Robust Pricing and Hedging Beyond one Marginal

Peter Spoida

University of Oxford

Lady Margaret Hall

A thesis submitted for the degree of Doctor of Philosophy

in the

Mathematical and Computational Finance Group
Mathematical Institute

Trinity 2014

Declaration of Authorship

I, Peter Spoida, declare that this thesis titled, 'Robust Pricing and Hedging Beyond one Marginal' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor Jan Obłój for the continuous support of my DPhil research and study. His patience, enthusiasm, and immense knowledge helped me at every stage of our work. I cannot imagine having a better advisor.

Besides my advisor, I am thankful to Pierre Henry-Labordère and Nizar Touzi for excellent collaboration on joint projects.

I thank Mathias Beiglböck, Samuel Cohen, Alexander Cox, Ben Hambly, Zhaoxu Hou, Marek Musiela, Harald Oberhauser and Johannes Ruf for numerous inspiring discussions.

Financial support from Oxford-Man Institute of Quantitative Finance, Deutscher Akademischer Austausch Dienst (DAAD) and the Engineering and Physical Sciences Research Council (EPSRC) is gratefully acknowledged.

Last but not the least, I would like to thank my parents and sister for supporting me throughout my life.

Abstract

The robust pricing and hedging approach in Mathematical Finance, pioneered by Hobson [50], makes statements about non-traded derivative contracts by imposing very little assumptions about the underlying financial model but directly using information contained in traded options, typically call or put option prices. These prices are informative about marginal distributions of the asset. Mathematically, the theory of Skorokhod embeddings provides one possibility to approach robust problems.

In this thesis we consider mostly robust pricing and hedging problems of Lookback options (options written on the terminal maximum of an asset) and Convex Vanilla Options (options written on the terminal value of an asset) and extend the analysis which is predominately found in the literature on robust problems by two features: firstly, options with multiple maturities are available for trading (mathematically this corresponds to multiple marginal constraints) and secondly, restrictions on the total realized variance of asset trajectories are imposed. Probabilistically, in both cases, we develop new optimal solutions to the Skorokhod embedding problem.

More precisely, in Part I we start by constructing an iterated Azéma-Yor type embedding – a solution to the n-marginal Skorokhod embedding problem, see Chapter 2. Subsequently, its implications are presented in Chapter 3. From a Mathematical Finance perspective we obtain explicitly the optimal superhedging strategy for Barrier/Lookback options. From a probability theory perspective, we find the *maximum maximum* of a martingale which is constrained by finitely many intermediate marginal laws. Further, as a by-product, we discover a new class of martingale inequalities for the terminal maximum of a càdlàg submartingale, see Chapter 4. These inequalities enable us to re-derive the sharp versions of Doob's inequalities. In Chapter 5 a different problem is solved. Motivated by the fact that in some markets both Vanilla and Barrier options with multiple maturities are traded, we characterize the set of market models in this case.

In Part II we incorporate the restriction that the total realized variance of every asset trajectory is bounded by a constant. This has been previously suggested by Mykland [68]. We further assume that finitely many put options with one fixed maturity are traded. After introducing the general framework in Chapter 6, we analyse the associated robust pricing and hedging problem for convex Vanilla and Lookback options in Chapters 7 and 8. Robust pricing is achieved through construction of appropriate Root solutions to the Skorokhod embedding problem. Robust hedging and pathwise duality are obtained by a careful development of dynamic pathwise superhedging strategies. Further, we characterize existence of market models with a suitable notion of arbitrage.

Contents

Chapter 1

Introduction

> Banks must explicitly assess the need for valuation adjustments to reflect two forms of model risk: the model risk associated with using a possibly incorrect valuation methodology; and the risk associated with using unobservable (and possibly incorrect) calibration parameters in the valuation model.
>
> ---
>
> Basel Committee on Banking Supervision
>
> `http://www.bis.org/publ/bcbs158.htm`, p. 29, December 2, 2012

The fundamental contributions of Black and Scholes [13] on option pricing laid the foundations of modern Mathematical Finance. Its implications on the academic community and the financial derivatives industry have been enormous. The paradigm is simple and powerful: choose your favourite probabilistic model, calibrate it and then use it to price and hedge financial derivatives contracts. This approach has been standard practice for a long time but has major shortfalls. Direct modelling can be difficult and markets often exhibit unexpected behaviour and models can miss these risks.

The *robust pricing and hedging methodology* tries to develop a *robust* alternative to the classical framework — a methodology which replaces model selection and calibration with an approach that is robust to a large class of outcomes. This is achieved by relaxing modelling assumptions on the one hand and taking into account directly different types of market information on the other hand. In its pure form, this robust methodology is (close to) a model-independent approach in the sense that pricing and hedging is done without (or with only minimal) assumptions about the future dynamics of the stock price but consistently with current market prices of liquidly traded instruments. In the simplest case hedging is achieved by (i) entering at time $t = 0$ into a static position consisting of liquid options and (ii) following a *simple* trading strategy in stocks. These strategies are referred to as *semi-static* hedging strategies and are "in adequacy with a familiar trader's quote: 'options are hedged with options'" (Henry-Labordère [48]).

From a theoretical perspective this viewpoint is not dissimilar to the uncertainty analysed by Keynes [60] or the *Knightian uncertainty* introduced by Knight [62], see also Hansen and Sargent [46, 47]. The fundamental idea is that any stochastic model $(\Omega, \mathbb{F}, \mathbb{P})$ which allows us to reason about risk, fails to capture uncertainty in the choice of the model itself. Hence, one is led to consider a class of models $\{\mathbb{P}_\alpha : \alpha \in \mathcal{I}\}$ simultaneously (here $\mathcal{I}$ is an index set) in order to accommodate for model uncertainty. Problems of this nature have been the subject of a lot of research recently in the probability/Mathematical Finance community, see e.g. the uncertain volatility approach of Lyons [63], the G-expectation of Peng [79] and the quasi-sure analysis of Denis and Martini [34] and Soner et al. [91].

To conclude, by considering a whole class of models, the robust methodology takes a conservative approach to pricing and hedging to protect against the worst-case scenario and has the additional advantage that "the traditional 'calibration business' in quantitative finance is no more required," (Henry-Labordère [48]) thus addressing both forms of model risk mentioned by the Basel Committee in the quotation above.

