\times\mathbb{R}_+\times D}) \in \mathsf{RST}_{\lambda} \quad \text{for all measurable } D \subseteq Y.\end{aligned}$$

Proposition 7.3.14. *Assume that ξ is a solution of* OptStop$^{\pi}$*. Then we have*

$$(r \times \mathsf{Id})_*(\pi)(\mathrm{SG}^{\xi}) = 0$$

for all $\pi \in \mathsf{JOIN}_{\lambda}(r_*(\xi))$.

Here we use $\times$ to denote the Cartesian product map, i.e., for sets X_i, Y_i and functions $F_i : X_i \to Y_i$ where $i \in \{1,2\}$ the map $F_1 \times F_2 : X_1 \times X_2 \to Y_1 \times Y_2$ is given by $(F_1 \times F_2)(x_1, x_2) = (F_1(x_1), F_2(x_2))$. Proposition 6.10 is an analogue of [9, Proposition 5.8] and it is where the material changes compared to [9] take place. We will give the proof at the end of this section.

Proposition 7.3.15. *See [9, Proposition 5.9]:*
Let (Y, v) be a Polish probability space and let $E \subseteq S \times Y$ be a measurable set. Then the following are equivalent

(a) $(r \times \mathsf{Id})_*(\pi)(E) = 0$ *for all* $\pi \in \mathsf{JOIN}_{\lambda}(v)$

(b) $E \subseteq (F \times Y) \cup (S \times N)$ *for some evanescent set $F \subseteq S$ and a measurable set $N \subseteq Y$ which satisfies $v(N) = 0$.*

Proposition 7.3.15 is proved in [9] and we will not repeat the proof here.

Proof of Theorem 7.3.11. Using Proposition 7.3.14 we see that $(r \times \mathrm{Id})_*(\pi)(\mathsf{SG}^\xi) = 0$ for all $\pi \in \mathsf{JOIN}_\lambda(r_*(\xi))$. Plugging this into Proposition 7.3.15 we find an evanescent set $F_1 \subseteq S$ and a set $N \subseteq S$ such that $r_*(\xi)(N) = 0$ and $\mathsf{SG}^\xi \subseteq (F_1 \times S) \cup (S \times N)$. Defining for any Borel set $E \subseteq S$ the analytic set

$$E^> := \left\{(g,t) \in S : \exists s < t, \left(g_{\restriction[0,s]}, s\right) \in E\right\},$$

we observe that $((E^>)^c)^< \subseteq E^c$ and find $r_*(\xi)(F_1^>) = 0$.
Setting $F_2 := \left\{(f,s) \in S : \xi_{(f,s)}([0,s]) = 1\right\}$ and arguing on the disintegration $(\xi_\omega)_{\omega \in C(\mathbb{R}_+)}$ we see that $r_*(\xi)(F_2^>) = 0$, so $r_*(\xi)(F^>) = 0$ for $F := F_1 \cup F_2$.
This shows that $S \setminus (N \cup F^>)$ has full $r_*(\xi)$-measure. Let Γ be a Borel subset of that set which also has full $r_*(\xi)$-measure.
Then

$$\Gamma^< \times \Gamma \subseteq \left((F^>)^c\right)^< \times N^c \subseteq F^c \times N^c \text{ and}$$
$$\widehat{\mathsf{SG}}^\xi \subseteq (F \times S) \cup (S \times N)$$

which shows $\widehat{\mathsf{SG}}^\xi \cap (\Gamma^< \times \Gamma) = \emptyset$. ❑

Lemma 7.3.16. *If $\alpha \in \mathsf{RST}_\lambda$ and $G : C(\mathbb{R}_+) \times \mathbb{R}_+ \to [0,1]$ is measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted, then the measure defined by*

$$F \mapsto \int F(\omega,t) G(\omega,t)\, d\alpha(\omega,t) \tag{7.3.17}$$

is still in RST_λ.

Proof. We use the criterion in Lemma 7.3.2. Let $(\alpha_\omega)_{\omega \in C(\mathbb{R}_+)}$ be a disintegration of α w.r.t. $\mathbb{W}_\lambda^0$ for which $(\omega,t) \mapsto \alpha_\omega([0,t])$ is measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted and maps into $[0,1]$. Then $(\hat\alpha_\omega)_\omega$ defined by $\hat\alpha_\omega := F \mapsto \int F(t) G(\omega,t)\, d\alpha_\omega(t)$ is a disintegration of the measure in (7.3.17) for which $(\omega,t) \mapsto \hat\alpha_\omega([0,t])$ is measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted and maps into $[0,1]$. ❑

Lemma 7.3.18 (Strong Markov property for RSTs). *Let $\alpha \in \mathsf{RST}_\lambda$. Then*

$$\int F(\omega,t)\, d\alpha(\omega,t) = \iint F((\omega,t) \odot \tilde\omega, t)\, d\mathbb{W}_0^t(\tilde\omega)\, d\alpha(\omega,t)$$

for all bounded measurable $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \to \mathbb{R}$.

Proof. Using integral notation instead of the more conventional $\mathbb{E}$, we may write the classical form of the strong markov property as

$$\int G(\Theta_{\tau(\omega)}(\omega)) H(\omega) \cdot \mathbb{1}_{\mathbb{R}_+}(\tau(\omega))\, d\mathbb{W}_\lambda^0(\omega) =$$
$$\int\int G(\tilde\omega) H(\omega) \cdot \mathbb{1}_{\mathbb{R}_+}(\tau(\omega))\, d\mathbb{W}_{\omega(\tau(\omega))}^{\tau(\omega)}(\tilde\omega)\, d\mathbb{W}_\lambda^0(\omega)$$

for all bounded measurable $G : C(\mathbb{R}_+) \to \mathbb{R}$ and all bounded $\mathcal{F}^0_\tau$-measurable $H : C(\mathbb{R}_+) \to \mathbb{R}$. Here Θ_t is the function which cuts off the initial segment of a path up to time t. From this a simple monotone class argument shows that

$$\int K(\Theta_{\tau(\omega)}(\omega),\omega)\cdot\mathbb{1}_{\mathbb{R}_+}(\tau(\omega))\,d\mathbb{W}^0_\lambda(\omega) = \int\int K(\tilde\omega,\omega)\cdot\mathbb{1}_{\mathbb{R}_+}(\tau(\omega))\,d\mathbb{W}^{\tau(\omega)}_{\omega(\tau(\omega))}(\tilde\omega)\,d\mathbb{W}^0_\lambda(\omega)$$

for all bounded $\mathcal{F}^0_\infty\otimes\mathcal{F}^0_\tau$-measurable $K : C(\mathbb{R}_+)\times C(\mathbb{R}_+) \to \mathbb{R}$.
We may then choose for $K(\tilde\omega,\omega)$ the function $F(\eta,\tau(\omega))$ where the path η is created by cutting off the tail of ω after time $\tau(\omega)$ and attaching $\tilde\omega$ in its place. Noting the relationship between $\mathbb{W}^{\tau(\omega)}_x$ and $\mathbb{W}^{\tau(\omega)}_0$ we then get

$$\int F(\omega,\tau(\omega))\cdot\mathbb{1}_{\mathbb{R}_+}(\tau(\omega))\,d\mathbb{W}^0_\lambda(\omega) = \int\int F((\omega,\tau(\omega))\odot\tilde\omega,\tau(\omega))\cdot\mathbb{1}_{\mathbb{R}_+}(\tau(\omega))\,d\mathbb{W}^{\tau(\omega)}_{\omega(\tau(\omega))}(\tilde\omega)\,d\mathbb{W}^0_\lambda(\omega).$$

Using Lemma 7.2.4 with $\Omega = [0,1]\times C(\mathbb{R}_+)$ and $\mathcal{G}_t = \mathcal{B}([0,1])\otimes\mathcal{F}_t$ we find a $(\mathcal{G}_t)_{t\geq 0}$-stopping time τ s.t. we may write α as

$$\alpha = ((y,\omega)\mapsto(\omega,\tau(y,\omega)))_*(\mathcal{L}\otimes\mathbb{W}^0_\lambda)_{\restriction C(\mathbb{R}_+)\times\mathbb{R}_+}$$

(where $\mathcal{L}$ is Lebesgue measure on $[0,1]$). For a fixed $y\in[0,1]$, $\omega\mapsto\tau(y,\omega)$ is an $(\mathcal{F}_t)_{t\geq 0}$-stopping time, so we may apply the previous equation to these stopping times and integrate over $y\in[0,1]$ to get

$$\int F(\omega,\tau(y,\omega))\cdot\mathbb{1}_{\mathbb{R}_+}(\tau(y,\omega))\,d\mathbb{W}^0_\lambda(\omega) = \int\int F((\omega,\tau(y,\omega))\odot\tilde\omega,\tau(y,\omega))\cdot\mathbb{1}_{\mathbb{R}_+}(\tau(y,\omega))\,d\mathbb{W}^{\tau(y,\omega)}_0(\tilde\omega)\,d(\mathcal{L}\otimes\mathbb{W}^0_\lambda)(y,\omega).$$

Using the equation for α we see that this is what we wanted to prove. ❑

Lemma 7.3.19 (Gardener's Lemma)**.**
Assume that we have $\xi\in\mathsf{RST}_\lambda(\mathcal{P})$, a measure α on $C(\mathbb{R}_+)\times\mathbb{R}_+$ and two families $\beta^{(\omega,t)}$, $\gamma^{(\omega,t)}$, where $(\omega,t)\in C(\mathbb{R}_+)\times\mathbb{R}_+$, with $\beta^{(\omega,t)},\gamma^{(\omega,t)}\in\mathsf{RST}^t(\mathcal{P})$ such that both maps

$$(\omega,t)\mapsto\int\mathbb{1}_D((\omega,t)\odot\tilde\omega,s)\,d\beta^{(\omega,t)}(\tilde\omega,s)\text{ and}$$
$$(\omega,t)\mapsto\int\mathbb{1}_D((\omega,t)\odot\tilde\omega,s)\,d\gamma^{(\omega,t)}(\tilde\omega,s)$$

are measurable for all Borel $D\subseteq C(\mathbb{R}_+)\times\mathbb{R}_+$ and that

$$\xi(D)-\iint\mathbb{1}_D((\omega,t)\odot\tilde\omega,s)\,d\beta^{(\omega,t)}(\tilde\omega,s)\,d\alpha(\omega,t)\geq 0 \tag{7.3.20}$$

for all Borel $D\subseteq C(\mathbb{R}_+)\times\mathbb{R}_+$. Then for $\hat\xi$ defined by

$$\int F\,d\hat\xi := \int F\,d\xi - \iint F((\omega,t)\odot\tilde\omega,s)\,d\beta^{(\omega,t)}(\tilde\omega,s)\,d\alpha(\omega,t) + \iint F((\omega,t)\odot\tilde\omega,s)\,d\gamma^{(\omega,t)}(\tilde\omega,s)\,d\alpha(\omega,t)$$

for all bounded measurable F, we have $\hat\xi\in\mathsf{RST}_\lambda(\mathcal{P})$.

Remark 7.3.21. The intuition behind the Gardener's Lemma is that we are replacing certain branches $\beta^{(\omega,t)}$ of the randomized stopping time ξ by other branches $\gamma^{(\omega,t)}$ to obtain a new stopping time $\hat{\xi}$. This process happens *along* the measure α. Note that (7.3.20) implies that $\int \mathbb{1}_D((\omega,t)\odot\tilde{\omega})\, d\mathbb{W}_0^t(\tilde{\omega})\, d\alpha(\omega,t) \le \mathbb{W}_\lambda^0(D)$ for all Borel $D \subseteq C(\mathbb{R}_+)$. The authors like to think of α as a stopping time and of the maps $(\omega,t)\mapsto\beta^{(\omega,t)}$ and $(\omega,t)\mapsto\gamma^{(\omega,t)}$ as adapted (in some sense that would need to be made precise). As these assumptions aren't necessary for the proof of the Gardener's Lemma, they were left out, but it might help the reader's intuition to keep them in mind.

Proof of Lemma 7.3.19. We need to check that the $\hat{\xi}$ we define is indeed a measure, that $(\operatorname{proj} C(\mathbb{R}_+))_*(\hat{\xi}) = \mathbb{W}_\lambda^0$ and that (7.2.3) holds for $\hat{\xi}$.
Checking that $\hat{\xi}$ is a measure is routine – we just note that (7.3.20) guarantees that $\hat{\xi}(D)\ge 0$ for all Borel sets D.
Let $G : C(\mathbb{R}_+)\to\mathbb{R}$ be a bounded measurable function.

$$\begin{aligned}\int G(\omega)\,d\hat{\xi}(\omega,t) &= \int G(\omega)\,d\xi(\omega,t) - \iint G((\omega,t)\odot\tilde{\omega})\,d\beta^{(\omega,t)}(\tilde{\omega},s)\,d\alpha(\omega,t)\\ &\qquad + \iint G((\omega,t)\odot\tilde{\omega})\,d\gamma^{(\omega,t)}(\tilde{\omega},s)\,d\alpha(\omega,t)\\ &= \int G\,d\mathbb{W}_\lambda^0 \qquad - \iint G((\omega,t)\odot\tilde{\omega})\,d\mathbb{W}_0^t\,d\alpha(\omega,t)\\ &\qquad + \iint G((\omega,t)\odot\tilde{\omega})\,d\mathbb{W}_0^t\,d\alpha(\omega,t)\\ &= \int G\,d\mathbb{W}_\lambda^0.\end{aligned}$$

Now let $F:\mathbb{R}_+\to\mathbb{R}$ and $G: C(\mathbb{R}_+)\to\mathbb{R}$ be bounded continuous functions, with F supported on $[0,r]$.

$$\begin{aligned}\int F(t)\Big(G-\mathbb{E}[G|\mathcal{F}_r^0]\Big)(\omega)\,d\hat{\xi}(\omega,t) &= \int F(t)\Big(G-\mathbb{E}[G|\mathcal{F}_r^0]\Big)(\omega)\,d\xi(\omega,t)\\ &\quad -\iint F(s)\Big(G-\mathbb{E}[G|\mathcal{F}_r^0]\Big)((\omega,t)\odot\tilde{\omega})\,d\beta^{(\omega,t)}(\tilde{\omega},s)\,d\alpha(\omega,t)\\ &\quad -\iint F(s)\Big(G-\mathbb{E}[G|\mathcal{F}_r^0]\Big)((\omega,t)\odot\tilde{\omega})\,d\gamma^{(\omega,t)}(\tilde{\omega},s)\,d\alpha(\omega,t).\end{aligned} \tag{7.3.22}$$

The first summand is 0 because $\xi\in\mathsf{RST}_\lambda(\mathcal{P})$. Looking at the second summand we expand the definition of $\mathbb{E}[G|\mathcal{F}_r^0]$.

$$\begin{aligned}\mathbb{E}[G|\mathcal{F}_r^0]((\omega,t)\odot\tilde{\omega}) &= \int G(((\omega,t)\odot\tilde{\omega},r)\odot\theta)\,d\mathbb{W}_0^r(\theta)\\ &= \int G((\omega,t)\odot((\tilde{\omega},r)\odot\theta))\,d\mathbb{W}_0^r(\theta).\end{aligned}$$

whenever $t\le r$, which is the case for those t which are relevant in the integrand above, because $F(s)\neq 0$ implies $s\le r$ and moreover $\beta^{(\omega,t)}$ is concentrated on $(\tilde{\omega},s)$ for which $t\le s$. Setting $\hat{G}^{(\omega,t)}(\tilde{\omega}) := G((\omega,t)\odot\tilde{\omega})$ and $\hat{F}^{(\omega,t)} := F_{\restriction[t,\infty)}$ we can write

$$\begin{aligned}&\iint F(s)\Big(G-\mathbb{E}[G|\mathcal{F}_r^0]\Big)((\omega,t)\odot\tilde{\omega})\,d\beta^{(\omega,t)}(\tilde{\omega},s)\,d\alpha(\omega,t) =\\ &\qquad \int \mathbb{1}_{[0,r]}(t)\int \hat{F}^{(\omega,t)}(s)\Big(\hat{G}^{(\omega,t)}-\mathbb{E}[\hat{G}^{(\omega,t)}|\mathcal{F}_r^t]\Big)(\tilde{\omega})\,d\beta^{(\omega,t)}(\tilde{\omega},s)\,d\alpha(\omega,t),\end{aligned}$$

which is 0 because $\beta^{(\omega,t)}\in\mathsf{RST}^t(\mathcal{P})$ and therefore

$$\int \hat{F}^{(\omega,t)}(s)\Big(\hat{G}^{(\omega,t)}-\mathbb{E}[\hat{G}^{(\omega,t)}|\mathcal{F}_r^t]\Big)(\tilde{\omega})\,d\beta^{(\omega,t)}(\tilde{\omega},s)=0$$

for all (ω,t) and $r\ge t$. The same argument works for the third summand in (7.3.22). ❑