Organization of the Chapter In Section 1.1 we collect central results from the theory of Skorokhod embeddings which will be used throughout the thesis. Section 1.2 introduces and motivates the robust framework. With this formalism, we can formulate the research problems considered in this thesis in Section 1.3. These questions appear in a greater generality in the literature and we point out main links to it in Section 1.4. Finally, Section 1.5 discusses in which sense a robust framework is capable of mastering current challenges in risk-management and trading.

1.1 The Skorokhod Embedding Problem

Mathematically, the Skorokhod Embedding Problem (SEP) is the central pillar of the thesis. It can be stated formally as follows, cf. Skorokhod [90] and Obłój [73] for a survey with many relevant references. Denote by $\mathcal{B}(\mathbb{R})$ the Borel sigma-algebra of $\mathbb{R}$. Throughout the introduction B denotes a Brownian motion.

Problem 1.1.1 (SEP)**.** *For a given probability measure* μ *on* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *such that* $\int_{\mathbb{R}} |x| \, \mu(dx) < \infty$ *and* $\int_{\mathbb{R}} x\mu(dx) = 0$*, find a stopping time* τ *such that* $B_\tau \sim \mu$ *and* $(B_{\tau\wedge t} : t \geq 0)$ *is a uniformly integrable martingale.*

Various solutions to the SEP have been proposed, some with certain optimality properties. After reviewing briefly some potential theory which will allow us to characterize when there exists a solution to the SEP, we recall two classical solutions, the Azéma-Yor solution [4] and Root solution [85] which will play an important role in the thesis.

1.1.1 Potential Theory

For probability measures μ and ν on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with finite first moment we denote by $\preccurlyeq_{\mathrm{c}}$ the convex order, i.e.,

$$\nu \preccurlyeq_{\mathrm{c}} \mu \qquad \Longleftrightarrow \qquad \int G \mathrm{d}\nu \leq \int G \mathrm{d}\mu \qquad \forall\ G \text{ convex}. \tag{1.1.1}$$

The following result is central, cf. Strassen [92], Meyer [65, Chapter XI], see also Obłój [73, Proposition 8.1]).

Proposition 1.1.2. *Let ν and μ be probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with finite first moment. Suppose $(X_t)_{t\geq 0}$ is a continuous, real-valued local martingale with $\langle X\rangle_\infty = \infty$ a.s. and $X_0 \sim \nu$.*

Then for any probability measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ there exists a stopping time τ in the natural filtration of X such that $X_\tau \sim \mu$ and $(X_{\tau\wedge t})_{t\geq 0}$ is a uniformly integrable martingale if and only if $\nu \preccurlyeq_{\mathrm{c}} \mu$.

The functions U_μ, c_μ and p_μ which we define now will be useful in view of the convex order $\preccurlyeq_{\mathrm{c}}$. For a probability measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with finite first moment write

$$U_\mu(x) := -\int_{\mathbb{R}} |y - x|\, \mu(\mathrm{d}y), \quad x \in \mathbb{R}, \tag{1.1.2}$$

for its potential function. Further, denote the call and put option price with strike x and maturity T corresponding to the measure μ by $c_\mu(x)$ and $p_\mu(x)$, respectively, i.e.

$$c_\mu(x) = \int_{\mathbb{R}} (s - x)^+ \, \mu(\mathrm{d}s), \qquad p_\mu(x) = \int_{\mathbb{R}} (x - s)^+ \, \mu(\mathrm{d}s) \tag{1.1.3}$$

and note that

$$2(p_\mu(x) + m - x) = 2c_\mu(x) = -U_\mu(x) + m - x, \tag{1.1.4}$$

where $m := \int x\mu(\mathrm{d}x)$ is the mean of the measure μ.

1.1.2 The Azéma-Yor Solution

Let μ be a probability measure with finite first moment which is assumed centred, i.e. $\int x\mu(\mathrm{d}x) = 0$. Then Azéma-Yor solution to the SEP of μ, cf. Azéma and Yor [4], is given by

$$\tau_{\mathrm{AY}} := \inf\left\{t > 0 \ : \ b_\mu(B_t) \le \bar{B}_t\right\} \tag{1.1.5}$$

where $\bar{B}_t := \sup_{u \le t} B_u$ and b_μ is the barycenter function of μ defined as

$$b_\mu(x) := \frac{1}{\mu([x,\infty))} \int_x^\infty y\mu(\mathrm{d}y). \tag{1.1.6}$$

The Azéma-Yor solution has the property that it maximises the law of the maximum in stochastic order, i.e.

$$\mathbb{P}\left[\bar{B}_{\tau_{\mathrm{AY}}} \ge y\right] \ge \mathbb{P}\left[\bar{B}_\tau \ge y\right] \qquad \forall y \ge 0 \tag{1.1.7}$$

where τ is any solution to the SEP of μ.

1.1.3 The Root Solution

The Root solution to the SEP of μ, cf. Root [85], is given by

$$\tau_{\mathrm{R}} = \inf\left\{t \ge 0 \ : \ (t, B_t) \in R_\mu\right\} \tag{1.1.8}$$

where $R_\mu \subseteq \mathbb{R}_+ \times \mathbb{R}$ is a certain closed set. This solution has the *minimal residual expectation property* as shown by Rost [86], see also Cox and Wang [24],

$$\mathbb{E}\left[\int_{\tau_{\mathrm{R}}\wedge\Xi}^{\tau_{\mathrm{R}}} f(B_t)\mathrm{d}t\right] \leq \mathbb{E}\left[\int_{\tau\wedge\Xi}^{\tau} f(B_t)\mathrm{d}t\right] \qquad \forall \Xi \geq 0, \quad f > 0, \tag{1.1.9}$$

where τ is any solution to the SEP of μ. It can be shown that (1.1.9) is equivalent to

$$\mathbb{E}\left[G\left(\tau_{\mathrm{R}}\right)\right] \leq \mathbb{E}\left[G\left(\tau\right)\right] \qquad \forall G \text{ convex, non-decreasing.} \tag{1.1.10}$$

Hence, in particular τ_{R} is an embedding of μ with minimal variance.

In fact, Rost [86, 87] proved the above results for general Markov processes, something we will use in Chapters 7 and 8.

1.2 A Coherent Framework for Incorporation of Information

As indicated above, the classical modelling framework has several shortcomings. We point out some of them and illustrate how the robust framework addresses them. To fix the main ideas, we exhibit the robust methodology introduced by Hobson [50] and Brown et al. [16]. Subsequently, we begin to formalize an attempt to combine risk-neutral and real-world information. This will lead us to distinguish between three inputs: *Information*, *Beliefs* and *Rules* which are the building blocks of our "framework for quantifying modelling assumptions." (Jan Obłój[1])

[1]Labex Louis Bachelier SIAM SMAI Conference on Financial Mathematics Advanced Modeling and Numerical Methods, Paris, June 19, 2014

1.2.1 The Classical Modelling Framework and its Critique

"Classical" Mathematical Finance uses a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and an adapted stochastic process (S_t) to model the stock price. When trying to price a derivative written on the stock, the fundamental insight of Black and Scholes [13] is that pricing is done through hedging. Indeed, the fair price for a payoff G is equal to the capital required to synthesize its cashflow through trading. This approach is dominant in the financial derivatives industry. However, it has been criticized regarding the following three points, see Obłój [74].

1. Information contained in the market is largely ignored. This information can be for example prices of liquidly traded options. By calibration of model parameters this information can be absorbed partially into the model but possibly at the expense of introducing time-inconsistencies.

2. Very specific modelling assumptions are made, i.e. the full distribution of the process is fixed.

3. Idealised markets are assumed, i.e. transaction costs, liquidity risk, credit risk, etc. is ignored.

These shortfalls have been obviously addressed in many directions, however, largely focusing on points 2. and 3. by introducing market frictions, see Kabanov and Safarian [56], and model uncertainty, see references given above. As we demonstrate in the next section, the robust approach attempts to address in particular point 1. above by incorporating directly and consistently available market information.

1.2.2 The Robust Pricing and Hedging Methodology — Skorokhod Embedding Approach

The seminal work of Hobson [50] sets the course for the robust methodology. In the following we outline the key arguments which lead to the robust approach developed in those contributions. In particular, we explain the link to the SEP and how robust prices are devised by pathwise arguments. Our exposition follows Cox and Obłój [22]. For a survey with many relevant references we refer to Hobson [52].

Suppose our goal is to price and hedge an exotic option written on the spot price $(S_t)_{t\leq T}$ with maturity T. We assume that interest rates are zero and that we have access to call options for all strikes with maturity T written on S. Knowledge of all these prices is equivalent to knowledge of μ — the risk-neutral distribution of S_T. This fact goes back to Breeden and Litzenberger [15]: Differentiating the call option's price $c(K) = \int_K^\infty (s-K)\mu(\mathrm{d}s)$ twice with respect to the strike we obtain the risk-neutral density of S_T as

$$\mu(\mathrm{d}K) = c''(\mathrm{d}K) \tag{1.2.1}$$

where the second derivative c'' exists as a measure because c is convex.

The spot price process $(S_t)_{t\leq T}$ is a uniformly integrable martingale under a risk-neutral measure $\mathbb{Q}$. Assume that S is continuous. By Dambis [31] and Dubins and Schwarz [37] S may be written as a time-change of a Brownian motion,

$$S_t = B_{\tau_t} \quad \text{where} \quad \tau_t = \langle S \rangle_t, \qquad 0 \leq t \leq T. \tag{1.2.2}$$

The next observation is key: $S_T = B_{\tau_T}$ has distribution μ, or in other words, τ_T is a solution to the SEP.

Pricing Let us assume that the payoff of the exotic option, $O(S_t : t \leq T) = O(S)_T$, is invariant under time-changes, i.e.

$$O(S)_T = O(B)_{\tau_T}. \tag{1.2.3}$$

This is certainly true for barrier options because the maximum is not affected by the time-change. Let $\mathcal{P}[O]$ denote an arbitrage-free price for the payoff O. To avoid arbitrage opportunities we certainly need that

$$\mathcal{P}[O_1] \leq \mathcal{P}[O_2] \quad \text{whenever} \quad O_1(\omega) \leq O_2(\omega) \quad \forall \text{ continuous } \omega. \tag{1.2.4}$$

Together, it follows that

$$\mathcal{P}[O(S)_T] \leq \sup_{\mathbb{Q}} \mathbb{E}_{\mathbb{Q}}[O(S)_T] = \sup_{\tau} \mathbb{E}_{\mathbb{Q}_0}[O(B)_\tau], \tag{1.2.5}$$

where $\mathbb{Q}$ is a martingale measure such that $S_T \sim_{\mathbb{Q}} \mu$, τ is a solution to the SEP of μ and B is a $\mathbb{Q}_0$-Brownian Motion starting at S_0.

To conclude, robust pricing boils down to solving a SEP with optimality properties.

Robust Hedging We want the upper price bound (1.2.5) to be enforced by a pathwise superhedging strategy of the form

$$O(S)_T \leq H(S_T) + N_T \qquad \text{pathwise}. \tag{1.2.6}$$

The superhedging strategy (1.2.6) consists of two components, a *static* component $H(S_T)$ and a *dynamic* component $(N_t)_{t \leq T}$. Using the Breeden-Litzenberger formula one can show that the static position can be synthesised from traded call options. As for the dynamic component, we want it to be meaningful as a self-financing trading strategy in the spot S, i.e.

$$N_T(\omega) = \int_0^T \phi_t(\omega)\mathrm{d}S_t(\omega),$$

for some appropriate, progressively measurable trading strategy ϕ, e.g. ϕ being simple or of bounded variation, see e.g. Section 6.3 for more details.

To conclude, robust hedging should be designed as cheaply as possible. Further, if there exists a market model which attains pathwise equality in (1.2.6) then this implies that the superhedging strategy is optimal.

An Illustrative Example of the Robust Approach – The One-Touch Digital Barrier Option Let us illustrate the underlying ideas of the robust approach with the example of an one-touch digital barrier option. We follow Brown et al. [16]. Denote $\bar{S}_t := \sup_{u\leq t} S_u$. This option has a payoff at maturity T given by

$$O(S)_T = O(S_t : t \leq T) = \mathbb{1}_{\{\bar{S}_T \geq y\}}$$

where $y > S_0$ is some barrier level. This payoff is invariant under time-changes (recall that S is assumed continuous), i.e. $O(S)_T = O(B)_{\tau_T}$. From market traded options we know that $S_T \sim \mu$.