Proof of Proposition 7.3.14. We prove the contrapositive. Assuming that there exists a $\pi' \in \mathsf{JOIN}_\lambda(r_*(\xi))$ with $(r\times \mathsf{Id})_*(\pi')(\mathsf{SG}^\xi) > 0$, we construct a $\xi^\pi \in \mathsf{RST}_\lambda(\nu)$ such that $\int c\, d\xi^\pi < \int c\, d\xi$. If $\pi' \in \mathsf{JOIN}_\lambda(r_*(\xi))$, then for any two measurable sets $D_1, D_2 \subseteq S$, because $\pi'_{\restriction (C(\mathbb{R}_+)\times\mathbb{R}_+)\times D_2} \in \mathsf{RST}_\lambda$ and by making use of Lemma 7.3.16 we can see that $(\mathrm{proj}_{C(\mathbb{R}_+)\times\mathbb{R}_+})_*(\pi'_{\restriction (r\times\mathsf{Id})^{-1}[D_1\times D_2]}) \in \mathsf{RST}_\lambda$. Using the monotone class theorem this extends to any measurable subset of $S\times S$ in place of $D_1\times D_2$. So we can set $\pi := \pi'_{\restriction (r\times\mathsf{Id})^{-1}[\mathsf{SG}^\xi]}$ and know that $(\mathrm{proj}_{C(\mathbb{R}_+)\times\mathbb{R}_+})_*(\pi) \in \mathsf{RST}_\lambda$ and that π is concentrated on SG^ξ.
We will be using a disintegration of π w.r.t. $r(\xi)$, which we call $\left(\pi_{(g,t)}\right)_{(g,t)\in S}$ and for which we assume that $\pi_{(g,t)}$ is a subprobability measure for all $(g,t)\in S$. It will also be useful to assume that $\pi_{(g,t)}$ is concentrated on the set $\{(\omega,s)\in C(\mathbb{R}_+)\times\mathbb{R}_+ : s = t\}$ not just for $r(\xi)$-almost all (g,t) but for all (g,t). Again, this is no restriction of generality. We will also push π onto $(C(\mathbb{R}_+)\times\mathbb{R}_+)\times(C(\mathbb{R}_+)\times\mathbb{R}_+)$, defining a measure $\bar\pi$ via

$$\int F\, d\bar\pi := \iint F\left((\omega,s),((g,t)\odot\tilde\eta,t)\right) d\mathbb{W}_0^t(\tilde\eta)\, d\pi\left((\omega,s),(g,t)\right)$$

for all bounded measurable F. Observe that by Lemma 7.3.18 the pushforward of π under projection onto the second coordinate (pair) is ξ and that a disintegration of $\bar\pi$ w.r.t. to ξ (again in the second coordinate) is given by $\left(\pi_{r(\eta,t)}\right)_{(\eta,t)\in C(\mathbb{R}_+)\times\mathbb{R}_+}$. Let us name $(\mathrm{proj}\, C(\mathbb{R}_+)\times\mathbb{R}_+)_*(\pi) =: \zeta \in \mathsf{RST}_\lambda$. We will now use the Gardener's Lemma to define two modifications ξ_0^π, ξ_1^π of ξ such that $\xi^\pi := \frac{1}{2}(\xi_0^\pi + \xi_1^\pi)$ is our improved randomized stopping time.
For all bounded measurable $F : C(\mathbb{R}_+)\times\mathbb{R}_+ \to \mathbb{R}$ define

$$\begin{aligned}\int F\, d\xi_0^\pi &:= \int F\, d\xi + \int (1-\xi_\omega([0,s]))\Big(-\int F((\omega,s)\odot\tilde\omega,u)\, d\xi^{r(\omega,s)}(\tilde\omega,u)\\ &\qquad\qquad + F(\omega,s)\Big)\, d\zeta(\omega,s)\\ \int F\, d\xi_1^\pi &:= \int F\, d\xi + \int (1-\xi_\omega([0,s]))\Big(-F(\eta,t)\\ &\qquad + \int F((\eta,t)\odot\tilde\omega,u)\, d\xi^{r(\omega,s)}(\tilde\omega,u)\Big)\, d\bar\pi((\omega,s),(\eta,t))\,.\end{aligned}$$

The concatenation on the last line is well-defined $\bar\pi$-almost everywhere because $\bar\pi$ is concentrated on $(r\times r)^{-1}\left[\mathsf{SG}^\xi\right]$ and so in the integrand above $s = t$ on a set of full measure. We need to check that the Gardener's Lemma applies in both cases. First of all observe that the product measure $\mathbb{W}_0^t \otimes \delta_t$ is in $\mathsf{RST}^t(\mathcal{P})$ and that Equation 7.3.18 implies

$$\int F(\omega,t)\, d\alpha(\omega,t) = \iint F((\omega,t)\odot\tilde\omega,s)\, d\left(\mathbb{W}_0^t\otimes\delta_t\right)(\tilde\omega,s)\, d\alpha(\omega,t)$$

for any randomized stopping time α. So for ξ_0^π the measures $\gamma^{(\omega,t)}$ are given by $\mathbb{W}_0^t\otimes\delta_t$ and for ξ_1^π the measures $\beta^{(\omega,t)}$ are given by $\mathbb{W}_0^t\otimes\delta_t$.
For ξ_0^π the measure along which we are replacing branches is given by

$$F \mapsto \int F(\omega,s)(1-\xi_\omega([0,s]))\, d\zeta(\omega,s)\,.$$

The branches $\beta^{(\omega,s)}$ we remove are $\xi^{r(\omega,s)}$. We need to check that

$$\int F\, d\xi - \int (1-\xi_\omega([0,s]))\int F((\omega,s)\odot\tilde\omega,u)\, d\xi^{r(\omega,s)}(\tilde\omega,u)\, d\zeta(\omega,s) \geq 0$$

for all positive, bounded, measurable $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \to \mathbb{R}$. Let us calculate.

$$\begin{aligned}
&\int (1-\xi_\omega([0,s])) \int F((\omega,s)\odot\tilde\omega,u)\, d\xi^{r(\omega,s)}(\tilde\omega,u)\, d\zeta(\omega,s) \\
&= \iiint F((\omega,s)\odot\tilde\omega,u)\, d\Big((\xi_{(\omega,s)\odot\tilde\omega})_{\upharpoonright(s,\infty)}\Big)(u)\, d\mathbb{W}_0^s(\tilde\omega)\, d\zeta(\omega,s) \\
&= \iint F(\omega,u)\, d\Big((\xi_\omega)_{\upharpoonright(s,\infty)}\Big)(u)\, d\zeta(\omega,s) \le \iint F(\omega,u)\, d(\xi_\omega)(u)\, d\zeta(\omega,s) \\
&\le \iint F(\omega,u)\, d(\xi_\omega)(u)\, d\mathbb{W}_\lambda^0(\omega) = \int F(\omega,u)\, d\xi(\omega,u).
\end{aligned}$$

Here we first used the definition of $\xi^{r(\omega,s)}$ and then Lemma 7.3.18 and finally that $(\mathrm{proj}_{C(\mathbb{R}_+)})_*(\zeta) \le \mathbb{W}_\lambda^0$.
For ξ_1^π we replace branches along

$$\begin{aligned}
F \mapsto &\int F(\eta,t)(1-\xi_\omega([0,s]))\, d\bar\pi\,((\omega,s),(\eta,t)) \\
&= \int F(\eta,t) \int (1-\xi_\omega([0,s]))\, d\pi_{r(\eta,t)}(\omega,s)\, d\xi(\eta,t)\;.
\end{aligned}$$

The calculation above shows that

$$\int F\, d\xi - \int (1-\xi_\omega([0,s]))F(\eta,t)\, d\bar\pi\,((\omega,s),(\eta,t)) \ge 0$$

for all positive, bounded, measurable $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \to \mathbb{R}$. For ξ_1^π the branches $\gamma^{(\eta,t)}$ that we add are given by

$$F \mapsto \frac{\int (1-\xi_\omega([0,s])) \int F(\tilde\omega,u)\, d\xi^{r(\omega,s)}(\tilde\omega,u)\, d\pi_{r(\eta,t)}(\omega,s)}{\int (1-\xi_\omega([0,s]))\, d\pi_{r(\eta,t)}(\omega,s)}$$

when $\int (1-\xi_\omega([0,s]))\, d\pi_{r(\eta,t)}(\omega,s) > 0$ and δ_t otherwise (again, the latter is arbitrary). In the more interesting case $\gamma^{(\eta,t)}$ is an average over elements of $\mathsf{RST}^t(\mathcal{P})$ and therefore itself in $\mathsf{RST}^t(\mathcal{P})$. Here it is again crucial that for $\pi_{r(\eta,t)}$-almost all (ω,s) we have $s=t$, otherwise we would be averaging randomized stopping times of our process started at unrelated times.
Putting this together we see that $\xi^\pi := \frac{1}{2}(\xi_0^\pi + \xi_1^\pi)$ is a randomized stopping time and that

$$\begin{aligned}
2\int F\, d(\xi^\pi - \xi) = \int (1-\xi_\omega([0,s]))\Big(&F(\omega,s) - \int F((\omega,s)\odot\tilde\omega,u)\, d\xi^{r(\omega,s)}(\tilde\omega,u) - F(\eta,t) \\
&+ \int F((\eta,t)\odot\tilde\omega,u)\, d\xi^{r(\omega,s)}(\tilde\omega,u)\Big) d\bar\pi((\omega,s),(\eta,t)) \quad (7.3.23)
\end{aligned}$$

for all bounded measurable $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \to \mathbb{R}$. Specializing to $F(\omega,s) = G(s)$ for $G : \mathbb{R}_+ \to \mathbb{R}$ bounded measurable we find that

$$\int G(s)\, d(\xi - \xi^\pi)(\omega,s) = 0\;,$$

again because for $\bar\pi$-almost all $((\omega,s),(\eta,t))$ we have $s=t$. This shows that $\xi^\pi \in \mathsf{RST}_\lambda(\nu)$.
Now, we want to extend (7.3.23) to c. We first show that (7.3.23) also holds for $F : C(\mathbb{R}_+) \times \mathbb{R}_+ \to \mathbb{R}$ which are measurable and positive and for which $\int F\, d\xi < \infty$. To

see this, approximate such an F from below by bounded measurable functions (for which (7.3.23) holds) and note that by previous calculations both

$$\int (1-\xi_\omega([0,s]))\int F((\omega,s)\odot\tilde\omega,u)\,d\xi^{r(\omega,s)}(\tilde\omega,u)\,d\bar\pi((\omega,s),(\eta,t)) \le \int F\,d\xi < \infty$$

and
$$\int (1-\xi_\omega([0,s]))F(\eta,t)\,d\bar\pi((\omega,s),(\eta,t)) \le \int F\,d\xi < \infty\,.$$

Looking at positive and negative parts of c and using Assumption 7.1.4.(d) to see that $\int c_-\,d(\xi^\pi - \xi) \in \mathbb{R}$ we get that indeed (7.3.23) holds for $F = c$.
Now, we will argue that the integrand in the right hand side of (7.3.23) is negative $\bar\pi$-almost everywhere. This will conclude the proof.
By inserting an r in appropriate places we can read off from Definition 7.3.5 what it means that $\bar\pi$ is concentrated on $(r\times r)^{-1}\left[\mathrm{SG}^\xi\right]$. In the course of verifying that (7.3.23) applies to c we already saw that cases b and c in Definition 7.3.5 can only occur on a set of $\bar\pi$-measure 0. Lemma 7.3.9 excludes case a $\bar\pi$-almost everywhere. This means that (7.3.6) holds $\bar\pi$-almost everywhere – or more correctly, that for $\bar\pi$-a.a. $((\omega,s),(\eta,t))$ we have $s=t$ and

$$c(\omega,s) - \int c((\omega,s)\odot\tilde\omega,u)\,d\xi^{r(\omega,s)}(\tilde\omega,u) - c(\eta,t) + \int c((\eta,t)\odot\tilde\omega,u)\,d\xi^{r(\omega,s)}(\tilde\omega,u) < 0\,, \tag{7.3.24}$$

completing the proof. ❑

7.4. Special Cases

Both Corollary 7.4.1 and Corollary 7.4.9 assert that the solutions of certain optimal stopping problems can be described by a barrier in an appropriate phase space.

7.4.1. Product of a Brownian Motion and a Deterministic function

Problem ($\textsc{OptStop}^{\psi(B_t,t)}$). Among all stopping times $\tau\sim\nu$ on $(\Omega,\mathcal{G},(\mathcal{G}_t)_{t\ge0},\mathbb{P})$ find the maximizer of

$$\tau \mapsto \mathbb{E}[Z_\tau]\,,$$

where the process Z is of the form $Z_t = \psi(B_t,t)$.

Corollary 7.4.1. *Assume that ν has finite first moment. There is an upper semicontinuous function $\beta : \mathbb{R}_+ \to [-\infty,\infty]$ such that the stopping time*

$$\tau := \inf\{t>0 : B_t \le \beta(t)\} \tag{7.4.2}$$

has distribution ν.
τ has the following uniqueness properties: On the one hand it is the a.s. unique stopping time which has distribution ν and which is of the form (7.4.2) (we will later say that such a stopping time is the hitting time of a downwards barrier*).*
On the other hand τ is also the a.s. unique solution of $\textsc{OptStop}^{\psi(B_t,t)}$ for a number of different ψ. Namely:

- *Let $p \geq 0$, assume μ has finite moment of order $\frac{1}{2}+p+\varepsilon$ for some $\varepsilon > 0$ and let $A : \mathbb{R}_+ \to \mathbb{R}$ be strictly increasing and $|A(t)| \leq K(1+t^p)$ for some constants K*[3]*. Then we may choose*

$$\psi(B_t, t) = B_t A(t).$$

- *Let $p \geq 2$, assume μ has finite moment of order $\frac{p}{2}+\varepsilon$ for some $\varepsilon > 0$ and let $\phi : \mathbb{R} \to \mathbb{R}$ satisfy $\phi''' > 0$ as well as $\left|\phi(y)\right| \leq K(1+|y|^p)$ for constants K. Then we may choose*

$$\psi(B_t, t) = \phi(B_t).$$

Remark 7.4.3. Finally, we note that Corollary 7.4.1 recovers Anulova's classical solution to the inverse first passage time problem [4], which has seen some recent interest (see [27, 18, 36]). Bayraktar and Miller [8] consider the same problem that we treat here. However their setup and methods are rather distinct from the ones used here: they assume that the target distribution is given by finitely many atoms and that the target functional depends solely on the terminal value of Brownian motion. Following the measure valued martingale approach of [20], [8] address the constrained optimal stopping problem using a Bellman perspective.

We will now demonstrate how to use the Monotonicity Principle of Theorem 7.1.18 to derive Corollary 7.4.1.
Both of the sets $\check{\mathcal{R}}$ and $\hat{\mathcal{R}}$ in Lemma 7.1.20 have the property that (writing $\mathcal{R}$ for the set in question) $(y,t) \in \mathcal{R}$ and $y' \leq y$ implies $(y',t) \in \mathcal{R}$. We call such sets (downwards) barriers. More specifically, for technical reasons in what follows it is slightly more convenient to talk about subsets of $[-\infty,\infty] \times \mathbb{R}_+$ instead of subsets of $\mathbb{R} \times \mathbb{R}_+$, giving the following definition.

Definition 7.4.4. Let X be a topological space. A *downwards barrier* is a set $\mathcal{R} \subseteq [-\infty,\infty] \times X$ such that $\{-\infty\} \times X \subseteq \mathcal{R}$ and

$$(y,t) \in \mathcal{R} \text{ and } y' \leq y \text{ imply } (y',t) \in \mathcal{R}.$$

Clearly, in Lemma 7.1.20, instead of talking about $\check{\mathcal{R}} \subseteq \mathbb{R} \times \mathbb{R}_+$, we could have talked about $\check{\mathcal{R}} \cup (\{-\infty\} \times \mathbb{R}_+) \subseteq [-\infty,\infty] \times \mathbb{R}_+$ without anything really changing, and likewise for $\hat{\mathcal{R}}$.

The reader will easily verify the following lemma.

Lemma 7.4.5. *Let X be a topological space. There is a bijection between the set of all upper semicontinuous functions $\beta : X \to [-\infty,\infty]$ and the set of all* closed *downwards barriers $\mathcal{R} \subseteq [-\infty,\infty] \times X$ (where closure is to be understood in the product topology). This bijection maps any upper semicontinuous function β to the barrier $\mathcal{R}$ which is the hypograph of β*

$$\mathcal{R} := \{(y,x) : y \leq \beta(x)\} ,$$

while the inverse maps a barrier $\mathcal{R}$ to the function β given by

$$\beta(x) := \sup\{y : (y,x) \in \mathcal{R}\} .$$

[3] One may of course choose $0 \leq p < 1$, $\varepsilon := 1-p$ and e.g. $A(t) := t^p$ so that no moment conditions beyond those at the very beginning of this theorem are imposed on μ.

On the way to proving Corollary 7.4.1, we will show now that the first hitting time after 0 of any downwards barrier by Brownian motion is a.s. equal to the first hitting time after 0 of the closure of that barrier. This serves to both resolve the question whether the times in Lemma 7.1.20 are stopping times and to show that $\check{\tau} = \hat{\tau}$, a.s.
Let us assume for the rest of this section that B is actually a Brownian motion started in 0.

Lemma 7.4.6. *Let $\mathcal{R}$ be a downwards barrier in $[-\infty,\infty]\times\mathbb{R}_+$. Let $\bar{\mathcal{R}}$ be the closure of $\mathcal{R}$ (in the product topology of the usual (order-induced) topologies on $[-\infty,\infty]$ and $\mathbb{R}_+$). Define*

$$\tau(\omega) := \inf\{t > 0 : (B_t(\omega), t) \in \mathcal{R}\},$$
$$\bar{\tau}(\omega) := \inf\{t > 0 : (B_t(\omega), t) \in \bar{\mathcal{R}}\}.$$

Then $\tau = \bar{\tau}$, a.s.

Proof. As $\bar{\mathcal{R}} \supseteq \mathcal{R}$ we clearly have $\bar{\tau}(\omega) \leq \tau(\omega)$ for all $\omega \in \Omega$. Define

$$\bar{\tau}_\varepsilon(\omega) := \inf\{t > 0 : (B_t(\omega) + A(t)\cdot\varepsilon, t) \in \bar{\mathcal{R}}\},$$

where $A(t) := \frac{t}{1+t}$ is a bounded, strictly increasing function. Using just that $\bar{\mathcal{R}}$ is the closure of $\mathcal{R}$ one proves by elementary methods that $\tau(\omega) \leq \bar{\tau}_\varepsilon(\omega)$ for all $\omega \in \Omega$ and any $\varepsilon > 0$. Because $A(t) = \int_0^t (1+s)^{-2}\,ds$ is the integral from 0 to t of a square integrable function we can apply Girsanov's theorem (see e.g. [63, Theorem 38.5]) to see that $\bar{\tau}_{1/n}$ converges to $\bar{\tau}$ in distribution as $n \to \infty$. As $(\bar{\tau}_{1/n})_n$ is a decreasing sequence bounded below by $\bar{\tau}$ we get that convergence holds almost surely. ❑

The following is a particular case of [31, Corollary 2.3] (which in turn relies on arguments given in [64, 48]). Note that this lemma is purely a statement about barrier-type stopping times and is not directly connected to the optimization problem under consideration.

Lemma 7.4.7 (Uniqueness of Barrier-type solutions)**.**
Let $(Y_t)_{t\geq 0}$ be a measurable, $(\mathcal{F}_t^0)_{t\geq 0}$-adapted process and assume that the process Z defined through $Z_t := Y_t(B)$ has a.s. continuous paths. Let $\mathcal{R}_1, \mathcal{R}_2 \subseteq [-\infty,\infty]\times\mathbb{R}_+$ be closed downwards barriers such that for

$$\tau_i(\omega) := \inf\{t > 0 : (Z_t(\omega), t) \in \mathcal{R}_i\}$$

we have $\tau_1 \sim \tau_2$. Then $\tau_1 = \tau_2$, a.s.