Firstly, let us look at the upper price bound. This boils down to computing the upper bound in (1.2.5). As mentioned in Section 1.1.2, the Azéma-Yor solution τ_{AZ} to the SEP has the following optimality property:

$$\mathbb{P}\left[\sup_{t\leq\tau} B_t \geq y\right] \leq \mathbb{P}\left[\sup_{t\leq\tau_{\mathrm{AZ}}} B_t \geq y\right] \tag{1.2.7}$$

where τ is any solution to the SEP. Hence, in view of (1.2.5), τ_{AZ} imposes the upper bound on the price of a one-touch digital barrier option.

Secondly, let us determine a robust superhedging strategy. It can be verified directly that the following inequality holds pathwise:

$$\mathbb{1}_{\{\bar{S}_T \geq y\}} \leq \frac{(S_T - \zeta)^+}{y - \zeta} + \frac{y - S_T}{y - \zeta} \mathbb{1}_{\{\bar{S}_T \geq y\}} =: F(S_T) + N_T \tag{1.2.8}$$

where $0 < \zeta < y$. The inequality (1.2.8) provides a simple robust superhedging strategy for the one-touch digital barrier option. It consists of an initial portfolio of $\frac{1}{y-\zeta}$ call options with strike ζ and if the asset rises to y we sell $\frac{1}{y-\zeta}$ forwards on S. Note that the first part of the hedge is static whereas the second part is semi-static. In addition the optimal ζ can be determined explicitly.

Finally, it can be shown that the minimal initial cost of (1.2.8) coincides with the upper price bound obtained by (1.2.7). Hence the superhedge (1.2.8) is optimal. Seeing hedging as dual to pricing, by the above results there is no duality gap, i.e. the upper price bound coincides with the price of the optimal superhedge (see below for more details).

1.2.3 Gap Between Market and Statistical Information

By the principle of risk neutral valuation the price of a financial derivative on S is obtained by the expectation of the (discounted) payoff under some risk-neutral measure $\mathbb{Q}$. Hence, current market prices relate to the future dynamics of S under $\mathbb{Q}$. They are modelled using probability theory. In contrast, classical econometric theory is concerned with time series data of S under the physical measure $\mathbb{P}$. This corresponds to the past dynamics of S under $\mathbb{P}$ and is analysed using statistical theory.

In general, it is unclear how to relate information measured under $\mathbb{P}$ to that measured under $\mathbb{Q}$. This is the fundamental reason why classical Mathematical Finance can make only limited use of time series data.

We make the observation that the quadratic variation of a semimartingale is invariant under an equivalent change of probability measure. In addition, estimation of realized variance has been extensively studied in econometrics and statistics, see e.g. Andersen et al. [3] or Barndorff-Nielsen and Shephard [6]. Then the idea is to use one's favourite estimation method to come up with an upper bound on realized variance, Ξ say, and incorporate this bound together with the other market constraints (as before) to compute a range of possible prices for an exotic option. This additional "view" on realized variance should yield tighter price bounds and pathwise closer superhedging strategies and it has a clear financial interpretation.

Next we describe how we will use the Robust Pricing and Hedging framework to allow for incorporation of these types of information.

1.2.4 Information, Beliefs and Rules

The framework we will use is based, implicitly or explicitly, on work done by several authors, e.g. Hobson [50]; Brown et al. [16]; Cox and Obłój [28]; Mykland [68, 69, 70]; Henry-Labordère [48]. We describe its building blocks and the reasoning behind.

The ultimate goal is to arrive at prices for exotic derivatives which are consistent with prices of observed instruments, the market mechanism and do not introduce arbitrage. The precise notion of arbitrage will depend on the particular setting and may be a *model-independent arbitrage*, see e.g. Davis and Hobson [33] and Cox and Obłój [28], or some weaker form of it. Note that we do not exclude arbitrage a priori, but rather want to use the robust framework to exploit arbitrage if it exists. In particular, if the market inputs are inconsistent with the specified form of arbitrage, then the framework shall allow us to compute a strategy to exploit this opportunity.

The three building blocks or model inputs are *Information, Beliefs* and *Rules.*

Information refers to market instruments which are observed in the market and which can be traded in at time $t = 0$ for a known price. In practice these will be liquidly traded instruments whose prices are considered as "unambiguous" by the economic principle of supply and demand. Here we are not making any assumption on the possibility to trade in them at a future point in time. A typical example is a call option written on a stock which can be traded now.

Beliefs allow us to incorporate generic properties of the future dynamics of the stock price process. Mathematically we will capture this by restricting the support of the model, i.e. imposing that every asset trajectory belongs to a path space $\mathfrak{P}$ which is a subset of càdlàg paths. This is a consistent way to do it because the robust framework a priori does not assume any probability measure. Part II of the thesis considers the restriction to paths which respect the realized variance constraint $\langle \log(\omega) \rangle_T \leq \Xi$. Another plausible example — in a multi-asset setup — is the restriction to paths (ω^1, ω^2) satisfying $\omega^1_T - \beta\omega^2_T \in [-1, 1]$ which is a certain cointegration relation.

We remark that *Beliefs* allow us to interpolate between a model-specific (e.g. Black and Scholes setting) and model independent setting (i.e. any càdlàg trajectory is possible) by varying the path space from paths which satisfy $\langle \log(\omega) \rangle_t = \sigma^2 t$ for all $t \geq 0$ to all càdlàg paths.

Rules state properties and mechanisms of the market. We include this aspect into the robust framework to demonstrate the flexibility of the framework, but in the following we will assume throughout that markets are frictionless, i.e. there are no transaction costs (no bid-ask spreads), no trading restrictions, trading is executed instantaneously and arbitrary amounts of the asset can be traded.

Having specified *Information, Beliefs* and *Rules*, absence of arbitrage considerations produce robust prices and hedges.

1.2.5 Discussion

Evidently, the relaxation of modelling assumptions results in less precise answers to the problem at hand. More concretely, rather than obtaining a unique arbitrage-free price for an exotic option, the robust methodology provides an interval of arbitrage-free prices. Whether or not these bounds are tight enough depends on the problem, but by enlarging the set of hedging instruments one hopes to decrease them.