Proof. Is to be found in [31, Corollary 2.3]. ❑

Now, we have the necessary prerequisites to use our main results to show that the considered optimization problem in this subsection admits a (unique) barrier-type solution.

Proof of Corollary 7.4.1. The strategy is as follows: We choose a cost function and leverage Theorem 7.1.10 to show that an optimizer exists, the Monotonicity Principle in the form of Theorem 7.1.18 and Lemma 7.1.20 will – with some help from Lemma 7.4.6 – show that any optimizer must be the hitting time of a barrier. Lemma 7.4.7 shows that any two barrier-type solutions must be equal.

Now, we provide the details. Start with a cost function $c(\omega,t) := -\omega(t)A(t)$ for a strictly monotone function $A : \mathbb{R}_+ \to \mathbb{R}$ which satisfies $|A(t)| \leq K(1+t^p)$ and assume that ν has moment of order $\frac{1}{2}+p+\varepsilon$ for some $\varepsilon > 0$. To prove that a barrier type solution exists when ν has first moment, choose a bounded strictly increasing A and $p = 0$, $\varepsilon = \frac{1}{2}$ in this step. (These assumptions guarantee in particular that the optimization problems considered below have a finite value.) Clearly the problem OptStop^τ for c corresponds to $\text{OptStop}^{\psi(B_t,t)}$ for $\psi(B_t,t) = B_tA(t)$ (i.e., ψ takes the role for $-c$ such that the minimal/maximal values agree up to a change of sign). We will deal with the case where $\psi(B_t,t) = \phi(B_t)$ at the end of this proof.

We now check that the conditions in Version B of Theorem 7.1.10 are satisfied. We also need to check that Assumption 7.1.4 holds. Here we need the assumption that ν has moment of order $\frac{1}{2}+p+\varepsilon$, as well as the Hölder and Burkholder-Davis-Gundy inequalities. The latter specialized to Brownian motion state that for all $q > 0$ there are positive constants K_0 and K_1 such that for any stopping time τ we have

$$K_0\mathbb{E}[\tau^{q/2}] \leq \mathbb{E}[(|B|^*_\tau)^q] \leq K_1\mathbb{E}[\tau^{q/2}],$$

where $|B|^*_t = \sup_{s\leq t}|B_s|$. With these in hands straightforward calculation allows us to bound $B_\tau A(\tau)$ in the $L^{1+\delta}$-norm for some $\delta > 0$, independently of the stopping time $\tau \sim \nu$. This shows both that the uniform integrability condition in Version B of Theorem 7.1.10 is satisfied and that Assumption 7.1.4.(d) is satisfied.

On $C(\mathbb{R}_+)$ we may choose the (Polish) topology of uniform convergence on compacts. For the topology on $\mathbb{R}_+$ we start with the usual topology and turn A into a continuous function (if it wasn't), by making use of the fact that any measurable function from a Polish space to a second countable space may be turned into a continuous function by passing to a larger Polish topology (with the same Borel sets) on the domain. (This can be found for example in [40, Theorem 13.11].)

In the statement of Corollary 7.4.1 we did not ask for the probability space $(\Omega,\mathcal{G},(\mathcal{G}_t)_{t\geq 0},\mathbb{P})$ to satisfy Assumption 7.1.4.(b). To remedy this, we can enlarge the probability space by setting $\tilde\Omega := \Omega\times[0,1]$, $\tilde{\mathcal{G}}_t := \mathcal{G}_t \otimes \mathcal{B}([0,1])$ and $\tilde{\mathbb{P}} := \mathbb{P}\otimes\mathcal{L}$, where $\mathcal{L}$ is Lebesgue measure on $[0,1]$. On this space we consider the Brownian motion $\tilde B_t(\omega,x) := B_t(\omega)$. Theorem 7.1.10 now gives us an optimal stopping time $\tilde\tau$ on the enlarged probability space. If we can show that this stopping time is in fact the hitting time of a barrier, then it follows that $\tilde\tau = \tau\circ((\omega,x)\mapsto\omega)$ for a stopping time τ which is defined as the hitting time of the Brownian motion B of the same barrier. As there are *more* stopping times on $(\tilde\Omega,\tilde{\mathcal{G}},(\tilde{\mathcal{G}}_t)_{t\geq 0})$ than on $(\Omega,\mathcal{G},(\mathcal{G}_t)_{t\geq 0})$ in the sense that any stopping time τ' on $(\Omega,\mathcal{G},(\mathcal{G}_t)_{t\geq 0})$ induces a stopping time $\tilde\tau' := \tau'\circ((\omega,x)\mapsto\omega)$ on $(\tilde\Omega,\tilde{\mathcal{G}},(\tilde{\mathcal{G}}_t)_{t\geq 0})$ we conclude that τ must also be optimal among the stopping times on $(\Omega,\mathcal{G},(\mathcal{G}_t)_{t\geq 0})$. Let us denote our Brownian motion by B, to the optimal stopping time by τ and to our filtered probability space by $(\Omega,\mathcal{G},(\mathcal{G}_t)_{t\geq 0},\mathbb{P})$, irrespective of whether this is the original process and space we started with, or an enlarged one.

Choosing $p_0 := \frac{1}{2}+p+\varepsilon$ in Assumption 7.1.4.(e) we apply Theorem 7.1.18 to obtain a set Γ on which (B,τ) is concentrated under $\mathbb{P}$ and for which (7.1.19) holds. As ν is concentrated on $(0,\infty)$, we may assume that $\Gamma\cap(C(\mathbb{R}_+)\times\{0\}) = \emptyset$. Next we want to show that Lemma 7.1.20 applies with $Y_t(\omega) = \omega(t)$.

Translating (7.1.21) to our situation, we want to prove that $\omega(t) < \eta(t)$ implies

$$-\omega(t)A(t) - \mathbb{E}\Big[\Big(\eta(t)+\tilde B_\sigma\Big)A(\sigma)\Big] < -\eta(t)A(t) - \mathbb{E}\Big[\Big(\omega(t)+\tilde B_\sigma\Big)A(\sigma)\Big]\,, \tag{7.4.8}$$

where $\tilde{B}$ is Brownian motion started in 0 at time t on $C([t,\infty))$ and σ is any stopping time thereon with $\mathbb{W}_0^t(\sigma = t) < 1$, $\mathbb{W}_0^t(\sigma = \infty) = 0$, $\int \sigma^{p_0}\, d\mathbb{W}_0^t < \infty$. Again the Burkholder-Davis-Gundy inequality shows that $\mathbb{E}[\tilde{B}_\sigma A(\sigma)] < \infty$. So (7.4.8) turns into

$$\omega(t)\,\mathbb{E}[A(\sigma) - A(t)] < \eta(t)\,\mathbb{E}[A(\sigma) - A(t)],$$

which clearly follows from the assumptions. So we know that Lemma 7.1.20 holds, i.e., using the names from said lemma we have $\check{\tau} \le \tau \le \hat{\tau}$ $\mathbb{P}$-a.s.
$\Gamma \cap (C(\mathbb{R}_+) \times \{0\}) = \emptyset$ implies $\check{\mathcal{R}} \cap (\mathbb{R} \times \{0\}) = \emptyset$ and therefore $\check{\tau}(\omega) = \inf\{t > 0 : (B_t(\omega), t) \in \check{\mathcal{R}}\}$, and likewise for $\hat{\mathcal{R}}$ and $\hat{\tau}$. As $\overline{\check{\mathcal{R}}} = \overline{\hat{\mathcal{R}}} =: \overline{\mathcal{R}}$ it follows from Lemma 7.4.6 that $\check{\tau} = \tau = \hat{\tau}$ a.s. and that τ is of the form claimed in (7.4.2) with $\beta(t) := \sup\{y \in \mathbb{R} : (y,t) \in \overline{\mathcal{R}}\}$. The uniqueness claims follow from Lemma 7.4.7 and what we have already proven.
We now treat the case where $\psi(B_t, t) = \phi(B_t)$ with $\phi''' > 0$, $\left|\phi(y)\right| \le K(1 + |y|^p)$ and ν has finite moment of order $\frac{p}{2} + \varepsilon$ for some $\varepsilon > 0$. Most of the proof remains unchanged. Setting $c(\omega, t) = -\phi(\omega(t))$ we may again use the Burkholder-Davis-Gundy inequalities to show that $c(B_\tau, \tau)$ is bounded in $L^{1+\delta}$-norm, independently of the stopping time $\tau \sim \nu$, thereby showing both that Assumption 7.1.4.(d) is satisfied and that the boundedness-condition in Version B of Theorem 7.1.10 is satisfied.
It remains to show that $\omega(t) < \eta(t)$ implies $((\omega,t),(\eta,t)) \in \mathrm{SG}$. $\phi''' > 0$ implies that the map $y \mapsto \phi(\eta(t) + y) - \phi(\omega(t) + y)$ is strictly convex. By the strict Jensen inequality $\mathbb{E}[\phi(\eta(t) + \tilde{B}_\sigma) - \phi(\omega(t) + \tilde{B}_\sigma)] > \phi(\eta(t)) - \phi(\omega(t))$ for any stopping time σ on $C([t,\infty))$ which is almost surely finite, satisfies optional stopping and is not almost surely equal to t. As we may choose $p_0 := \frac{p}{2} + \varepsilon$, which is greater than 1, we may assume that the σ in the definition of SG has finite first moment, which is enough to guarantee that it satisfies optional stopping. Rearranging the last inequality gives (7.1.15). ❑

7.4.2. Supremum Process of Brownian Motion

To give an example of a slightly more complicated functional amenable to analysis with our tools consider the following problem.

Problem (OptStop$^{B_t^*}$)**.** Among all stopping times $\tau \sim \nu$ on $(\Omega, \mathcal{G}, (\mathcal{G}_t)_{t\ge 0}, \mathbb{P})$ find the maximizer of

$$\tau \mapsto \mathbb{E}[B_\tau^*]\,,$$

where $B_t^* = sup_{s \le t} B(s)$.

Then

Corollary 7.4.9. *Assume that ν has finite moment of order $\frac{3}{2}$. Then* OptStop$^{B_t^*}$ *has a solution τ given by*

$$\tau = \inf\{t > 0 : B_t - B_t^* \le \beta(t)\}$$

for some upper semicontinuous function $\beta : \mathbb{R}_+ \to [-\infty, 0]$.

We proceed to prove Corollary 7.4.9. This is closely modeled on the treatment of the Azema-Yor embedding in [9, Theorem 6.5]. As in this case we run into a technical obstacle, though one which can be overcome by combining the ideas we have already seen in slightly new ways. Thus the proof of Corollary 7.4.9 is very similar to the proof of Corollary 7.4.1. To demonstrate the problem let us begin an attempt to prove Corollary 7.4.9. Again, we read off $c(\omega,t) = -\omega^*(t)$, with $\omega^*(t) = \sup_{s\le t}\omega(s)$. We may use Theorem 7.1.10 to find a solution τ of the problem $\text{OptStop}^{B_t^*}$ and we use Theorem 7.1.18 to find a set $\Gamma \subseteq C(\mathbb{R}_+)\times\mathbb{R}_+$ for which $\mathbb{P}[(B,\tau)\in\Gamma] = 1$ and $\mathrm{SG}\cap(\Gamma^< \times \Gamma) = \emptyset$. Now we would like to apply Lemma 7.1.20 with $Y_t(\omega) = \omega(t) - \omega^*(t)$, as proposed by Corollary 7.4.9, so we want to prove that $\omega(t)-\omega^*(t) < \eta(t)-\eta^*(t)$ implies $((\omega,t),(\eta,t))\in\mathrm{SG}$.
Let us do the calculations: We start with an $(\mathbb{F}^t_s)_{s\ge t}$-stopping time σ, for which $\mathbb{W}^t_0(\sigma = t) < 1$, $\mathbb{W}^t_0(\sigma=\infty) = 0$ and for which both sides in (7.1.15) are defined and finite. To reduce clutter, let us name $(\omega \mapsto (\omega,\sigma(\omega)))_*(\mathbb{W}^t_0) =: \alpha$, so that (7.1.15), which we want to prove, reads

$$-\omega^*(t) + \int ((\omega,t)\odot\theta)^*(s)\,d\alpha(\theta,s) < -\eta^*(t) + \int ((\eta,t)\odot\theta)^*(s)\,d\alpha(\theta,s). \tag{7.4.10}$$

We may rewrite the left hand side as

$$\int\Big(\omega^*(t) \vee \big(\omega(t)+\theta^*(s)\big)\Big) - \omega^*(t)\,d\alpha(\theta,s) = \int 0 \vee \big(\omega(t)-\omega^*(t)+\theta^*(s)\big)\,d\alpha(\theta,s)\,.$$

For the right hand side we get the same expression with ω replaced by η. Looking at the integrands we see that if

$$0 < \eta(t) - \eta^*(t) + \theta^*(s) \tag{7.4.11}$$

then

$$0 \vee \big(\omega(t)-\omega^*(t)+\theta^*(s)\big) < 0 \vee \big(\eta(t)-\eta^*(t)+\theta^*(s)\big)\,,$$

but in the other case

$$0 \vee \big(\omega(t)-\omega^*(t)+\theta^*(s)\big) = 0 = 0 \vee \big(\eta(t)-\eta^*(t)+\theta^*(s)\big)\,.$$

So if (7.4.11) holds for (θ,s) from a set of positive α-measure, then we proved what we wanted to prove. But if $\theta^*(s) \le \eta^*(t)-\eta(t)$ for α-a.a. (θ,s) then in (7.1.15) we have equality instead of strict inequality.
As in [9, Theorem 6.5], one way of getting around this is to introduce a secondary optimization criterion. One way to explain the idea of secondary optimization is to think about what happens if, instead of considering a cost function $c : C(\mathbb{R}_+)\times\mathbb{R}_+ \to \mathbb{R}$ we consider a cost function $c : C(\mathbb{R}_+)\times\mathbb{R}_+ \to \mathbb{R}^n$. Of course, to be able to talk about optimization, we will want to have an order on $\mathbb{R}^n$. For reasons that should become clear soon, we decide on the lexicographical order. For the case $n = 2$ that we are actually interested in for Corollary 7.4.9 this means that

$$(x_1,x_2) \le (y_1,y_2) \iff x_1 < y_1 \text{ or } (x_1 = y_1 \text{ and } x_2 \le y_2)\,.$$

We claim that Theorem 7.1.18 is still true if we replace $c : C(\mathbb{R}_+)\times\mathbb{R}_+ \to \mathbb{R}$ by $c : C(\mathbb{R}_+)\times\mathbb{R}_+ \to \mathbb{R}^n$ and read any symbol $\le$ which appears between vectors in $\mathbb{R}^n$ as the lexicographic order on $\mathbb{R}^n$ (and of course likewise for all the derived symbols and notions $<$, $\geq$, $>$, inf, etc.). Moreover, the arguments are exactly the same. Indeed the crucial part that may deserve some mention is at the end of the proof of Proposition 7.3.14, where we use the assumption that (7.3.24) holds on a set of positive measure, i.e., that the integrand is < 0 on a set of positive measure, and that the integrand is 0 outside that set, to conclude that the integral itself must be < 0. This implication is also true for the lexicographical order on $\mathbb{R}^n$. One more detail to be aware of is that integrating functions which map into $\mathbb{R}^2$ may give results of the form (∞, x), $(x, -\infty)$, etc. In the case of a one-dimensional cost function we excluded such problems by making Assumption 7.1.4.(d). What we really want in the proof of Proposition 7.3.14 is that $\int c\, d\xi$ and $\int c\, d\xi^\pi$ should be finite. Clearly a sufficient condition to guarantee this is to replace Assumption 7.1.4.(d) by

(d') $\mathbb{E}[c(B,\tau)] \in \mathbb{R}^n$ for all stopping times $\tau \sim \nu$.

This is not the most general version possible but it will suffice for our purposes.
To get an existence result we may assume that $c = (c_1, c_2)$ is component-wise lower semi-continuous and that both c_1 and c_2 are bounded below (in either of the ways described in the two versions of Theorem 7.1.10). Note that – because we are talking about the lexicographic order – $\xi \in \mathsf{RST}_\lambda(\nu)$ is a solution of $\mathrm{O{\scriptstyle PT}S{\scriptstyle TOP}}^\pi$ for c if and only if ξ is a solution of $\mathrm{O{\scriptstyle PT}S{\scriptstyle TOP}}^\pi$ for c_1 and among all such solutions ξ', ξ minimizes $\int c_2\, d\xi'$. By Theorem 7.1.10 in the form that we have already proved the set of solutions of $\mathrm{O{\scriptstyle PT}S{\scriptstyle TOP}}^\pi$ for c_1 is non-empty. It is also a closed subset of a compact set and therefore itself compact. This allows us to reiterate the argument that we used in the proof of Theorem 7.1.10 to find inside this set a minimizer of $\xi' \mapsto \int c_2\, d\xi'$. This minimizer is the solution of $\mathrm{O{\scriptstyle PT}S{\scriptstyle TOP}}^\pi$ for c.

With this in hand we may pick up our

Proof of Corollary 7.4.9. The same arguments as in the proof of Corollary 7.4.1 apply, so we may assume that our probability space satisfies Assumption 7.1.4.(b). We start with a cost function $c(\omega, t) := (c_1(\omega, t), c_2(\omega, t)) := (-\omega^*(t), (\omega^*(t) - \omega(t))^3)$.

$$\|c_1(B,\tau)\|_{L^3} \leq \||B|^*_\tau\|_{L^3} \leq K_1 \|\tau\|^{1/2}_{L^{3/2}},$$

by the Burkholder-Davis-Gundy inequalities, so $(c_1)_-$ satisfies the uniform integrability condition and $\mathbb{E}[c(B,\tau)]$ is finite for all stopping times $\tau \sim \nu$. $c_2 \geq 0$ and by the Burkholder-Davis-Gundy inequalities

$$\mathbb{E}[c_2(B,\tau)] \leq \mathbb{E}[(B^*(\tau))^3] \leq K_1 \mathbb{E}[\tau^{3/2}] = K_1 \int t^{3/2}\, d\nu(t)$$

for some constant K_1. The last term is finite by assumption.
By our discussion in the preceding paragraphs we find a solution τ of $\mathrm{O{\scriptstyle PT}S{\scriptstyle TOP}}^\tau$ for c and a measurable, $(\mathcal{F}^0_t)_{t\geq 0}$-adapted set $\Gamma \subseteq C(\mathbb{R}_+) \times \mathbb{R}_+$, for which $\mathbb{P}[(B,\tau) \in \Gamma] = 1$ and $\mathsf{SG} \cap (\Gamma^< \times \Gamma) = \emptyset$, where now $((\omega, t), (\eta, t)) \in \mathsf{SG}$ if and only if for all $(\mathbb{F}^t_s)_{s\geq t}$-stopping times σ for which $\mathbb{W}^t_0(\sigma = t) < 1$, $\mathbb{W}^t_0(\sigma = \infty) = 0$, $\int \sigma^{3/2}\, d\mathbb{W}^t_0 < \infty$, setting $\alpha := (\omega \mapsto (\omega, \sigma(\omega)))_*(\mathbb{W}^t_0)$ we have that either equation (7.4.10) holds or

$$-\omega^*(t) + \int ((\omega,t) \odot \theta)^*(s)\, d\alpha(\theta, s) = -\eta^*(t) + \int ((\eta,t) \odot \theta)^*(s)\, d\alpha(\theta, s) \tag{7.4.12}$$

and

$$c_2(\omega,t) - \int c_2((\omega,t)\odot\theta,s)\,d\alpha(\theta,s) < c_2(\eta,t) - \int c_2((\eta,t)\odot\theta,s)\,d\alpha(\theta,s)\,. \tag{7.4.13}$$

Now, we want to apply Lemma 7.1.20, so we want to show that $\omega(t)-\omega^*(t) < \eta(t)-\eta^*(t)$ implies $((\omega,t),(\eta,t)) \in \mathrm{SG}$. We already dealt with the case where α is such that (7.4.11) holds on a set of positive α-measure. Now, we deal with the other case, so we have

$$\theta^*(s) \le \eta^*(t) - \eta(t) < \omega^*(t) - \omega(t) \tag{7.4.14}$$

for α-a.a. (θ,s) and we know that (7.4.12) holds. We show that (7.4.13) holds. Because of (7.4.14), $((\omega,t)\odot\theta)^*(s) = \omega^*(t)$, and so $c_2((\omega,t)\odot\theta,s) = (\omega^*(t)-\omega(t)-\theta(s))^3$. We calculate the left hand side of (7.4.13).

$$\begin{aligned}
&\int (\omega^*(t)-\omega(t))^3 - (\omega^*(t)-\omega(t)-\theta(s))^3\,d\alpha(\theta,s)\\
&\quad= \int 3(\omega^*(t)-\omega(t))^2\theta(s) - 3(\omega^*(t)-\omega(t))(\theta(s))^2 + (\theta(s))^3\,d\alpha(\theta,s)\\
&\quad= (\omega(t)-\omega^*(t))3\int(\theta(s))^2\,d\alpha(\theta,s) + \int(\theta(s))^3\,d\alpha(\theta,s).
\end{aligned}$$

Here the Burkholder-Davis-Gundy inequalities show that both $\int(\theta(s))^3\,d\alpha(\theta,s)$ and $\int(\theta(s))^2\,d\alpha(\theta,s)$ are finite so that we may split the integral and they also show that $\{\tilde{B}_{\sigma\wedge T} : T \ge t\}$ is uniformly integrable so that by the optional stopping theorem $\int\theta(s)\,d\alpha(\theta,s) = 0$. ($\tilde{B}$ is again a Brownian motion started in 0 at time t on $C([t,\infty))$.)
For the right hand side of (7.4.13) we get the same expression with ω replaced by η. This concludes the proof that $\omega(t)-\omega^*(t) < \eta(t)-\eta^*(t)$ implies $((\omega,t),(\eta,t)) \in \mathrm{SG}$ and Lemma 7.1.20 gives us barriers $\check{\mathcal{R}}$, $\hat{\mathcal{R}}$ such that for their hitting times $\check{\tau}$, $\hat{\tau}$ by $B_t - B_t^*$ we have $\check{\tau} \le \tau \le \hat{\tau}$, a.s.
Again we want to show that $\check{\tau} = \hat{\tau}$, a.s. and that they are actually stopping times. Again we do so by showing that they are both a.s. equal to the hitting time of the closure of the respective barrier. If $\bar{\check{\mathcal{R}}} \cap (\{0\}\times\mathbb{R}_+) = \emptyset$ then this works in exactly the same way as in Lemma 7.4.6. (This time we define $\bar{\tau}_\varepsilon := \inf\{t > 0 : (B_t^\varepsilon(\omega) - (B^\varepsilon)_t^*(\omega), t) \in \bar{\mathcal{R}}\}$ where $B_t^\varepsilon(\omega) := B_t(\omega) + A(t)\varepsilon$.) If $\bar{\check{\mathcal{R}}} \cap (\{0\}\times\mathbb{R}_+) \neq \emptyset$ then $(B_t^\varepsilon(\omega) - (B^\varepsilon)_t^*(\omega), t) \in \bar{\mathcal{R}}$ and $t > 0$ need not imply $B_t(\omega) - B_t^*(\omega) < B_t^\varepsilon(\omega) - (B^\varepsilon)_t^*(\omega)$, which is essential for the topological argument showing that the hitting time of $\mathcal{R}$ is less than or equal $\bar{\tau}_\varepsilon$.
But if $\bar{\hat{\mathcal{R}}} \cap (\{0\}\times\mathbb{R}_+) = \bar{\check{\mathcal{R}}} \cap (\{0\}\times\mathbb{R}_+) \neq \emptyset$, then $\check{\tau}$ and $\hat{\tau}$ are both almost surely $\le T$ where $T := \inf\{t > 0 : (0,t) \in \bar{\hat{\mathcal{R}}}\}$, so in the step where we show that the hitting time of $\mathcal{R}$ is less than $\bar{\tau}_\varepsilon$ we can argue under the assumption that $\bar{\tau}_\varepsilon(\omega) < T$. In this case we do have that $(B_t^\varepsilon(\omega) - (B^\varepsilon)_t^*(\omega), t) \in \bar{\mathcal{R}}$ and $t > 0$ implies $B_t(\omega) - B_t^*(\omega) < B_t^\varepsilon(\omega) - (B^\varepsilon)_t^*(\omega)$. ❑

Remark 7.4.15. We hope that the proofs of Corollary 7.4.1 and Corollary 7.4.9 have given the reader some idea of how to apply the main results of this section to arrive at barrier-type solutions of constrained optimal stopping problems, as depicted in Figure 7.1. We would like to conclude by giving a couple of pointers to the interested reader who may want to work through the proofs corresponding to the remaining pictures in Figure 7.1. For the problem of minimizing $\mathbb{E}[B_{\tau^*}]$, it may actually happen that the times $\check{\tau}$, $\hat{\tau}$ from Lemma 7.1.20 do not coincide. Specifically one has to expect this to happen on non-negligible set when $\check{\mathcal{R}}$ contains parts of the time axis which $\hat{\mathcal{R}}$ does not contain.

Under these circumstances an optimizer may turn out to be a true randomized stopping time, with a proportion of a path hitting the time axis at a certain point needing to be stopped while the rest continues. In this situation the picture alone does not completely describe the optimal stopping time.
For the problems involving absolute values one needs to make a minor modification in the proof of Proposition 7.3.14. Specifically one can allow “mirroring” the paths which are “transplanted” using the Gardener’s Lemma. This leads to a slightly different definition of Stop-Go pairs, which is perhaps most easily described by saying that the paths which are stopped by σ may be flipped upside-down on either side.

Appendix

A

Appendix

Finally, some detailed considerations are given, which are used in the main part of the thesis, but would have had disturbed the reading flow there. The material can therefore be skipped and considered when needed. The following topics are examined: Dominance in stochastic order, some general results, expected shortfall, multi-dimensional log-normal distribution and transition kernels.

A.1. Dominance in Stochastic Order

Dominance in stochastic order is an important tool in many areas of probability and statistics. Stochastic orders generated by integrals are considered in [52]. We are guided by [72] and [53] to give a definition of stochastic order with regard to random variables and probability measures in our considerations. We will transfer it to stochastic order for signed measures. It finds application in the Section 3.5.2 of Part I and the Section 6.2 of Part II.
Let I be the considered index set or a continuous time interval.

Definition A.1.1 (Dominance in stochastic order)**.** See [72, p. 3]:
Let X and Y be two random variables such that

$$\mathbb{P}(X > t) \leq \mathbb{P}(Y > t) \quad \text{for all } t \in I. \tag{A.1.2}$$

Then X is said to be smaller than Y in the usual stochastic order or Y dominates X in stochastic order (denoted by $X \leq_{st} Y$).

Remark A.1.3. Note that (A.1.2) is the same as

$$\mathbb{P}(X \leq t) \geq \mathbb{P}(Y \leq t) \quad \text{for all } t \in I.$$

Lemma A.1.4. *See [72, p. 4]:*
$X \leq_{st} Y$ if, and only if,

$$\mathbb{E}[f(X)] \leq \mathbb{E}[f(Y)] \tag{A.1.5}$$

holds for all increasing functions f for which the expectations exist.

Remark A.1.6.

(a) Let I be a continuous time interval. Note that $X \leq_{st} Y$ if, and only if,

$$\int_{I_{>s}} \mathbb{P}(Y > u)\,du - \int_{I_{>s}} \mathbb{P}(X > u)\,du \quad \text{is decreasing in } s \in I. \tag{A.1.7}$$

(b) Let I be a totally ordered countable index set. If X and Y are random variables taking on values in I, then we have the following. Let $p_t = \mathbb{P}(X = t)$ and $q_t = \mathbb{P}(Y = t)$, $t \in I$. Then $X \leq_{st} Y$ if, and only if,

$$\sum_{t \in I_{\leq s}} p_t \geq \sum_{t \in I_{\leq s}} q_t, \quad \text{for all } s \in I, \tag{A.1.8}$$

or, equivalently $X \leq_{st} Y$ if, and only if,

$$\sum_{t \in I_{>s}} p_t \leq \sum_{t \in I_{>s}} q_t, \quad \text{for all } s \in I. \tag{A.1.9}$$

In continuous time the order relation for probability distributions are often immediately given in the following way, cf. [46] and [52, Example 5.1]. This definition is equivalent to the definition of the usual stochastic order given in Definition A.1.1.

Definition A.1.10. Let E be a Polish space endowed with a closed partial order. For probability measures $\mathbb{P}$ and $\mathbb{Q}$ on $(E, \mathcal{E})$, $\mathbb{P}$ is stochastically dominated by $\mathbb{Q}$ if

$$\int f\,d\mathbb{P} \leq \int f\,d\mathbb{Q} \quad \text{for all measurable bounded increasing functions } f.$$

We write that $\mathbb{P} \leq_{st} \mathbb{Q}$.

In the literature there are definitions of stochastic order with regard to random variables and probability measures. Now, we want to introduce it for signed measures.

Definition A.1.11 (dominance in first order; signed measure)**.**
Let I be a totally ordered countable index set or a continuous time interval. Let μ^* and μ be two signed measures of finite total variation with $\mu(I) = \mu^*(I)$. Then μ^* dominates μ in first order, (denoted by $\mu \leq_{st} \mu^*$), if

$$\mu(I_{>s}) \leq \mu^*(I_{>s}), \quad \text{for all } s \in I. \tag{A.1.12}$$

Remark A.1.13. (a) Consider it as a function of s, i.e., $s \mapsto \mu(I_{>s})$, such that we want to prove if one function lies always above the other one.

(b) Note that (A.1.12) is equivalent to

$$\mu(I_{\leq s}) \geq \mu^*(I_{\leq s}), \quad \text{for all } s \in I.$$

(c) Let I be a totally ordered countable index set. For any signed measure μ we define $\mu(\{t\}) := \mu_t$, $t \in I$, and we have that $\mu(I_{>s}) = \sum_{t \in I_{>s}} \mu_t$.

(d) Let I be a continuous time interval. For a signed measure μ we denote the positive and negative variation by μ^+ and μ^-, cf. [52]. As usual $|\mu| := \mu^+ + \mu^-$ is the total variation. Integrals are mostly written in the functional form $\mu(f) := \int f \, d\mu := \int f \, d\mu^+ - \int f \, d\mu^-$. Notice that $\mu(f)$ exists and is finite if and only if $\mu^+(|f|) + \mu^-(|f|) < \infty$.

Remark A.1.14. Note that there are signed measures without finite total variation, for example signed measures which are induced by the covariation of two local martingales.

Lemma A.1.15. *Let I be a totally ordered countable index set and μ^*, μ be two signed measures of finite total variation with $\mu(I) = \mu^*(I)$. Then $\mu \leq_{st} \mu^*$ if, and only if,*

$$\sum_{t \in I} f(t)\mu_t \leq \sum_{t \in I} f(t)\mu_t^* \tag{A.1.16}$$

holds for all increasing functions $f : I \to \mathbb{R}$ for which the expectations exist.

Proof. "$\Leftarrow$" For all increasing functions $f : I \to \mathbb{R}$ for which the expectations exist, it holds that

$$\sum_{t \in I} f(t)\mu_t \leq \sum_{t \in I} f(t)\mu_t^*.$$

Choosing $f_s := \mathbb{1}_{I_{>s}}$ we have for μ that $\sum_{t \in I} f_s(t)\mu_t = \sum_{t \in I_{>s}} \mu_t = \mu(I_{>s})$. Analogously we get the same for μ^*. Therefore for all $s \in I$ it follows from (A.1.16) that

$$\mu(I_{>s}) \leq \mu^*(I_{>s}).$$

"$\Rightarrow$" We will decompose a signed measure μ into its positive μ^+ and negative part μ^-, whereby $\mu_t^+ := \max\{\mu_t, 0\}$ and $\mu_t^- := \max\{-\mu_t, 0\}$ for every $t \in I$. Then we have $\mu_t = \mu_t^+ - \mu_t^-$ for every $t \in I$. Now we have that μ^* dominates μ in first order, i.e.,

$$\mu^*(I_{>s}) \geq \mu(I_{>s}), \quad \forall s \in I. \tag{A.1.17}$$

In the following we will assume that $\mu(I) + \mu^-(I) + (\mu^*)^-(I) > 0$, because the case $\mu(I) + \mu^-(I) + (\mu^*)^-(I) = 0$ is trivial. If $\mu(I) + \mu^-(I) + (\mu^*)^-(I) = 0$ then everything is null, i.e., $\mu(I) = 0$, $\mu^-(I) = 0$ and $(\mu^*)^-(I) = 0$. It follows that $\mu^+(I) = \mu(I) + \mu^-(I) = 0$ and also $(\mu^*)^+(I) = 0$, because $\mu^*(I) = \mu(I) = 0$. This also holds for every subset of I.
Adding $\mu^-(I_{>s}) + (\mu^*)^-(I_{>s})$ to (A.1.17) and multiplying with $\frac{1}{\mu(I)+\mu^-(I)+(\mu^*)^-(I)}$ we get for all $s \in I$

$$\begin{aligned} \hat{\mu} &:= \frac{1}{\mu(I) + \mu^-(I) + (\mu^*)^-(I)} \Big(\mu^*(I_{>s}) + \mu^-(I_{>s}) + (\mu^*)^-(I_{>s}) \Big) \\ &\geq \frac{1}{\mu(I) + \mu^-(I) + (\mu^*)^-(I)} \Big(\mu(I_{>s}) + \mu^-(I_{>s}) + (\mu^*)^-(I_{>s}) \Big) =: \tilde{\mu}. \end{aligned}$$

With $\mu + \mu^- + (\mu^*)^- = \mu^+ + (\mu^*)^-$ it is clear that $\mu + \mu^- + (\mu^*)^-$ is a finite, non-negative measure. Through the appropriate scaling we have that $\tilde{\mu}$ is a probability measure. The same is true for $\hat{\mu}$. Furthermore we have that $\hat{\mu}$ dominates $\tilde{\mu}$ in stochastic order.

Now, let X^* and X be random variables taking values in I such that $\mathbb{P}(X^* = t) = \hat{\mu}_t$ and $\mathbb{P}(X = t) = \tilde{\mu}_t$ for every $t \in I$. Due to Lemma A.1.4 we get

$$\sum_{t\in I} f(t)\hat{\mu}_t \geq \sum_{t\in I} f(t)\tilde{\mu}_t$$

for all increasing functions $f : I \to \mathbb{R}$ for which the expectations exist. Multiply the inequality with $\mu(I) + \mu^-(I) + (\mu^*)^-(I)$ and subtract $\sum_{t\in I} f(t)(\mu^-(I_{>s}) + (\mu^*)^-(I_{>s}))$ gives us the assertion that

$$\sum_{t\in I} f(t)\mu^*_t \geq \sum_{t\in I} f(t)\mu_t,$$

for all increasing functions $f : I \to \mathbb{R}$ for which the expectations exist. ❑

Lemma A.1.18. *Let I be a continuous time interval and μ^*, μ be two signed measures of finite total variation with $\mu(I) = \mu^*(I)$. Then $\mu \leq_{st} \mu^*$ if, and only if,*

$$\mu(f) \leq \mu^*(f) \tag{A.1.19}$$

holds for all measurable bounded increasing functions $f : I \to \mathbb{R}$ for which the integrals exist.

Proof. We get it analogously like in the proof of Lemma A.1.15, if we replace the sums by the integrals.