The described methodology based on the SEP is not universal in the simple form described above because for e.g. Asian-type options the time-change approach fails. More precisely, assume one is interested in an Asian option paying

$$\int_0^T f(S_u)\mathrm{d}s$$

at maturity T. We observe,

$$\mathcal{O}(S)_T = \int_0^T f(S_u)\mathrm{d}s = \int_0^T f(W_{\rho_u})\mathrm{d}u \neq \int_0^{\rho_T} f(W_u)\mathrm{d}u = \mathcal{O}(B)_{\rho_T}$$

where $\rho_u = \langle S \rangle_u$, and hence equality in (1.2.5) is not valid for general f and the approach to robust pricing presented above fails.

As for robust hedging, classically, the superhedging strategy (1.2.6) has been guessed barehanded. Although the SEP approach can provide some insight into the extremal model the hedge has to be designed relative to, it does not provide a systematic machinery how to do this in practice.

1.3 Research: Going Beyond one Marginal

The theme of this thesis can be described as an attempt to go beyond a robust framework which considers one marginal constraints only.

Using the formalism introduced above, the robust framework which has been predominantly considered in the literature can be described by the following specifications.

Information:	call option prices for one fixed maturity and all strikes
Beliefs:	price trajectories are continuous
Rules:	no market frictions

We are aware of only a few exceptions which go beyond this, e.g. Brown et al. [18] and Henry-Labordère [48].

Our goal is to use the model inputs *Information* and *Beliefs* to incorporate more knowledge and impose generic properties on the future dynamics.

1.3.1 Part I: Multiple Marginals

In Part I we consider the robust problem with multiple marginals, i.e. we use the following robust framework.

Information:	call option prices for maturities $t_1 < \cdots < t_n = T$ and all strikes
Beliefs:	price trajectories are continuous
Rules:	no market frictions

In Chapter 3 we arrive at robust prices and hedges for the simple barrier option $\mathbb{1}_{\{\sup_{t\leq T} S_t \geq B\}}$ by obtaining a robust superhedging strategy for simple barrier options consisting of a portfolio of calls with different maturities and a self-financing trading strategy. Probabilistically, we find the "*maximum maximum* of a martingale constrained by finitely many intermediate marginal laws"[2].

In Chapter 2, under some additional assumption, we construct an extremal model — a solution to the SEP in the n-marginal case — hence demonstrating that our superhedge is the cheapest possible in that case.

In Chapter 4 we analyse a class of martingale inequalities for the maximum which come out of the results of Chapter 3.

In Chapter 5 we present a characterization of market models which are consistent with both a finite number of Vanilla and barrier option prices of multiple maturities. Here the reasoning is that in certain markets there are also liquidly traded barrier options.

1.3.2 Part II: Realized Variance Constraints

In a series of articles Mykland [67, 68, 69, 70] proposes to constrain the realized variance of price trajectories. Inspired by this, in Part II of the thesis we incorporate both a constraint on realized variance and finitely many Vanilla option prices. We use the following robust framework, see Chapter 6 for a more formal

[2] phrasing adapted from the title of Brown et al. [18]

description.

Information:	put option prices for one fixed maturity and strikes $K_1, \ldots, K_n$
Beliefs:	price trajectories are continuous,
	$\langle \log(\omega) \rangle_T \leq \Xi$ and $t \mapsto \langle \log(\omega) \rangle_t$ is strictly increasing
Rules:	no market frictions

In Chapters 7 and 8 we obtain robust prices and hedges in this framework for convex Vanilla options $G(S_T)$, i.e. G convex, and Lookback options $G\left(\sup_{t\leq T} S_t\right)$, G non-decreasing.

More precisely, in a first step, we rediscover the (unpublished) results of Mykland [67] for convex Vanilla options with techniques from the theory of Skorokhod embeddings. In subsequent steps we extend the analysis to Lookback options, describe pathwise hedging strategies and investigate forms of arbitrage in this setting.

1.4 A Wider Perspective

The questions of Section 1.3 are special cases of more fundamental problems which we outline now. Finding answers to these questions in full generality is currently an active research field.

1.4.1 Link to Martingale Optimal Transport

In a slightly different technical formulation the class of robust pricing problems we consider can be formulated generally as an *martingale optimal transportation* *problem*,

$$\sup_{\mathbb{P}\in\mathcal{P}(\boldsymbol{\mu})} \mathbb{E}_{\mathbb{P}}[G] \tag{1.4.1}$$

where $\mathcal{P}(\boldsymbol{\mu})$ is a class of martingale measures $\mathbb{P}$ which satisfy $S_{t_i} \sim_{\mathbb{P}} \mu_i,\ i = 1,\dots,n$, and G is some payoff function. Problems of this type have been analysed by e.g. Beiglböck et al. [10]; Galichon et al. [44]; Tan and Touzi [93]. Roughly speaking, the objective is to transport the Dirac measure $\delta_{\{S_0\}}$ optimally along the marginals $\mu_1,\dots,\mu_n$ where this transportation has to respect the martingale nature of the process S. Each transport plan, here described by a measure $\mathbb{P} \in \mathcal{P}(\boldsymbol{\mu})$, is associated with a cost $\mathbb{E}_{\mathbb{P}}[G]$ given by the path-dependent payoff G.

1.4.2 Duality

We have reasoned in Section 1.2.2 that it makes sense to derive prices from pathwise inequalities. Several authors have studied the question whether *the supremum over all market models of the expected option's payoff equals the minimal cost of superreplicating this payoff*. This problem is referred to as *Duality problem* and we are deliberately imprecise about its formulation. In fact, there are multiple formalizations of this problem and their approaches rely on different techniques. In a continuous time setup, Galichon et al. [44] frame the superhedging problem in the *quasi-sure* framework and prove duality by using regular conditional probability distributions. Dolinsky and Soner [35, 36] use a pathwise superhedging notion and approximate the continuous time model by discrete time models. In discrete time, it turns out to be convenient to apply results from the classical theory of optimal transport, in particular Monge-Kantorovich duality, cf. Villani [95]. This perspective yields an elegant proof of duality as shown by Beiglböck et al. [10].

Let us note that for specific payoffs duality results have been proven (explicitly or implicitly) in the works on Robust Pricing and Hedging, cf. e.g. Brown et al. [16], Cox and Obłój [28, 22], Davis et al. [32], Cox and Wang [24].