"⇐" For all measurable bounded increasing functions $f : I \to \mathbb{R}$ for which the expectations exist, it holds that

$$\int f\,d\mu \leq \int f\,d\mu^*.$$

Choosing $f_s := \mathbb{1}_{I_{>s}}$, which is bounded by 1, increasing and measurable, we have for μ that $\int f_s d\mu = \int_{I_{>s}} d\mu = \mu(I_{>s})$. Analogously we get the same for μ^*. Therefore for all $s \in I$ it follows from (A.1.19) that

$$\mu(I_{>s}) \leq \mu^*(I_{>s}).$$

"⇒" We will decompose a signed measure μ into its positive μ^+ and negative part μ^-, whereby $\mu_t^+ := \max\{\mu_t, 0\}$ and $\mu_t^- := \max\{-\mu_t, 0\}$ for every $t \in I$. Then we have $\mu_t = \mu_t^+ - \mu_t^-$ for every $t \in I$. Now we have that μ^* dominates μ in first order, i.e.,

$$\mu^*(I_{>s}) \geq \mu(I_{>s}), \quad \forall s \in I. \tag{A.1.20}$$

In the following we will assume that $\mu(I) + \mu^-(I) + (\mu^*)^-(I) > 0$, because the case $\mu(I) + \mu^-(I) + (\mu^*)^-(I) = 0$ is trivial. If $\mu(I) + \mu^-(I) + (\mu^*)^-(I) = 0$ then everything is null, i.e., $\mu(I) = 0$, $\mu^-(I) = 0$ and $(\mu^*)^-(I) = 0$. It follows that $\mu^+(I) = \mu(I) + \mu^-(I) = 0$ and also $(\mu^*)^+(I) = 0$, because $\mu^*(I) = \mu(I) = 0$. This also holds for every subset of I.

Adding $\mu^-(I_{>s}) + (\mu^*)^-(I_{>s})$ to (A.1.20) and multiplying with $\frac{1}{\mu(I)+\mu^-(I)+(\mu^*)^-(I)}$ we get for all $s \in I$

$$\begin{aligned}\hat{\mu} &:= \frac{1}{\mu(I) + \mu^-(I) + (\mu^*)^-(I)}\Big(\mu^*(I_{>s}) + \mu^-(I_{>s}) + (\mu^*)^-(I_{>s})\Big) \\ &\geq \frac{1}{\mu(I) + \mu^-(I) + (\mu^*)^-(I)}\Big(\mu(I_{>s}) + \mu^-(I_{>s}) + (\mu^*)^-(I_{>s})\Big) =: \tilde{\mu}.\end{aligned}$$

With $\mu + \mu^- + (\mu^*)^- = \mu^+ + (\mu^*)^-$ it is clear that $\mu + \mu^- + (\mu^*)^-$ is a finite, non-negative measure. Through the appropriate scaling we have that $\tilde{\mu}$ is a probability measure. The same is true for $\hat{\mu}$. Furthermore we have that $\hat{\mu}$ dominates $\tilde{\mu}$ in stochastic order and we can also use the equivalent definition. Due to Definition A.1.10 we have

$$\int f \, d\hat{\mu} \geq \int f \, d\tilde{\mu}$$

for all measurable bounded increasing functions $f : I \to \mathbb{R}$ for which the integrals exist. Multiply the inequality with $\mu(I)+\mu^-(I)+(\mu^*)^-(I)$ and subtract $\int_{I_{>s}} f \, d(\mu^-+(\mu^*)^-)$ gives us the assertion that

$$\mu(f) \leq \mu^*(f)$$

for all measurable bounded increasing functions $f : I \to \mathbb{R}$ for which the integrals exist.

Note that we could also discuss it about the existence of two random variables X^* and X taking values in I such that $\mathbb{P}(X^* = t) = \hat{\mu}_t$ and $\mathbb{P}(X = t) = \tilde{\mu}_t$ for every $t \in I$ like in the proof of Lemma A.1.15 using Lemma A.1.4.

❑

A.2. Some General Results

First, we want to show the Jensen's inequality for substochastic measures.

Lemma A.2.1 (Jensen's inequality for substochastic measures)**.** *Let* $(\Omega, \mathcal{F}, \mu)$ *be a subprobability space, such that* $\mu(\Omega) \in [0,1)$. *If* g *is a real-valued function that is* μ*-integrable, and if* φ *is a convex function on the real line with* $\varphi(0) \leq 0$, *then:*

$$\varphi\left(\int_\Omega g\, d\mu\right) \leq \int_\Omega \varphi \circ g\, d\mu.$$

Proof. Let $\mu(\Omega) \neq 0$. We define $\tilde{\mu} = \frac{\mu}{\mu(\Omega)}$ which is obviously a probability measure. Then

$$\begin{aligned}
&\varphi\left(\int_\Omega g\, d\mu\right) = \varphi\left(\mu(\Omega)\int_\Omega g\, d\tilde{\mu} + (1-\mu(\Omega))\cdot 0\right)\\
&\leq \mu(\Omega)\cdot\varphi\left(\int_\Omega g\, d\tilde{\mu}\right) + (1-\mu(\Omega))\cdot\varphi(0) \quad \text{(convexity)}\\
&\leq \mu(\Omega)\cdot\frac{1}{\mu(\Omega)}\int_\Omega \varphi(g)\, d\mu + (1-\mu(\Omega))\cdot\varphi(0) \quad \text{(Jensen)}\\
&\leq \int_\Omega \varphi(g)\, d\mu. \quad (\varphi(0)\leq 0)
\end{aligned}$$

❑

Let $I \neq \emptyset$ denote a countable, i.e., a finite or countably infinite, totally-ordered index set, $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in I}$ and ν a given probability measure on I. In Section 3.2 we will extend the given filtration $\mathbb{F}$ in an eligible way to embed a special set, the set $\mathcal{M}_I^\nu$ of all adapted random probability measures γ with $\mathbb{E}[\gamma_t] = \nu_t$ for all $t \in I$, into another set, the set $\mathcal{T}_I^\nu$ of all stopping times τ with distribution ν. To have an unique assignment, one of the following assumptions should demand:

Assumption A.2.2. We assume one of the conditions:

(a) $\mathcal{F}_t$ includes all null sets of $\mathcal{F}_\infty := \sigma(\bigcup_{t\in I}\mathcal{F}_t)$ for all $t \in I$.

(b) There exists a sequence $(t_n)_{n\in\mathbb{N}}$ in I such that $t_n \leq t_{n+1}$ for all $n \in \mathbb{N}$ and I includes a maximum element.

By virtue of definition of an adapted random probability measure, it is necessary, in particular because of Defintion 3.1.1(b), otherwise would be the constructed stopping time not unique defined on the corresponding null sets. In the case of a finite index set no additional assumption is necessary.

Remark A.2.3. By condition (a) of Assumption A.2.2, the null sets would be identified. The condition (b) of Assumption A.2.2 gives us the possibility to redefine the random adapted probability measure.

If the Assumption A.2.2 does not hold, we have to enlarge the filtration.

Definition A.2.4. A filtration $\mathbb{G} = (\mathcal{G}_t)_{t\in I}$ of $\mathbb{F}$ is called an enlargement of the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in I}$, if $\mathcal{F}_t \subseteq \mathcal{G}_t$ for all $t \in I$.

Extending the filtration by adding null sets is easier to describe for an interval and can be done as follows:

Remark A.2.5 (Adding null sets to filtrations). Let $I \subseteq \mathbb{R} \cup \{\infty, -\infty\}$ be an interval and $\mathbb{F} = (\mathcal{F}_t)_{t \in I}$ a filtration. Define $\mathbb{G} = (\mathcal{G}_t)_{t \in I}$ by

$$\mathcal{G}_t = \left\{ G \in \mathcal{F} \mid \text{There exists } F \in \mathcal{F}_t \text{ with } \mathbb{P}(F \Delta G) = 0 \right\}, \quad t \in I.$$

Then the following holds:

(a) $\mathbb{G}$ is an enlargement of $\mathbb{F}$ such that $\mathcal{F}_t \subseteq \mathcal{G}_t$ for all $t \in I$ and its σ-algebras contain all null sets of $\mathcal{F}$.

(b) If $\mathbb{F}$ is right-continuous, then $\mathbb{G}$ is right-continuous.

(c) Every $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ satisfies $\mathbb{E}[X|\mathcal{F}_t] \overset{\text{a.s.}}{=} \mathbb{E}[X|\mathcal{G}_t]$ for all $t \in I$.

(d) Let $\mathcal{H} \subseteq \mathcal{F}$ be a σ-algebra and $t \in I$. If $\mathcal{H}$ is independent of $\mathcal{F}_t$, then it is also independent of $\mathcal{G}_t$.

This can easily be verified.

We still want to specify the definition of progressive measurable processes here.

Definition A.2.6 (Progressively measurable)**.** Let $I \subseteq \mathbb{R}$ be a continuous time interval and $I_{\leq t} = (-\infty, t] \cap I$ be the set of all times up to $t \in I$. A stochastic process X defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \in I}, \mathbb{P})$ is $\mathbb{F}$-progressively measurable or simply $\mathbb{F}$-progressive with respect to a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in I}$, if the function $X(s, \omega) : I_{\leq t} \times \Omega \to \mathbb{R}$ is $\mathcal{B}_{I_{\leq t}} \otimes \mathcal{F}_t$-measurable for every $t \in I$.

The section should be completed with a proposition on conditional expectations involving independent random variables. The following proposition is a modified version of [25, Example 5.1.5]:

Proposition A.2.7 (Conditional expectation involving independent random variables)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G} \subseteq \mathcal{F}$ a sub-σ-algebra, $(S_1, \mathcal{S}_1)$ and $(S_2, \mathcal{S}_2)$ measurable spaces, $X : \Omega \to S_1$ and $Y : \Omega \to S_2$ random variables. Suppose X and Y are independent. Let $F : S_1 \times S_2 \to \mathbb{R}$ be an $\mathcal{S}_1 \otimes \mathcal{S}_2$-measurable function with $\mathbb{E}[|F(X,Y)|] < \infty$ and let $h(x) = \mathbb{E}[F(x, Y)]$. Then we have that*

$$\mathbb{E}[F(X,Y)|\sigma(X)] = h(X).$$

Proof. It is clear that $h(X) \in \sigma(X)$. We have to check that for every $A \in \sigma(X)$

$$\int_A F(X,Y)\, d\mathbb{P} = \int_A h(X)\, d\mathbb{P}.$$

Note that if $A \in \sigma(X)$ then there exist a $C \in \mathcal{B}(\mathbb{R})$ with $A = X^{-1}(C)$. Then also the preimage of C under (X, Y) is the event A.
Using the change of variables formula ([25, Theorem 1.6.9]) and the fact that the distribution $\mathbb{P}_{(X,Y)}$ of (X, Y) is a product measure ([25, Theorem 2.1.7]), then the definition of h,

and change of variables again, yields

$$\begin{aligned}\int_A F(X,Y)\,d\mathbb{P} &= \mathbb{E}[F(X,Y)\mathbb{1}_C(X)] = \int_{C\times\mathbb{R}} F(x,y)\,d\mathbb{P}_{(X,Y)}(x,y) \\ &= \int_C \int_{\mathbb{R}} F(x,y)\,d\mathbb{P}_Y(y)\,d\mathbb{P}_X(x) = \int_C h(x)\,d\mathbb{P}_X(x) = \int_A h(X)\,d\mathbb{P},\end{aligned}$$

which proves the desired result. ❑

A.3. Notes on the Expected Shortfall

In this section we will return the definitions of quantiles and the expected shortfall, given in [33], to mind. Furthermore we will use in Section 3.5.2 and 3.6 of Part I the result of Lemma A.3.3 to derive upper bounds which differ from these given in [33].

Definition A.3.1. Given a random variable $X : \Omega \to \mathbb{R}$, $\delta \in [0,1]$.

(a) Define the δ-quantile of X by

$$q_\delta(X) := \inf\{x \in \mathbb{R} | \mathbb{P}(X \leq x) \geq \delta\}.$$

Note that $q_0(X) = -\infty$ and if $\mathbb{P}(X \leq x) < 1$ for all $x \in \mathbb{R}$, then $q_1(X) = \infty$.

(b) Define $f_{\delta,X} : \Omega \to [0,1]$ by

$$f_{\delta,X} := \begin{cases} 0 & \text{if } \delta = 1, \\ \mathbb{1}_{X > q_\delta(X)} + \beta_{\delta,X}\mathbb{1}_{X = q_\delta(X)} & \text{if } \delta \in [0,1), \end{cases}$$

where

$$\beta_{\delta,X} := \begin{cases} \frac{\mathbb{P}(X \leq q_\delta(X)) - \delta}{\mathbb{P}(X = q_\delta(X))} & \text{if } \mathbb{P}(X = q_\delta(X)) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

(c) The expected shortfall of X at level δ is given by

$$(1-\delta)\,\mathrm{ES}[X;\delta] = \mathbb{E}[f_{\delta,X}X].$$

Note that $\mathrm{ES}[X;0] = \mathbb{E}[X]$ and $\mathrm{ES}[X;1] = 0$ as $\beta_{1,X} = f_{1,X} = 0$.

Remark A.3.2.

(a) Note that $\beta_{\delta,X} \in [0,1]$, because $\mathbb{P}(X < q_\delta(X)) \leq \delta \leq \mathbb{P}(X \leq q_\delta(X))$. Therefore $f_{\delta,X}$ is $[0,1]$-valued.

(b) For $\delta \in [0,1]$ we have $\mathbb{E}[f_{\delta,X}] = 1-\delta$. This is due to the fact that for $\delta \in [0,1)$

$$\mathbb{E}[f_{\delta,X}] = \mathbb{P}(X > q_\delta(X)) + \beta_{\delta,X}\mathbb{P}(X = q_\delta(X)).$$

Lemma A.3.3. *Cf. [33, Lemma 4.6]:*
Let X and Y be real-valued random variables, $\delta \in [0,1]$. Assume $Y \geq 0$, $\mathbb{E}[Y] < \infty$ and $\mathbb{E}[|X|] < \infty$. Define

$$\mathcal{F}^Y_{\delta,X} := \left\{ f : \Omega \to [0,1] \,\middle|\, f \textit{ measurable }, \mathbb{E}[fY] = \mathbb{E}[f_{\delta,X}Y] \right\}.$$

Then the following holds:

(a) $\mathbb{E}[f_{\delta,X}XY]$ *is well-defined and*

$$\sup_{f \in \mathcal{F}^Y_{\delta,X}} \mathbb{E}[fXY] = \mathbb{E}[f_{\delta,X}XY].$$

(b) *If* $f^* \in \mathcal{F}^Y_{\delta,X}$ *satisfies* $\mathbb{E}[f^*XY] = \mathbb{E}[f_{\delta,X}XY] < \infty$, *then*

$$f^* = f_{\delta,X}, \quad \text{a.s., on } \{Y > 0,\, X \neq q_\delta(Y)\}.$$

(c) *If* Y *and* $f_{\delta,X}X$ *are uncorrelated, then*

$$\mathbb{E}[f_{\delta,X}XY] = \mathbb{E}[Y]\mathbb{E}[f_{\delta,X}X] = (1-\delta)\mathbb{E}[Y]\,\mathrm{ES}[X;\delta].$$

Remark A.3.4. Note that Y and $f_{\delta,X}X$ are uncorrelated, if X and Y are independent.

At this point we would like to reproduce a few important results of [33].

Lemma A.3.5. *See [33, Lemma 5.10]:*
Given a discrete time interval $I \subseteq \mathbb{N}_0$ *with* $0 \in I$ *and a probability distribution* ν *on* I. *For a given adapted stochastic process* Z *we define the increments of* Z *by* $\Delta Z_0 := Z_0$ *and* $\Delta Z_t := Z_t - Z_{t-1}$ *for all* $t \in I \setminus \{0\}$. *Assume the increments are integrable and there exists a sequence* $(c_t)_{t\in I} \subseteq [1,\infty)$ *such that they satisfy*

$$\mathbb{E}[\Delta Z_t|\mathcal{F}_{t-1}] = \mathbb{E}[\Delta Z_t], \quad a.s.,$$

and

$$\mathbb{E}[|\Delta Z_t||\mathcal{F}_{t-1}] \leq c_t\mathbb{E}[|\Delta Z_t|], \quad a.s.,$$

for all $t \in I \setminus \{0\}$ *with* $\nu_{leqt-1} < 1$, *as well as*

$$\sum_{t\in I} c_t\mathbb{E}[|\Delta Z_t|](1-\nu_{\leq t-1}) < \infty$$

with the understanding that $1-\nu_{\leq t-1} = 1$ *for* $t = 0$. *Then, for all* $\gamma \in \mathcal{M}^\nu_I$, Z_γ *is well-defined, integrable and*

$$\mathbb{E}[Z_\gamma] = \sum_{t\in I} E[Z_t]\nu_t.$$

Theorem A.3.6. *See [33, Theorem 5.25]:*
Given a discrete time interval $I \subseteq \mathbb{N}_0$ *with* $0 \in I$ *and a probability distribution* ν *on* I, *assume that the adapted process* $Z = (Z_t)_{t\in I}$ *can be decomposed into* $Z_t = M_t + N_t + A_t$ *for* $t \in I$, *where* M *is a martingale such that* M *and* ν *satisfy one of the conditions of [33, Theorem 2.49],* N *is a process such that* N *and* ν *satisfy the conditions of [33, Lemma 5.10] and* A *is a predictable process, with* $A_0 = 0$. *Denote the increments of the process* A *by* $\Delta A_0 = A_0 = 0$ *and* $\Delta A_t := A_t - A_{t-1}$ *for* $t \in I \setminus \{0\}$. *Assume that for the density of [33, Definition 4.3(b)] we have for every* $t \in I \setminus \{0\}$ *with* $t+1 \in I$

$$\mathbb{E}[f_{\delta_t,\Delta A_{t+1}}|\mathcal{F}_{t-1}] = 1-\delta_t, \quad a.s.$$

and that for some sequence $(c_t)_{t\in I}$ *for every* $t \in I \setminus \{0\}$

$$\mathbb{E}[|\Delta A_t||\mathcal{F}_{t-1}] \leq c_t\mathbb{E}[|\Delta A_t|], \quad a.s.$$

and that for each $t \in I \setminus \{0,1\}$ we have that $f_{\delta_{t-1},\Delta A_t}\Delta A_t$ and $(1-\gamma_0-\ldots-\gamma_{t-2})$ are uncorrelated. Further assume that the process A satisfies either

$$\sum_{t\in I\setminus\{0\}} c_t \mathbb{E}[|\Delta A_t|]\nu_{\geq t} < \infty$$

or

$$\mathbb{E}[\sup_{t\in I}|A_t|] < \infty.$$

With these assumptions we have that Z_γ is well-defined and integrable. Then an optimal adapted random probability measure γ^* is given by

$$\gamma_t^* = \begin{cases} (1-\gamma_{\leq t-1}^*)(1-f_{\delta_t,\Delta A_{t+1}}) & \text{for } t+1 \in I, \\ (1-\gamma_{\leq t-1}^*) & \text{for } t+1 \notin I, \end{cases}$$

where $f_{\delta_t,\Delta A_{t+1}}$ is defined as in [33, Definition 4.3(b)] and

$$\delta_t = \begin{cases} \frac{\nu_t}{1-\nu_{\leq t-1}} & \text{if } \nu_{\leq t-1} < 1, \\ 0 & \text{if } \nu_{\leq t-1} = 1. \end{cases}$$

Using this strategy we have

$$V_{\mathcal{M}}^{\nu}(Z) = \mathbb{E}[M_0] + \sum_{t\in I}\mathbb{E}[N_t]\nu_t + \sum_{t\in I\setminus\{0\}}(1-\nu_{\leq t-1})\mathrm{ES}[\Delta A_t;\delta_{t-1}].$$

If $\mathbb{P}(\Delta A_{t+1} = q_{\delta_t}(\Delta A_{t+1})) = 0$ for all $t \in I$ with $t+1 \in I$, then the optimal strategy γ^* is a.s. unique.

Lemma A.3.7. *See [33, Lemma 5.37]:*
Given a discrete time interval $I \subseteq \mathbb{N}_0$ with $0 \in I$ and a probability distribution ν on I. Assume that the adapted process $Z = (Z_t)_{t\in I}$ is a process of independent random variables such that $\mathbb{E}[\sup_{t\in I}|Z_t|] < \infty$. Further let $U = (U_t)_{t\in I}$ be an adapted process of independent random variables uniformly distributed on $[0,1]$, such that Z and U are independent. For $t \in I$ set $\delta_t = 1 - \frac{\nu_t}{1-\nu_0-\ldots-\nu_{t-1}}$ and

$$E_t := \{Z_t > q_{\delta_t}(Z_t)\} \cup \{Z_t = q_{\delta_t}(Z_t), 1-\beta_{\delta_t,Z_t} < U_t \leq 1\}$$

with β_{δ_t,Z_t} as in Definition A.3.1. Then an optimal stopping time τ^* solving

$$\sup_{\tau\in T_I^{\nu}} \mathbb{E}[Z_\tau] = \mathbb{E}[Z_{\tau^*}]$$

is given by

$$\{\tau^* = t\} = \begin{cases} \bigcap_{I_{<t}} E_s^c \cap E_t & \text{if } \nu_{\leq t} < 1, \\ \bigcap_{I_{<t}} E_s^c & \text{if } \nu_{\leq t} = 1. \end{cases}$$

$\mathbb{E}[Z_{\tau^*}]$ can be computed as

$$\mathbb{E}[Z_{\tau^*}] = \sum_{t\in I}\mathbb{E}[Z_t \mathbb{1}_{\tau^*=t}] = \sum_{t=0}^{T}\nu_t \cdot \mathrm{ES}[Z_t;\delta_t].$$

A.4. Multi-Dimensional Log-Normal Distribution

A log-normal distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. Therefore we will define at first the multi-dimensional normal distribution.

Definition A.4.1. See [71, Definition 2.6]: Let $\mu \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times d}$ and $X \sim \mathcal{N}(0, I_d)$. Then the distribution of $Y := AX + \mu$ is called a n-dimensional normal distribution with expected value μ and covariance matrix $C := AA^T$. We use the notation $Y \sim \mathcal{N}(\mu, C)$ and $\mathcal{L}(Y) = \mathcal{N}(\mu, C)$. If $\mu = 0$, then the normal distribution is called centered.

Proposition A.4.2. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be affine, i.e., $f(y) = By + \nu$ with $\nu \in \mathbb{R}^m$ and $(m \times n)$-matrix B. If Y is an $\mathbb{R}^n$-valued random variable with $Y \sim \mathcal{N}(\mu, C)$, then*

$$f(Y) \sim \mathcal{N}(B\mu + \nu, BCB^T).$$

Proof. By Definition A.4.1 there exist a $\mu \in \mathbb{R}^n$, a dimension $d \in \mathbb{N}$ as well as $A \in \mathbb{R}^{n\times d}$ with $C = AA^T$ and $X \sim \mathcal{N}(0, I_d)$ such that $Y := AX + \mu \sim \mathcal{N}(\mu, C)$. Then $f(Y) = B(AX + \mu) + \nu = (BA)X + (B\mu + \nu)$ and $(BA)(BA)^T = BAA^TB^T = BCB^T$ which prove the claim. ❑

Proposition A.4.3. *Let $X = (X_1, \dots, X_j)^T$ and $Y = (Y_1, \dots, Y_k)^T$ be random vectors. If $Z := (X, Y)^T \sim \mathcal{N}(\mu, C)$ and $Cov(X, Y) = 0$, then X and Y are independent.*

Proof. With $C_X = \mathrm{Cov}(X)$ and $C_Y = \mathrm{Cov}(Y)$, we can partition the covariance matrix as

$$C = \begin{pmatrix} C_X & 0 \\ 0 & C_Y \end{pmatrix}.$$

Furthermore, let

$$\mu_X = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_j \end{pmatrix}, \ \mu_Y = \begin{pmatrix} \mu_{j+1} \\ \vdots \\ \mu_{j+k} \end{pmatrix} \text{ and } t_X = \begin{pmatrix} t_1 \\ \vdots \\ t_j \end{pmatrix}, \ t_Y = \begin{pmatrix} t_{j+1} \\ \vdots \\ t_{j+k} \end{pmatrix},$$

such that $X \sim \mathcal{N}(\mu_X, C_X)$ and $Y \sim \mathcal{N}(\mu_Y, C_Y)$. Then it holds for $t := (t_X, t_Y)^T \in \mathbb{R}^{j+k}$ that

$$\begin{aligned} \mathbb{E}[\exp(i\langle t, Z\rangle)] &= \exp(i\langle t, \mu\rangle - \frac{1}{2}\langle t, Ct\rangle) \\ &= \exp(i\langle t_X, \mu_X\rangle + i\langle t_Y, \mu_Y\rangle - \frac{1}{2}\langle t_X, C_X t_X\rangle - \frac{1}{2}\langle t_Y, C_Y t_Y\rangle) \\ &= \mathbb{E}[\exp(i\langle t_X, X\rangle)]\mathbb{E}[\exp(i\langle t_Y, Y\rangle)]. \end{aligned}$$

Because of the fact that the characteristic function determines the distribution uniquely, it follows that

$$\mathscr{L}(Z) = \mathscr{L}(X) \otimes \mathscr{L}(Y),$$

which means that X and $Y$are independent. ❑

With the preliminaries to the normal distribution we can now introduce the definition of the log-normal distribution in higher dimension.

Definition A.4.4. See [71, Definition 2.33]:
Let $\mu \in \mathbb{R}^n$, $C \in \mathbb{R}^{n\times n}$ positive semidefinite, and let $Y = (Y_1,\dots,Y_n) \sim \mathcal{N}(\mu, C)$. Then the distribution of $Z := (\exp(Y_1),\dots,\exp(Y_n))$ is called a n-dimensional log-normal distribution with parameters μ and C and we use the notation $Z \sim M\operatorname{Log}\mathcal{N}(\mu, C)$. In the one-dimensional case of a log-normal distribution, we write $Z \sim \operatorname{Log}\mathcal{N}(\mu, C)$. By definition, a random vector with an n-dimensional log-normal distribution takes values in $(0,\infty)^n$.

Lemma A.4.5 (Properties of the multi-dimensional log-normal distribution)**.**
See [71, Exercise 2.34]: Let $Z = (Z_1,\dots,Z_n) \sim M\operatorname{Log}\mathcal{N}(\mu, C)$ and use the multi-index notation

$$Z^p := \prod_{k=1}^{n} Z_k^{p_k}, \quad p = (p_1,\dots,p_n) \in \mathbb{R}^n.$$

(a) Show that $Z^p \sim \operatorname{Log}\mathcal{N}(\langle p,\mu\rangle, \langle p, Cp\rangle)$ for every $p \in \mathbb{R}^n$.

(b) Show that $\mathbb{E}[Z^p] = \exp(\langle p,\mu\rangle + \frac{1}{2}\langle p, Cp\rangle)$ for every $p \in \mathbb{R}^n$.

(c) Show for all $p, q \in \mathbb{R}^n$ that

$$\begin{aligned} Cov(Z^p, Z^q) &= \mathbb{E}[Z^{p+q}] - \mathbb{E}[Z^p]\mathbb{E}[Z^q] \\ &= \exp(\langle p+q,\mu\rangle + \frac{1}{2}\langle p, Cp\rangle + \frac{1}{2}\langle q, Cq\rangle)(\exp(\langle p, Cq\rangle) - 1). \end{aligned}$$

(d) Show for all $p, q \in \mathbb{R}^n$ that Z^p and Z^q are independent if and only if $Cov(Z^p, Z^q) = 0$, which is the case if and only if $\langle p, Cq\rangle = 0$.

Proof. (a) By Definition A.4.4 there is a random vector $Y = (Y_1,\dots,Y_n) \sim \mathcal{N}(\mu, C)$, such that $Z = (Z_1,\dots,Z_n) = (\exp(Y_1),\dots,\exp(Y_n)) \sim M\operatorname{Log}\mathcal{N}(\mu, C)$. Then we have that

$$Z^p := \prod_{k=1}^{n} Z_k^{p_k} = \prod_{k=1}^{n} \exp(Y_k)^{p_k} = \exp\left(\sum_{k=1}^{n} p_k Y_k\right) = \exp(\langle p, Y\rangle).$$

Using Proposition A.4.2 with $f(y) = \langle p, y\rangle$ for some $p = (p_1,\dots,p_n) \in \mathbb{R}^n$ and $Y \sim \mathcal{N}(\mu, C)$ implies that $\langle p, Y\rangle \sim \mathcal{N}(\langle p,\mu\rangle, \langle p, Cp\rangle)$. Furthermore it follows that $Z^p = \exp(\langle p, Y\rangle) \sim M\operatorname{Log}\mathcal{N}(\mu, C)$.

(b) For $X \sim \mathcal{N}(\mu, \sigma^2)$ we know that

$$\mathbb{E}[tX] = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right).$$

Using this and (a) implies that

$$\begin{aligned} \mathbb{E}[Z^p] &= \mathbb{E}[\exp(1 \cdot \underbrace{\langle p, Y\rangle}_{\sim \mathcal{N}(\langle p,\mu\rangle, \langle p, Cp\rangle)})] \\ &= \exp\left(\langle p,\mu\rangle \cdot 1 + \frac{1}{2}\langle p, Cp\rangle \cdot 1^2\right) \\ &= \exp\left(\langle p,\mu\rangle + \frac{1}{2}\langle p, Cp\rangle\right). \end{aligned}$$

(c) The covariance is given by $\mathrm{Cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$, such that

$$\begin{aligned}
\mathrm{Cov}(Z^p, Z^q) = \mathbb{E}[Z^{p+q}] - \mathbb{E}[Z^p]\mathbb{E}[Z^q] \quad (\text{ using (b)}) \\
&= \exp\Big(\langle p+q, \mu\rangle + \frac{1}{2}\langle p+q, C(p+q)\rangle\Big) \\
&\quad - \exp\Big(\langle p, \mu\rangle + \frac{1}{2}\langle p, Cp\rangle\Big) \cdot \exp\Big(\langle q, \mu\rangle + \frac{1}{2}\langle q, Cq\rangle\Big) \\
&\text{(using bilinearity of } \langle\cdot,\cdot\rangle \text{ and exclude a factor)} \\
&= \exp\Big(\langle p+q, \mu\rangle + \frac{1}{2}\langle p, Cp\rangle + \frac{1}{2}\langle q, Cq\rangle\Big)\Big(\exp(\langle p, Cq\rangle) - 1\Big).
\end{aligned}$$

(d) For all $p, q \in \mathbb{R}^n$ it follows from (a) and the proof of Proposition A.4.3 that Z^p and Z^q are independent if and only if $\mathrm{Cov}(Z^p, Z^q) = 0$. Finally, it follows from (c) that $\mathrm{Cov}(Z^p, Z^q) = 0$ if and only if $\langle p, Cq\rangle = 0$. ❑

A.5. Fubini and Transition Kernel

To introduce our distribution-constrained optimization problem OptStop$^\gamma$ in continuous time similar to the one in discrete time, we must exchange the adapted random probability measures by stochastic transition kernels, see Section 6.1 of Part II. Here are some preliminary considerations on this subject from various sources.

Definition A.5.1. Cf. [65, Definition 8.4]:
Let $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{J}, \lambda)$ be σ-finite measure spaces, and let f be an $(\mathcal{S} \times \mathcal{J})$-measurable function on $X \times Y$.
With each function f on $X \times Y$ and with each $x \in X$ we associate a function f_x defined on Y by $f_x(y) = f(x, y)$. Similarly, if $y \in Y$, f^y is the function defined on X by $f^y(x) = f(x, y)$.

Theorem A.5.2 (Fubini). *See [65, Theorem 8.8]:*
Let $(X, \mathcal{S}, \mu)$ and $(Y, \mathcal{J}, \lambda)$ be σ-finite measure spaces, and let f be an $(\mathcal{S} \times \mathcal{J})$-measurable function on $X \times Y$.

(a) If $0 \le f \le \infty$, and if

$$\phi(x) = \int_Y f_x \, d\lambda, \quad \psi(y) = \int_X f^y \, d\mu \quad (x \in X, y \in Y), \tag{A.5.3}$$

then ϕ is $\mathcal{S}$-measurable, ψ is $\mathcal{J}$-measurable, and

$$\int_X \phi \, d\mu = \int_{X \times Y} f \, d(\mu \times \lambda) = \int_Y \psi \, d\lambda. \tag{A.5.4}$$

(b) If f is complex and if

$$\phi^*(x) = \int_Y |f|_x \, d\lambda \quad \text{and} \quad \int_X \phi^* d\mu < \infty, \tag{A.5.5}$$

then $f \in L^1(\mu \times \lambda)$.

(c) If $f \in L^1(\mu \times \lambda)$, then $f_x \in L^1(\lambda)$ for almost all $x \in X$, $f^y \in L^1(\mu)$ for almost all $y \in Y$; the functions ϕ and ψ, defined by (A.5.3) a.e., are in $L^1(\mu)$ and $L^1(\lambda)$, respectively, and (A.5.4) holds.

Remark A.5.6. Notes: The first and last integrals in (A.5.4) can also be written in the more usual form

$$\int_X d\mu(x) \int_Y f(x, y) \, d\lambda(y) = \int_Y d\lambda(y) \int_X f(x, y) \, d\mu(x). \tag{A.5.7}$$

These are the so-called "iterated integrals" of f. The middle integral in (A.5.4) is often referred to as a double integral. The combination of (b) and (c) gives the following useful result: If f is $(\mathcal{S} \times \mathcal{J})$-measurable and if

$$\int_X d\mu(x) \int_Y |f(x, y)| \, d\lambda(y) < \infty,$$

then the two iterated integrals (A.5.7) are finite and equal.
In other words, "the order of integration may be reversed" for $(\mathcal{S} \times \mathcal{J})$-measurable functions f whenever $f \ge 0$ and also whenever one of the iterated integrals of $|f|$ is finite.

See [41, Chapter 14]: Consider now the situation of finitely many σ-finite measure spaces $(\Omega_i, \mathcal{A}_i, \mu_i)$, $i = 1, \ldots, n$, where $n \in \mathbb{N}$.

Lemma A.5.8. *See [41, Lemma 14.13]:*
Let $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ and let $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ be an $\mathcal{A}_1 \otimes \mathcal{A}_2$-measurable map. Then, for all $\tilde{\omega}_1 \in \Omega_1$ and $\tilde{\omega}_2 \in \Omega_2$,

$$\begin{aligned} A_{\tilde{\omega}_1} &:= \{\omega_2 \in \Omega_2 : (\tilde{\omega}_1, \omega_2) \in A\} \in \mathcal{A}_2, \\ A_{\tilde{\omega}_2} &:= \{\omega_1 \in \Omega_1 : (\omega_1, \tilde{\omega}_2) \in A\} \in \mathcal{A}_1, \\ f_{\tilde{\omega}_1} &: \Omega_2 \to \overline{\mathbb{R}}, \quad \omega_2 \mapsto f(\tilde{\omega}_1, \omega_2) \text{ is } \mathcal{A}_2\text{-measurable}, \\ f_{\tilde{\omega}_2} &: \Omega_1 \to \overline{\mathbb{R}}, \quad \omega_1 \mapsto f(\omega_1, \tilde{\omega}_2) \text{ is } \mathcal{A}_1\text{-measurable}. \end{aligned}$$

Theorem A.5.9 (Finite product measures)**.** *See [41, Theorem 14.14]:*
There exists a unique σ-finite measure μ on $\mathcal{A} := \bigotimes_{i=1}^n \mathcal{A}_i$ such that

$$\mu(A_1 \times \ldots \times A_n) = \prod_{i=1}^n \mu_i(A_i) \quad \text{for } A_i \in \mathcal{A}_i, i = 1, \ldots, n.$$

$\bigotimes_{i=1}^n \mu_i := \mu_1 \otimes \ldots \otimes \mu_n := \mu$ is called the product measure of the μ_i. If all spaces involved equal $(\Omega_0, \mathcal{A}_0, \mu_0)$, then we write $\mu^{\otimes n} := \bigotimes_{i=1}^n \mu_0$.

We come next to a concept that generalizes the notion of product measure. Recall the definition of a transition kernel, which is given in the following way:

Definition A.5.10 (Transition kernel, Markov kernel)**.** See [41, Definition 8.25]:
Let $(\Omega_1, \mathcal{A}_1)$, $(\Omega_2, \mathcal{A}_2)$, be measurable spaces. A map $\kappa : \Omega_1 \times \mathcal{A}_2 \to [0, \infty]$ is called a (σ-)finite transition kernel (from Ω_1 to Ω_2) if:

(i) $\omega_1 \mapsto \kappa(\omega_1, A_2)$ is $\mathcal{A}_1$-measurable for any $A_2 \in \mathcal{A}_2$.

(ii) $A_2 \mapsto \kappa(\omega_1, A_2)$ is a (σ-)finite measure on $(\Omega_2, \mathcal{A}_2)$ for any $\omega_1 \in \Omega_1$.

If in (ii) the measure is a probability measure for all $\omega_1 \in \Omega_1$, then κ is called a stochastic kernel or a Markov kernel. If in (ii) we also have $\kappa(\omega_1, \Omega_2) \leq 1$ for any $\omega_1 \in \Omega_1$, then κ is called sub-Markov or substochastic.