One application of *Duality* is in the context of numerical methods. By solving a (semi-infinite) linear programming problem Davis et al. [32] and Henry-Labordère [48] demonstrate how to compute robust prices, superhedging strategies and potential arbitrage strategies.

1.4.3 Arbitrage and Robust Fundamental Theorem of Asset Pricing

The reader might have noticed that the appropriate notion of arbitrage will depend on the exact specification of the robust framework in terms of the *Information, Beliefs* and *Rules*. One might want to characterize arbitrage in terms of these model inputs only which might then be termed *Robust Fundamental Theorem of Asset Pricing*.

In particular settings, i.e. for certain choices of *Information, Beliefs* and *Rules* and for specific exotic options, characterizations have been obtained by e.g. Davis and Hobson [33], Cox and Obłój [28, 22], Davis et al. [32]. A general result, in particular a suitable notion of arbitrage, is still awaited, see however Acciaio et al. [1].

1.5 Automated Option Market Making

Finally, we discuss whether the robust framework incorporates practical features which are requested by modern risk-management and trading applications.

Carr [19] raises the question "how to automate the pricing and risk management of derivative securities", i.e. "given the current market price of the underlying asset and also given market option quotes at several given strikes and terms, provide options quotes at any strike and term in a specified set". He also specifies a list of desirable properties a black-box which solves this task should have. We contrast this with the robust framework.

1.5.1 "A Partial List of Desirable Properties"

According to Carr [19] desirable properties of an automated option market making black-box should include at least the following.

Properties regarding inputs

I1: "ability to accept related inputs eg. American options, variance swaps, CDS, options on other underlyings"

I2: "well-posedness, i.e. small changes in inputs lead only to small changes in output"

I3: "parsimony"

Properties regarding outputs

O1: "no model-free arbitrages in the output"

O2: price "consistency with given mid-prices"

O3: "ability to uniquely and accurately price related derivatives such as American options, variance swaps, barrier options, and other exotics"

O4: ability to "produce deltas, gammas, thetas"

O5: "resonance with financial orthodoxy e.g. non-negative gammas"

O6: "implied risk-neutral dynamics not too far from observable $\mathbb{P}$ dynamics"

Properties regarding numerics

N1: "real-time computational speed"

N2: "numerical robustness"

1.5.2 Suitability of Robust Framework

The robust methodology constitutes in many respects a different way of thinking about pricing and hedging of financial derivatives. Whereas properties I1, O1–O3 are met by construction, for other properties such as I3 or O4 it is not entirely clear what the counterpart or meaning in the robust framework is.

Major Differences As for O6, since the robust methodology is not assuming a model a priori, one could argue that there are no implied risk-neutral dynamics. Therefore, on the one hand, one is led to conclude that the robust framework is parsimonious in its model-inputs (information, beliefs and rules).

On the other hand robust methodology allows for an uncountable number of classical models and in a sense identifies an *extremal* model. Pricing and hedging are built relative to this model and therefore correspond to "extremal Deltas[3]". Other "Greeks" are ignored because in robust hedging as introduced here the paradigm is to make use of *static* positions in options. In some cases, this extremal model could be termed *degenerate* because of its artificial structure, e.g. one-to-one relation of spot and its running maximum in the robust problem for the barrier

[3]Delta refers to the classical terminology in mathematical derivatives pricing theory.

option. This issue has been addressed by Henry-Labordère [48] who proposes to *believe* that any model can deviate only to a certain degree from a reference model (in relative entropy terms).

Potential Strengths The most evident advantage of the robust methodology is that it directly takes into account market information and trading rules. In markets where there is low liquidity or trading restrictions (either explicit though e.g. short selling constraints or implicit through e.g. high transaction costs) the robust approach could be effective because of its ability to control the amount of trading.

We are not aware of a general answer to I2 and do not investigate this question here, but point out that in the special setup of Chapters 6–8 model outputs are continuous in model inputs.

Henry-Labordère [48] proposes to apply the robust methodology on a portfolio level instead of on a derivative-by-derivative level. The rationale behind this is that netting will result in "cancellations of model risk" and hence a reasonably precise statement about the riskiness of a whole portfolio. This perspective is further underpinned by the interpretation of the robust price as a coherent risk measure as pointed out by Cont [21]. More precisely, for a random variable X representing the value of a portfolio, it can be shown that

$$\rho(X) = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}} \left[-X \right], \tag{1.5.1}$$

where $\mathcal{P}$ is a class of probability measures, defines a coherent risk measure, see e.g. Cont [21].

Numerically, these problems could be (and have already been) tackled through a (semi-infinite) linear programming problem, see Davis et al. [32], Henry-Labordère [48] or Piterbarg [82]. Indeed, in principle after discretising all assets in time and space the linear program searches over martingale measures which satisfy a number of constraints and maximize an objective function. Market options' constraints are linear by linearity of the expectation. Path-constraints specify null-sets and therefore also can be formulated as linear constraints. The objective is then to maximize the expectation of an exotic option's payoff, say. If no martingale measure exists, then the dual problem readily identifies an arbitrage.

Although the robust framework as we presented it is static in the sense that we optimize as of today for the best superhedging strategy, a natural and straight-forward extension is to periodically "reoptimize" for the new optimal superhedge and rebalance accordingly if it is cheaper.

To conclude, in its current form, the robust framework resembles more a risk management tool than a live trading tool.

Part I

Multiple Marginal Constraints — Incorporating Options with Multiple Maturities

Chapter 2

An Iterated Azéma-Yor Type Embedding for Finitely Many Marginals

Joint Publication Large parts of this chapter are submitted as part of a joint publication with Jan Obłój, see Obłój and Spoida [75].

The subject of Chapter 2 is the iterated Azéma-Yor type embedding as we call it. It is an explicit solution to an n-marginal Skorokhod embedding problem (SEP) with the desirable optimal property that it maximises the distribution of the maximum in stochastic order among all embeddings. In Chapter 3 implications of this embedding for robust pricing and hedging will be explained.

To describe the problem we consider, take a standard Brownian motion B and a sequence of probability measures $\mu_1, \dots, \mu_n$. A solution to the n-marginal SEP is a sequence of stopping times $\tau_1 \leq \cdots \leq \tau_n$ such that $B_{\tau_i} \sim \mu_i$, $1 \leq i \leq n$, and $(B_{t \wedge \tau_n})_{t \geq 0}$ is a uniformly integrable martingale. It follows from Jensen's inequality that a solution may exist only if all μ_i are centred and the sequence is in convex order. And then it is easy to see how to solve the problem: it suffices to iterate a solution to the classical case $n = 1$ developed for a non-trivial initial distribution of B_0, of which several exist.