Lemma A.5.11. *See [41, Lemma 14.20]:*
Let κ be a finite transition kernel from $(\Omega_1, \mathcal{A}_1)$ to $(\Omega_2, \mathcal{A}_2)$ and let $f : \Omega_1 \times \Omega_2 \to [0, \infty]$ be measurable with respect to $(\mathcal{A}_1 \otimes \mathcal{A}_2)$-$\mathcal{B}([0, \infty])$. Then the map

$$\begin{aligned} I_f &: \Omega_1 \to [0, \infty], \\ \omega_1 &\mapsto \int f(\omega_1, \omega_2) \kappa(\omega_1, d\omega_2), \end{aligned}$$

is well-defined and $\mathcal{A}_1$-measurable.

Remark A.5.12. See [41, Remark 14.21]:
In the following, we often write $\int \kappa(\omega_1, d\omega_2) f(\omega_1, \omega_2)$ instead of $\int f(\omega_1, \omega_2) \kappa(\omega_1, d\omega_2)$ since for multiple integrals this notation allows us to write the integrator closer to the corresponding integral sign.

Theorem A.5.13. *See [41, Theorem 14.22]:*
Let $(\Omega_i, \mathcal{A}_i)$, $i = 0,1,2$, be measurable spaces. Let κ_1 be a finite transition kernel from $(\Omega_0, \mathcal{A}_0)$ to $(\Omega_1, \mathcal{A}_1)$ and let κ_2 be a finite transition kernel from $(\Omega_0 \times \Omega_1, \mathcal{A}_0 \otimes \mathcal{A}_1)$ to $(\Omega_2, \mathcal{A}_2)$. Then the map

$$\kappa_1 \otimes \kappa_2 : \Omega_0 \times (\mathcal{A}_1 \otimes \mathcal{A}_2) \to [0,\infty),$$
$$(\omega_0, A) \mapsto \int_{\Omega_1} \kappa_1(\omega_0, d\omega_1) \int_{\Omega_2} \kappa_2((\omega_0,\omega_1), d\omega_2) \mathbb{1}_A((\omega_1,\omega_2)),$$

is well-defined and is a σ-finite (but not necessarily a finite) transition kernel from $(\Omega_0, \mathcal{A}_0)$ to $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$. If κ_1 and κ_2 are (sub)stochastic, then $\kappa_1 \otimes \kappa_2$ is (sub)stochastic. We call $\kappa_1 \otimes \kappa_2$ the product of κ_1 and κ_2.
If κ_2 is a kernel from $(\Omega_1, \mathcal{A}_1)$ to $(\Omega_2, \mathcal{A}_2)$, then we define the product $\kappa_1 \otimes \kappa_2$ similarly by formally understanding κ_2 as a kernel from $(\Omega_0 \times \Omega_1, \mathcal{A}_0 \otimes \mathcal{A}_1)$ to $(\Omega_2, \mathcal{A}_2)$ that does not depend on the Ω_0-coordinate.

Corollary A.5.14 (Products via kernels)**.** *See [41, Corollary 14.23]:*
Let $(\Omega_1, \mathcal{A}_1, \mu)$ be a finite measure space, let $(\Omega_2, \mathcal{A}_2)$ be a measurable space and let κ be a finite transition kernel from Ω_1 to Ω_2. Then there exists a unique σ-finite measure $\mu \otimes \kappa$ on $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ with

$$\mu \otimes \kappa(A_1 \times A_2) = \int_{A_1} \kappa(\omega_1, A_2)\mu(d\omega_1) \quad \text{for all } A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2.$$

If κ is stochastic and if μ is a probability measure, then $\mu \otimes \kappa$ is a probability measure.

Corollary A.5.15. *See [41, Corollary 14.24]:*
Let $n \in \mathbb{N}$ and let $(\Omega_i, \mathcal{A}_i)$, $i = 0, \dots, n$, be measurable spaces. For $i = 1, \dots, n$, let κ_i be a substochastic kernel from $(\bigtimes_{k=0}^{i-1} \Omega_k, \bigotimes_{k=0}^{i-1} \mathcal{A}_k)$ to $(\Omega_i, \mathcal{A}_i)$ or from $(\Omega_{i-1}, \mathcal{A}_{i-1})$ to $(\Omega_i, \mathcal{A}_i)$. Then the recursion $\kappa_1 \otimes \ldots \otimes \kappa_i := (\kappa_1 \otimes \ldots \otimes \kappa_{i-1}) \otimes \kappa_i$ for any $i = 1, \dots, n$ defines a substochastic kernel $\bigotimes_{k=1}^{i} \kappa_k := \kappa_1 \otimes \ldots \otimes \kappa_i$ from $(\Omega_0, \mathcal{A}_0)$ to $(\bigtimes_{k=1}^{i} \Omega_k, \bigotimes_{k=1}^{i} \mathcal{A}_k)$. If all κ_k are stochastic, then all $\bigotimes_{k=1}^{i} \kappa_k$ are stochastic.
If μ is a finite measure on $(\Omega_0, \mathcal{A}_0)$, then $\mu_i := \mu \otimes \bigotimes_{k=1}^{i} \kappa_k$ is a finite measure on $(\bigtimes_{k=0}^{i} \Omega_k, \bigotimes_{k=0}^{i} \mathcal{A}_k)$. If μ is a probability measure and if every κ_i is stochastic, then μ_i is a probability measure.

Definition A.5.16 (Composition of kernels)**.** See [41, Definition 14.25]:
Let $(\Omega_i, \mathcal{A}_i)$, $i = 0,1,2$, be measurable spaces and let κ_i be a substochastic kernel from $(\Omega_{i-1}, \mathcal{A}_{i-1})$ to $(\Omega_i, \mathcal{A}_i)$, $i = 1,2$. Define the composition of κ_1 and κ_2 by

$$\kappa_1 \cdot \kappa_2 : \Omega_0 \times \mathcal{A}_2 \to [0,\infty),$$
$$(\omega_0, A_2) \mapsto \int_{\Omega_1} \kappa_1(\omega_0, d\omega_1)\kappa_2(\omega_1, A_2).$$

Theorem A.5.17. *See [41, Theorem 14.26]:*
If we denote by $\pi_2 : \Omega_1 \times \Omega_2 \to \Omega_2$ the projection to the second coordinate, then

$$(\kappa_1 \cdot \kappa_2)(\omega_0, A_2) = (\kappa_1 \otimes \kappa_2)(\omega_0, \pi_2^{-1}(A_2)) \quad \text{for all } A_2 \in \mathcal{A}_2.$$

In particular, the composition $\kappa_1 \cdot \kappa_2$ is a (sub)stochastic kernel from $(\Omega_0, \mathcal{A}_0)$ to $(\Omega_2, \mathcal{A}_2)$.

Theorem A.5.18 (Fubini for transition kernels)**.** *See [41, Theorem 14.29]:*
Let $(\Omega_i, \mathcal{A}_i)$ be measurable spaces, $i = 1,2$. Let μ be a finite measure on $(\Omega_1, \mathcal{A}_1)$ and let κ be a finite transition kernel from Ω_1 to Ω_2 . Assume that $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ is measurable with respect to $\mathcal{A}_1 \otimes \mathcal{A}_2$. If $f \geq 0$ or $f \in L^1(\mu \otimes \kappa)$, then

$$\int_{\Omega_1 \times \Omega_2} f \, d(\mu \otimes \kappa) = \int_{\Omega_1} \left(\int_{\Omega_2} f(\omega_1, \omega_2) \kappa(\omega_1, d\omega_2) \right) \mu(d\omega_1).$$

See [39]: Given two measurable spaces $(S, \mathcal{S})$ and $(T, \mathcal{T})$, a mapping $\mu : S \times \mathcal{T} \to \overline{\mathbb{R}}_+$ is called a (probability) kernel from S to T if the function $\mu_s B = \mu(s, B)$ is $\mathcal{S}$-measurable in $s \in S$ for fixed $B \in \mathcal{T}$ and a (probability) measure in $B \in \mathcal{T}$ for fixed $s \in S$. Any kernel μ determines an associated operator that maps suitable functions $f : T \to \mathbb{R}$ into their integrals $\mu f(s) = \int \mu(s, dt) f(t)$. Note in particular that the class $\mathcal{P}(S)$ of probability measures on S is a measurable subset of $\mathcal{M}(S)$. Kernels play an important role in probability theory, where they may appear in the guises of random measures, conditional distributions, Markov transition functions, and potentials. The following characterizations of the kernel property are often useful. For simplicity we are restricting our attention to probability kernels.

Lemma A.5.19 (Kernels)**.** *See [39, Lemma 1.37]:*
Fix two measurable spaces $(S, \mathcal{S})$ and $(T, \mathcal{T})$, a π-system $\mathcal{C}$ with $\sigma(\mathcal{C}) = \mathcal{T}$, and a family $\mu = \{\mu_s; s \in \mathcal{S}\}$ of probability measures on T. Then these conditions are equivalent:

(i) *μ is a probability kernel from S to T;*

(ii) *μ is a measurable mapping from S to $\mathcal{P}(T)$;*

(iii) *$s \mapsto \mu_s B$ is a measurable mapping from S to $[0,1]$ for every $B \in \mathcal{C}$.*

More primitive classes than σ-fields often arise in applications. A class $\mathcal{C}$ of subsets of some space Ω is called a π-system if it is closed under finite intersections, so that A, $B \in \mathcal{C}$ implies $A \cap B \in \mathcal{C}$. Furthermore, a class $\mathcal{D}$ is a λ-system if it contains Ω and is closed under proper differences and increasing limits. Thus, we require that $\Omega \in \mathcal{D}$, that A, $B \in \mathcal{D}$ with $A \supset B$ implies $A \setminus B \in \mathcal{D}$, and that $A_1, A_2, \ldots \in \mathcal{D}$ with $A_n \uparrow A$ implies $A \in \mathcal{D}$.
Let us now introduce a third measurable space $(U, \mathcal{U})$, and consider two kernels μ and ν, one from S to T and the other from $S \times T$ to U. Imitating the construction of product measures, we may attempt to combine μ and ν into a kernel $\mu \otimes \nu$ from S to $T \times U$ given by

$$(\mu \otimes \nu)(s, B) = \int \mu(s, dt) \int \nu(s, t, du) 1_B(t, u), \quad B \in \mathcal{T} \otimes \mathcal{U}.$$

The following lemma justifies the formula and provides some further useful information.

Lemma A.5.20 (Kernels and functions)**.** *See [39, Lemma 1.38]:*
Fix three measurable spaces $(S,\mathcal{S})$, $(T,\mathcal{T})$, and $(U,\mathcal{U})$. Let μ and ν be probability kernels from S to T and from $S\times T$ to U, respectively, and consider two measurable functions $f : S\times T \to \mathbb{R}_+$ and $g : S\times T \to U$. Then

(i) $\mu_s f(s,\cdot)$ *is a measurable function of* $s\in S$*;*

(ii) $\mu_s \circ (g(s,\cdot))^{-1}$ *is a kernel from* S *to* U*;*

(iii) $\mu\otimes\nu$ *is a kernel from* S *to* $T\times U$*.*

For any measurable function $f\geq 0$ on $T\times U$, we get as in [39, Theorem 1.27]

$$(\mu\otimes\nu)_s f = \int \mu(s,dt)\int \nu(s,t,du)f(t,u), \quad s\in S,$$

or simply $(\mu\otimes\nu)f = \mu(\nu f)$. By iteration we may combine any kernels μ_k from $S_0\times\cdots\times S_{k-1}$ to S_k , $k = 1,\ldots,n$, into a kernel $\mu_1\otimes\cdots\otimes\mu_n$ from S_0 to $S_1\times\cdots\times S_n$, given by

$$(\mu_1\otimes\cdots\otimes\mu_n)f = \mu_1(\mu_2(\cdots(\mu_n f)\cdots))$$

for any measurable function $f\geq 0$ on $S_1\times\cdots\times S_n$. In applications we may often encounter kernels μ_k from S_{k-1} to S_k , $k = 1,\ldots,n$, in which case the composition $\mu_1\cdots\mu_n$ is defined as a kernel from S_0 to S_n given for measurable $B\subseteq S_n$ by

$$\begin{aligned}(\mu_1\cdots\mu_n)_s B &= (\mu_1\otimes\cdots\otimes\mu_n)(S_1\times\cdots\times S_{n-1}\times B)\\ &= \int \mu_1(s,ds_1)\int \mu_2(s_1,ds_2)\cdots\int \mu_{n-1}(s_{n-2},ds_{n-1})\mu_n(s_{n-1},B).\end{aligned}$$

Let $(E,\mathcal{E})$ be a measurable space and $Q : E\times\mathcal{E}\to[-1,1]$ be a signed bounded kernel, i.e., $Q_x(\cdot)$ is a finite measure on $(E,\mathcal{E})$ for any $x\in E$ and $x\mapsto Q_x(A)$ is a measurable function for any set $A\in\mathcal{E}$. For any fixed x, let the measure Q_x^+ be the positive part of the signed measure Q_x as in Hahn-Jordan decomposition. Is it true that Q^+ is a kernel, i.e., is the function $x\mapsto Q_x^+(A)$ measurable for any $A\in\mathcal{E}$? It clearly holds if Q is an integral kernel, i.e., $Q(x,dy) = q(x,y)\mu(dy)$, where μ is a finite measure on $(E,\mathcal{E})$ and $q : E\times E\to\mathbb{R}$ is a jointly measurable function, but I am interested in the general case.
The following answer is based on [23]. The result holds true under the assumption that $(E,\mathcal{E})$ is countably generated. The algebra generated by a countable set is countable, so we can assume without loss of generality that there is a countable algebra $\mathcal{A}$ with $\mathcal{E} = \sigma(\mathcal{A})$. Since the difference of two measurable functions is measurable, it suffices to show that Q^+ is a kernel.
For all $x\in E$ and $B\in\mathcal{E}$, we have

$$Q_x^+(B) = \sup_{A\in\mathcal{E},A\subseteq B} Q_x(A).$$

Let $\alpha\in[0,1]$. We have

$$\{x : Q_x(B)^+ < \alpha\} = \bigcup_{A\in\mathcal{E},A\subseteq B}\{x : Q_x(A) < \alpha\}.$$

Now the union on the righthand side is generally over an uncountable set. But for each $x \in E$, $B \in \mathcal{E}$, and $\epsilon > 0$ there is $A \in \mathcal{A}$ such that $Q_x(B \Delta A) < \epsilon$. It follows that

$$\bigcup_{A \in \mathcal{E}, A \subseteq B} \{x : Q_x(A) < \alpha\} = \bigcup_{A \in \mathcal{E}} \{x : Q_x(A \cap B) < \alpha\}.$$

Here is another proof, taken from [62, Lemma 1.5, page 190], always under the assumption that $(E, \mathcal{E})$ is countably generated. The following answer is almost a paraphrase of the proof of Revuz.
By assumption, there is a sequence of finite partitions $\mathcal{P}_n$ of E, such that $\mathcal{P}_{n+1}$ is a refinement of $\mathcal{P}_n$, and $\mathcal{E}$ is generated by $\bigcup_{n \geq 0} \mathcal{P}_n$. For every $x \in E$, there exists a unique $E_n^x \in \mathcal{P}_n$ with $x \in E_n^x$.
Let $x \in E$ be fixed for the moment. Define λ_x as the probability measure which is a multiple of $|Q_x|$ (if $Q_x = 0$, choose it as you want). Then define a function f_n on E by

$$f_n(y) = \begin{cases} \frac{Q_x(E_n^y)}{\lambda_x(E_n^y)} & \text{if } \lambda_x(E_n^y) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

By martingale convergence theorem, we have that f_n converges λ_x-a.s. to the density of Q_x with respect to λ_x. Hence, f_n^+ converges λ_x-a.s. to the density of Q_x^+, and since f_n are uniformly bounded, we have for all $A \in \mathcal{E}$:

$$Q_x^+(A) = \lim_n \int_A f_n^+ \, d\lambda_x.$$

Now, if $A \in \mathcal{P}_k$, then for all $n > k$, $\int_A f_n^+ \, d\lambda_x = Q_{x,n}^+(A)$, where $Q_{x,n}^+$ is the positive part of the restriction of Q_x to the σ-algebra generated by $\mathcal{P}_n$. It is easy to see that the map $x \mapsto Q_{x,n}^+(A)$ is measurable, and so is the map $x \mapsto Q_x^+(A)$.

Hence, we have proven that $x \mapsto Q_x^+(A)$ is measurable for every $A \in \bigcup_{n \geq 0} \mathcal{P}_n$, and then a Dynkin class argument finishes the proof. (By monoton class theorem Q^+ is a kernel.)

Lemma A.5.21. *See [23, 2.2]:*
Let M be the set of all countably additive, finite, signed measures on a sigma-field Σ of subsets of a set X. There is a natural definition of measurability in M, namely, a subset of M is measurable if it is an element of Σ^*, the smallest σ-field of subsets of M such that; for each $A \in \Sigma$ the function $\mu \mapsto \mu(A)$ is measurable from M to the Borel line.
Let X be a non-empty set, $\mathcal{F}$ a countable field of subsets of X, and Σ the smallest σ-field including $\mathcal{F}$.
If ϕ is a measurable map from $(\Omega, \mathcal{A})$ to (M, Σ^*), and f is a bounded, measurable function from $(\Omega \times X, \mathcal{A} \times \Sigma)$ to the Borel line, then $\omega \mapsto \int_X f(\omega, x)\phi(\omega)(dx)$ is a measurable function from $(\Omega, \mathcal{A})$ to the Borel line.

Bibliography

[1] L. Alili and P. Patie, *On the first crossing times of a Brownian motion and a family of continuous curves*, Comptes Rendus Mathematique, 340 (2005), pp. 225–228.

[2] ——, *Boundary crossing identities for Brownian motion and some nonlinear ode's*, Proceedings of the American Mathematical Society, 142 (2014), pp. 3811–3824.

[3] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, Berlin; London, 3. ed., 2006.

[4] S. V. Anulova, *On Markov Stopping Times with a Given Distribution for a Wiener Process*, Theory of Probability & Its Applications, 25 (1981), pp. 362–366.

[5] M. Avellaneda and J. Zhu, *Distance to Default*, RISK, 14 (2001), pp. 125–129.

[6] A. R. Bacinello, P. Millossovich, A. Olivieri, and E. Pitacco, *Variable annuities: A unifying valuation approach*, Insurance: Mathematics and Economics, 49 (2011), pp. 285–297.

[7] A. R. Bacinello and I. Zoccolan, *Variable Annuities with State-Dependent Fees: Valuation, Numerical Implementation, Comparative Static Analysis and Model Risk*, EUT Edizioni Università di Trieste, (2017).

[8] E. Bayraktar and C. W. Miller, *Distribution-constrained optimal stopping*, Mathematical Finance, 29, pp. 368–406.

[9] M. Beiglböck, A. M. G. Cox, and M. Huesmann, *Optimal transport and Skorokhod embedding*, Inventiones Mathematicae, (2016), pp. 1–74.

[10] M. Beiglböck, M. Eder, C. Elgert, and U. Schmock, *Geometry of distribution-constrained optimal stopping problems*, Probability Theory and Related Fields, 172 (2018), pp. 71–101.

[11] M. Beiglböck and C. Griessler, *A land of monotone plenty*, Annali della SNS, to appear, (2015).

[12] M. Beiglböck, P. Henry-Labordère, and F. Penkner, *Model-independent bounds for option prices – a mass transport approach*, Finance and Stochastics, 17 (2013), pp. 477–501.

[13] M. Beiglböck and A. Pratelli, *Duality for rectified cost functions*, Calculus of Variations and Partial Differential Equations, 45 (2012), pp. 27–41.

[14] M. Beiglböck and N. Juillet, *On a problem of optimal transport under marginal martin-*

gale constraints, The Annals of Probability, 44 (2016), pp. 42–106.

[15] M. Beiglböck, M. Nutz, and N. Touzi, *Complete duality for martingale optimal transport on the line*, The Annals of Probability, 45 (2017), pp. 3038–3074.

[16] L. Breiman, *First exit times from a square root boundary*, in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Contributions to Probability Theory, Part 2, Berkeley, Calif., 1967, University of California Press, pp. 9–16.

[17] L. Campi, I. Laachir, and C. Martini, *Change of numeraire in the two-marginals martingale transport problem*, Finance and Stochastics, 21 (2017), pp. 471–486.

[18] X. Chen, L. Cheng, J. Chadam, and D. Saunders, *Existence and uniqueness of solutions to the inverse boundary crossing problem for diffusions*, The Annals of Applied Probability, 21 (2011), pp. 1663–1693.

[19] M. Christiansen and M. Steffensen, *Safe-side scenarios for financial and biometrical risk*, ASTIN Bulletin, 43 (2013), pp. 323–357.

[20] A. Cox and S. Källblad, *Model-independent bounds for Asian options: A dynamic programming approach*, SIAM Journal on Control and Optimization (SICON), 55 (2017), pp. 3409–3436.

[21] C. Dellacherie and P.-A. Meyer, *Probabilities and Potential*, vol. 29 of North-Holland Mathematics Studies, North-Holland Publishing Co., Amsterdam, 1978.

[22] Y. Dolinsky and H. M. Soner, *Martingale optimal transport and robust hedging in continuous time*, Probability Theory and Related Fields, 160 (2014), pp. 391–427.

[23] L. Dubins and D. Freedman, *Measurable sets of measures.*, Pacific Journal of Mathematics, 14 (1964), pp. 1211–1222.

[24] R. M. Dudley and S. Gutmann, *Stopping times with given laws*, in Séminaire de Probabilités XI, C. Dellacherie, P. A. Meyer, and M. Weil, eds., Springer Berlin Heidelberg, 1977, pp. 51–58.

[25] R. Durrett, *Probability: Theory and Examples*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2010.

[26] T. Efferth, M. Banerjee, and N. Paul, *Broken heart, tako-tsubo or stress cardiomyopathy? Metaphors, meanings and their medical impact*, International Journal of Cardiology, 230 (2017), pp. 262–268.

[27] E. Ekström and S. Janson, *The inverse first-passage problem and optimal stopping*, arXiv e-prints, (2015). arXiv:1508.07827.

[28] A. Galichon, P. Henry-Labordère, and N. Touzi, *A Stochastic Control Approach to No-Arbitrage Bounds Given Marginals, with an Application to Lookback Options*, The Annals of Applied Probability, 24 (2014), pp. 312–336.

[29] W. Gangbo and R. J. McCann, *The geometry of optimal transportation*, Acta Mathematica, 177 (1996), pp. 113–161.

[30] N. Ghoussoub, Y.-H. Kim, and T. Lim, *Structure of optimal martingale transport plans*

in general dimensions, The Annals of Probability, 47 (2019), pp. 109–164.

[31] A. Grass, *Uniqueness and Stability Properties of Barrier Type Skorokhod Embeddings*, Master thesis, University of Vienna, (2016).

[32] G. Guo, X. Tan, and N. Touzi, *On the Monotonicity Principle of Optimal Skorokhod Embedding Problem*, SIAM Journal on Control and Optimization (SICON), 54 (2016), pp. 2478–2489.

[33] K. Hirhager, *Adapted Dependence with Applications to Financial and Actuarial Risk Management*, PhD Thesis, TU Wien, Vienna, (2013).

[34] D. Hobson, *The Skorokhod embedding problem and model-independent bounds for option prices*, in Paris-Princeton Lectures on Mathematical Finance 2010, vol. 2003 of Lecture Notes in Math., Springer, Berlin, 2011, pp. 267–318.

[35] D. Hobson and A. Neuberger, *Robust bounds for forward start options*, Mathematical Finance, 22 (2012), pp. 31–56.

[36] S. Jaimungal, A. Kreinin, and A. Valov, *The generalized Shiryaev problem and Skorokhod embedding*, Theory of Probability & Its Applications, 58 (2014), pp. 493–502.

[37] R. Kainhofer, M. Predota, and U. Schmock, *The new Austrian annuity valuation table AVÖ 2005R*, Mitteilungen der Aktuarvereinigung Österreich, 13 (2006), pp. 55–135. Also available as `http://www.avoe.at/pdf/mitteilungen/H13_w3.pdf`.

[38] S. Källblad, *A Dynamic Programming Principle for Distribution-Constrained Optimal Stopping*, arXiv e-prints, (2017). arXiv:1703.08534.

[39] O. Kallenberg, *Foundations of Modern Probability*, Probability and Its Applications, Springer New York, 2002.

[40] A. Kechris, *Classical Descriptive Set Theory*, Graduate Texts in Mathematics, Springer New York, 1995.

[41] A. Klenke, *Probability Theory: A Comprehensive Course*, Universitext, Springer London, 2013.

[42] M. Knott and C. S. Smith, *On the optimal mapping of distributions*, Journal of Optimization Theory and Applications, 43 (1984), pp. 39–49.

[43] M. Koller, *Stochastic Models in Life Insurance*, Springer, Berlin Heidelberg, 2. ed., 2010.

[44] N. Kuroiwa, S. Fukuzawa, and S. Okino, *The impact of northeast Japan earthquake on acute myocardial infarction and heart failure*, Journal of the American College of Cardiology, 61 (2013), p. E45.

[45] H. R. Lerche, *Boundary Crossing of Brownian Motion: Its Relation to the Law of the Iterated Logarithm and to Sequential Analysis*, Lecture Notes in Statistics, Springer, New York, 1986.

[46] T. Lindvall, *On Strassen's Theorem on Stochastic Domination*, Electron. Commun. Probab., 4 (1999), pp. 51–59.

[47] X. Liu, R. Mamon, and H. Gao, *A comonotonicity-based valuation method for guaranteed annuity options*, Journal of Computational and Applied Mathematics, 250 (2013), pp. 58–69.

[48] R. M. Loynes, *Stopping times on Brownian motion: Some properties of Root's construction*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 16 (1970), pp. 211–218.

[49] LTGA, *Long-Term Guarantees Assessment – Technical Specifications*. https://eiopa.europa.eu/consultations/qis/insurance/long-term-guarantees-assessment/technical-specifications/index.html.

[50] Y. Lu, *Broken-heart, common life, heterogeneity: Analyzing the spousal mortality dependence*, ASTIN Bulletin, 47 (2017), pp. 837–874.

[51] E. Makaris, G. Michas, R. Micha, I. Pisimisis, I. Tsichlis, D. Svoronos, C. Panotopoulos, D. Gkotsis, G. Koudounis, and S. Zompolos, *Socioeconomic crisis and incidence of acute myocardial infarction in Messinia, Greece*, Journal of the American College of Cardiology, 61 (2013), p. E47.

[52] A. Müller, *Stochastic Orders Generated by Integrals: A Unified Study*, Advances in Applied Probability, 29 (1997), pp. 414–428.

[53] A. Müller and M. Scarsini, *Stochastic order relations and lattices of probability measures*, SIAM Journal on Optimization, 16 (2006), pp. 1024–1043.

[54] M. Niiyama, F. Tanaka, K. Sato, T. Onoda, T. Takahashi, T. Segawa, M. Honma, and M. Nakamura, *Increase in the incidence of acute coronary syndrome after the 2011 east Japan natural disaster: Concordance with sequential quake shocks*, Journal of the American College of Cardiology, 61 (2013), p. E46.

[55] M. Nutz and F. Stebegg, *Canonical supermartingale couplings*, The Annals of Probability, 46 (2018), pp. 3351–3398.

[56] J. Obłój, *The Skorokhod embedding problem and its offspring*, Probability Surveys, 1 (2004), pp. 321–390.

[57] G. Pammer, *Distribution-Constrained Optimal Stopping Problems in Discrete Time*, Thesis (Diplom), TU Wien, Vienna, (2017).

[58] G. Peskir, *On Integral Equations Arising in the First-Passage Problem for Brownian Motion*, Journal of Integral Equations and Applications, 14 (2002), pp. 397–423.

[59] M. N. Peters, M. J. Katz, J. C. Moscona, M. E. Alkadri, R. K. Syed, T. A. Turnage, V. S. Nijjar, M. B. Bisharat, P. Delafontaine, and A. M. Irimpen, *Alteration in the Chronobiology of Onset of Acute Myocardial Infarction in New Orleans Residents Following Hurricane Katrina*, Journal of the American College of Cardiology, 61 (2013), p. E48.

[60] QIS5, *Quantitative impact study 5 – technical specifications*. https://eiopa.europa.eu/consultations/qis/quantitative-impact-study-5/technical-specifications/index.html.

[61] M. Reed and B. Simon, *Methods of Modern Mathematical Physics: Functional Analysis*, Academic Press, 1980.

[62] D. Revuz, *Markov Chains*, North-Holland Mathematical Library, Elsevier Science, 2008.

[63] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales*, vol. 2 of Cambridge Mathematical Library, Cambridge University Press, 2. ed., 2000.

[64] D. H. Root, *The existence of certain stopping times on Brownian motion*, Annals of Mathematical Statistics, 40 (1969), pp. 715–718.

[65] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, Inc., New York, 3. ed., 1987.

[66] ——, *Functional Analysis*, McGraw-Hill Inc., New York, 2. ed., 1991.

[67] L. Rüschendorf, *Fréchet-bounds and their applications*, in Advances in probability distributions with given marginals (Rome, 1990), vol. 67 of Math. Appl., Kluwer Acad. Publ., Dordrecht, 1991, pp. 151–187.

[68] ——, *Optimal solutions of multivariate coupling problems*, Applicationes Mathematicae (Warsaw), 23 (1995), pp. 325–338.

[69] P. Salminen, *On the First Hitting Time and the Last Exit Time for a Brownian Motion to/from a Moving Boundary*, Advances in Applied Probability, 20 (1988), pp. 411–426.

[70] R. Schilling, *Measures, Integrals and Martingales*, Cambridge University Press, Bd. 13, 2005.

[71] U. Schmock, *Stochastic Analysis for Financial and Actuarial Mathematics*. `https://fam.tuwien.ac.at/~schmock/tmp/StochasticAnalysis_20190122.pdf`, Lecture Notes.

[72] M. Shaked and G. Shanthikumar, *Stochastic Orders*, Springer Series in Statistics, Springer New York, 2007.

[73] C. Villani, *Optimal Transport: Old and New*, vol. 338 of Grundlehren der mathematischen Wissenschaften, Springer Science & Business Media, 2008.

[74] R. B. Washburn and A. S. Willsky, *Optional Sampling of Submartingales Indexed by Partially Ordered Sets*, The Annals of Probability, 9 (1981), pp. 957–970.

[75] D. Williams, *Probability with Martingales*, Cambridge University Press, 1991.

[76] D. A. Zaev, *On the Monge–Kantorovich problem with additional linear constraints*, Mathematical Notes, 98 (2015), pp. 725–741.

Christiane Elgert

Curriculum Vitae

Personal

Date of birth 10th of April 1990

Place of birth Kyritz, Germany

Nationality German

Family status single

Education

since Oct. 2014 **Ph.D. Studies in Financial and Actuarial Mathematics**
Ph.D. thesis: *Theory of Distribution-Constrained Optimization Problems*, supervisor: Prof. Dr. Uwe Schmock, TU Wien, Vienna, Austria.

Oct. 2012 – Sept. 2014 **Master Program** in Mathematics (20% second subject: biology)
Master's thesis: Numerische Lösung der Diffusionsgleichung bei variabler räumlicher Struktur, University of Rostock, Germany.
overall grade 1.2

Oct. 2009 – Sep. 2012 **Bachelor Program** in Mathematics (20% second subject: biology)
Bachelor's thesis: Numerische Modellierung von Leberinfektionen mittels Reaktions-Diffusionsgleichung, University of Rostock, Germany.
overall grade 1.3

2009 **Allgemeine Hochschulreife** (university-entrance diploma)
Friedrich-Ludwig-Jahn-Gymnasium Kyritz, Germany.
overall grade 1.0

Professional Activities

Jul. 2016 – Feb. 2019 **Research Assistant**, *Jubiläumsfonds Project No. 16549*, "Aggregation of Credit Risk in Times of Financial Crises", TU Wien, Vienna, Austria.

Nov. 2014 – Feb. 2015 **University Assistant**, *Research Unit Financial and Actuarial Mathematics*, TU Wien, Vienna Austria.

Oct. 2014 – Jun. 2016 **Research Assistant**, *Project "Approximation von aggregierten Risiken unter Nebenbedingungen" sponsored by Arithmetica*, TU Wien, Austria, (reduced working time during employment as University Assistant and during Scholarship of DAAD).

Sep. 2013 – Mar. 2014 **Research Assistant**, *Institute of Computer Science, Systems Biology & Bioinformatics, University of Rostock, Germany.*

Aug. 2013 **Internship**, *Leibniz Institute for Farm Animal Biology (FBN), Institute of Genetics and Biometry, Dummerstorf, Germany.*

Nov. 2012 – Apr. 2013 **Student Assistant**, *Institute of Mathematics, University of Rostock, Germany.*

christiane.elgert@gmx.de

Scholarships

Mar. 2015 – Feb. 2016 **Scholarship by the DAAD**, *Jahresstipendien für Doktorandinnen und Doktoranden Studienjahr 2015/16.*

Publications

M. Beiglböck, M. Eder, C. Elgert, and U. Schmock. *Geometry of Distribution-Constrained Optimal Stopping Problems.* Probability Theory and Related Fields, 172 (2018), pp. 71-101. `https://doi.org/10.1007/s00440-017-0805-x`

in preparation C. Elgert, K. Hirhager, and U. Schmock. *Existence of an Optimal Strategy of Distribution-Constrained Discrete Optimization Problems.*

Invited Talks

Nov. 2018 *Distribution-constrained optimal stopping problems*, IFAM seminar, Institute for Financial and Actuarial Mathematics, University of Liverpool, UK.

Apr. 2017 *Adapted dependence by optimal stopping*, Dresden–Wien Workshop 2017, TU Dresden, Germany.

Conferences and Workshops

Jul. 2018 *VISS – Vienna International Summer School "Machine Learning Methods and Data Analytics in Risk and Insurance"*, TU Wien, Vienna, Austria.

Jul. 2017 *IME – Educational Workshop on Insurance: Mathematics and Economics*, TU Wien, Vienna, Austria.

Jul. 2017 *IME – 21th International Congress on Insurance: Mathematics and Economics*, TU Wien, Vienna, Austria.

Apr. 2017 *Dresden–Wien Workshop in Wahrscheinlichkeitstheorie, Statistik und Finanzmathematik*, TU Dresden, Germany.

Sep. 2016 *VCMF – Educational Workshop*, WU Vienna, Austria.

Sep. 2016 *VCMF – Vienna Congress on Mathematical Finance*, WU Vienna, Austria.

Jun. 2015 *IME – 19th International Congress on Insurance: Mathematics and Economics*, University of Liverpool, UK.

May 2015 *Advanced Modelling in Mathematical Finance – A conference in honour of Ernst Eberlein*, Kiel, Germany.

Apr. 2015 *Seminar "A Benchmark Approach to Investing, Pricing and Hedging" of the Actuarial Association of Austria (AVOe)*, Vienna, Austria.

Jan. 2015 *14th Winter school on Mathematical Finance – Special Topics: Nonlinear Pricing Dependence and Model Risk*, Lunteren, Netherlands.

Sep. 2014 *EAJ – 2nd European Actuarial Journal Conference*, TU Wien, Vienna, Austria.

Sep. 2014 *EAJ – Educational Workshop*, TU Wien, Vienna, Austria.

✉ christiane.elgert@gmx.de

Teaching

Mar. 2018 – Jun. 2018 *Assistant of Prof. Dr. Uwe Schmock for the Exercise in* Financial Mathematics I, Salzburg Institute of Actuarial Sciences, Austria.

Mar. 2015 – Jun. 2015 *Assistant of Prof. Dr. Uwe Schmock for the Exercise in* Financial Mathematics I, Salzburg Institute of Actuarial Sciences, Austria.

Nov. 2012 – Feb. 2013 *Exercise Classes*, University of Rostock, Germany.

since 2006 *Private Tutor*, Private teaching of mathematics, biology and physics for pupils (all classes) and students (different study fields).

Personal Skills

Programming Skills

R	good knowledge
Matlab	good knowledge
Maple	good knowledge
C, C++	basic knowledge
SPSS	basic knowledge

Language Skills

German	native
English	very good knowledge
Schwedish	basic knowledge
Latin	qualification

Vienna, February 10, 2019

christiane.elgert@gmx.de