In contrast, the question of optimality is much more involved. In general there is no guarantee that a simple iteration of optimal embeddings would be globally optimal. Indeed, this is usually not the case. Consider the embedding of Azéma and Yor [4] which consists of a first exit time for the joint process $(B_t, \bar{B}_t)_{t\geq 0}$, where $\bar{B}_t = \sup_{s\leq t} B_s$. More precisely, their solution $\tau^{\mathrm{AY}} = \inf\left\{t \geq 0 : B_t \leq \xi_\mu(\bar{B}_t)\right\}$ leads to a functional relation $B_{\tau^{\mathrm{AY}}} = \xi_\mu(\bar{B}_{\tau^{\mathrm{AY}}})$. This then translates into the optimal property that the distribution of $\bar{B}_{\tau^{\mathrm{AY}}}$ is maximized in stochastic order amongst all solutions to SEP for μ, i.e. for all y,

$$\mathbb{P}\left[\bar{B}_{\tau^{\mathrm{AY}}} \geq y\right] = \sup\left\{\mathbb{P}\left[\bar{B}_\rho \geq y\right] : \rho \text{ s.t. } B_\rho \sim \mu, (B_{t\wedge\rho}) \text{ is UI } \right\}.$$

It is not hard to generalise the Azéma-Yor embedding to a non-trivial starting law, see Obłój [73, Sec. 5]. Consequently we can find η_i such that $\tau_i = \inf\left\{t \geq \tau_{i-1} : B_t \leq \eta_i(\sup_{\tau_{i-1}\leq s\leq t} B_s)\right\}$ solve the n-marginal SEP. However this construction will maximise stochastically the distributions of $\sup_{\tau_{i-1}\leq t\leq \tau_i} B_t$, for each $1 \leq i \leq n$, but not of the global maximum $\bar{B}_{\tau_n}$. The latter is achieved with a new solution which we develop here.

Our construction involves an interplay between all n-marginals and hence is not an iteration of a one-marginal solution. However it preserves the spirit of the Azéma-Yor embedding in the following sense. Each τ_i is still a first exit for $(B_t, \bar{B}_t)_{t\geq\tau_{i-1}}$ which is designed in such a way as to obtain a "strong relation" between B_{τ_i} and $\bar{B}_{\tau_i}$, ideally a functional relation. Under our technical assumption about the measures $\mu_1, \ldots, \mu_n$, Assumption $\circledast$, we describe this relation in detail in Lemma 2.2.1.

For $n = 2$ we recover the results of Brown et al. [17]. We also recover the trivial case $\tau_i = \tau^{AY}_{\mu_i}$ which happens when $\xi_{\mu_i} \leq \xi_{\mu_{i+1}}$, we refer to Madan and

Yor [64] who in particular then investigate properties of the arising time-changed process. However, as a counterexample shows, our construction does not work for all laws $\mu_1, \ldots, \mu_n$ which are in convex order. Assumption $\circledast$ fails when a special interdependence between the marginals is present and the analysis then becomes more technical and the resulting quantities are, in a way, less explicit. We only detail the appropriate arguments for the case $n = 3$.

Organisation of the Chapter In Section 2.1 we explain the main quantities for the embedding and state the main result. We also present the restriction on the measures $\mu_1, \ldots, \mu_n$ which we require for our construction to work (Assumption $\circledast$). In Section 2.2 we prove the main result and Section 2.3 provides a discussion of extensions together with comments on Assumption $\circledast$. The proof of an important but technical lemma is relegated to the appendix. We conclude with a numerical example.

Organisation of First Part of the Thesis After having obtained the iterated Azéma-Yor type embedding in Chapter 2, we investigate its optimal properties and implications for robust pricing and hedging of Lookback options in Chapter 3. Motivated by these results we obtain a class of martingale inequalities for the terminal maximum $\bar{X}_{t_n}$ of a càdlàg submartingale X which features information of the process at intermediate times $t_1, \ldots, t_n$. Finally, in Chapter 5 we consider the problem of calibration to both finitely many Vanilla and Barrier options of different maturities.

2.1 Main Result

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F} = (\mathcal{F}_t)$, be a filtered probability space satisfying the usual hypotheses and B a continuous $\mathbb{F}$–local martingle, $B_0 = 0$, $\langle B \rangle_\infty = \infty$ a.s. and

B has no intevals of constancy a.s. We denote $\bar{B}_t := \sup_{s\leq t} B_t$. We are primarily interested in the case when B is a standard Brownian motion and it is convenient to keep this example in mind, hence the notation. We allow for more generality as this introduces no changes to the statements or the proofs.

2.1.1 Definitions

The following definition will be crucial in the remainder of the article. We define the stopping boundaries $\xi_1, \ldots, \xi_n$ for our iterated Azéma-Yor type embedding together with quantities $K_1, \ldots, K_n$ which will be later linked to the law of the maximum at subsequent stopping times.

Definition 2.1.1. *Fix $n \in \mathbb{N}$. For convenience we set*

$$c_0 \equiv 0, \quad K_0 \equiv 0, \quad \xi_0 \equiv -\infty. \tag{2.1.1}$$

For $\zeta \in \mathbb{R}$ and $i = 1, \ldots, n$ we write

$$c_i(\zeta) := \int_{\mathbb{R}} (x-\zeta)^+ \, \mu_i(\mathrm{d}x). \tag{2.1.2}$$

Let $y \geq 0$ and assume that for $i = 1, \ldots, n-1$ the quantities $\xi_i, K_i, \imath_i$ and $\jmath_i$ are already defined. Then we define

$$\begin{aligned} &\imath_n(\cdot; y) : (-\infty, y] \to \{0, 1, \ldots, n-1\}, \\ &\zeta \mapsto \imath_n(\zeta; y) := \max\{k \in \{0, 1, \ldots, n-1\} : \xi_k(y) < \zeta\}, \end{aligned} \tag{2.1.3}$$

and

$$\xi_n(y) := \sup\left\{\operatorname*{arg\,inf}_{\zeta<y}\left(\frac{c_n(\zeta)}{y-\zeta} - \left[\frac{c_{\imath_n(\zeta;y)}(\zeta)}{y-\zeta} - K_{\imath_n(\zeta;y)}(y)\right]\right)\right\}. \tag{2.1.4}$$

With

$$\jmath_n(y) := \imath_n(\xi_n(y); y) \tag{2.1.5}$$

we set

$$K_n(y) := \frac{1}{y - \xi_n(y)} \left\{ c_n(\xi_n(y)) - \left[c_{\jmath_n(y)}(\xi_n(y)) - (y - \xi_n(y)) K_{\jmath_n(y)}(y) \right] \right\}. \tag{2.1.6}$$

Definition 2.1.2 (Embedding)**.** *Set $\tau_0 \equiv 0$ and for $i = 1, \dots, n$ define*

$$\tau_i := \begin{cases} \inf \left\{ t \geq \tau_{i-1} : B_t \leq \xi_i(\bar{B}_t) \right\} & \textit{if } B_{\tau_{i-1}} > \xi_i(\bar{B}_{\tau_{i-1}}), \quad (2.1.7a) \\ \tau_{i-1} & \textit{else.} \quad (2.1.7b) \end{cases}$$

Figure 2.1.1 illustrates a set of possible stopping boundaries ξ_1, ξ_2, ξ_3 in the case of $n = 3$. If Assumption ⊛ is in place, see Section 2.1.2, we will show that the stopping boundaries are continuous (except possibly for $i = 1$) and non-decreasing, cf. Section 2.1.5.

The n^{th} stopping boundary ξ_n is obtained from an optimization problem which features $\xi_1, \dots, \xi_{n-1}$ and $K_1, \dots, K_{n-1}$. $K_n(y)$ is the value of the objective function at the optimal value $\xi_n(y)$. Note that all previously defined stopping boundaries $\xi_1, \dots, \xi_{n-1}$ and the quantities $K_1, \dots, K_{n-1}$ remain unchanged.

Denote the right and left endpoints of the support of the measure μ_i by

$$r_{\mu_i} := \inf \left\{ x : \mu_i \left((x, \infty) \right) = 0 \right\}, \qquad l_{\mu_i} := \sup \left\{ x : \mu_i \left([x, \infty) \right) = 1 \right\}, \tag{2.1.8}$$

respectively, and the barycentre function of μ_i by

$$b_i(x) := \frac{\int_{[x,\infty)} u \mathrm{d}\mu_i(u)}{\mu_i([x,\infty))} \mathbb{1}_{\{x < r_{\mu_i}\}} + x \mathbb{1}_{\{x \geq r_{\mu_i}\}}, \tag{2.1.9}$$

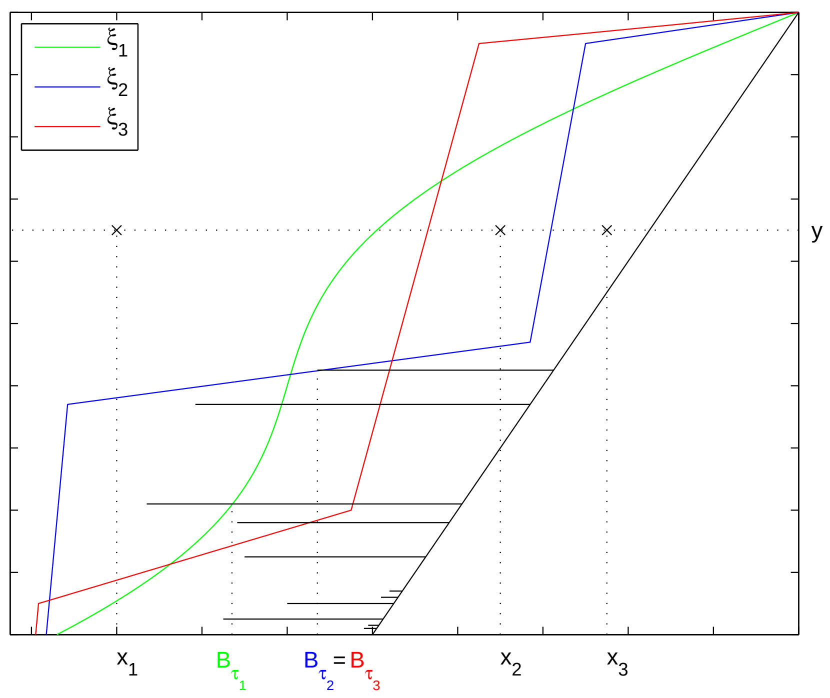

FIGURE 2.1.1: We illustrate possible stopping boundaries ξ_1, ξ_2, ξ_3. The horizontal lines represent a sample path of the process $\left(B_t, \bar{B}_t\right)$ where the x-axis is the value of B and the y-axis the value of $\bar{B}$. Each horizontal segment is an excursion of B away from its maximum $\bar{B}$. According to the definition of the embedding, the first stopping time τ_1 is found when the process first hits ξ_1. Since $\xi_1(\bar{B}_{\tau_1}) > \xi_2(\bar{B}_{\tau_1})$ the process continues and targets ξ_2. The stopping time τ_2 is found when the process first hits ξ_2. Since $\xi_2(\bar{B}_{\tau_2}) \leq \xi_3(\bar{B}_{\tau_2})$ we get $\tau_3 = \tau_2$. For the y we fixed we have $\imath_3(x_1; y) = 0, \imath_3(x_2, y) = 1, \imath_3(x_3; y) = 2$.

As shown by Brown et al. [17], the right-continuous inverse of b_i, denoted by b_i^{-1}, can be represented as

$$b_i^{-1}(y) = \sup \left\{ \arg\inf_{\zeta < y} \frac{c_i(\zeta)}{y - \zeta} \right\}. \tag{2.1.10}$$

It is clear and has been studied in more detail by Madan and Yor [64] that if the sequence of barycentre functions is increasing in i, then the intermediate law constraints do not have an impact on the corresponding iterated Azéma-Yor embedding. However, in general the barycentre functions will not be increasing in i, cf. Brown et al. [18], and hence will affect the embedding. We think of $\jmath_n(y)$ as the index of the last law $\mu_i, i < n$, which represents, locally at the level of the maximum y, a *binding constraint* for the embedding. As compared to the optimization from which b_n^{-1} is obtained, cf. (2.1.10), the optimization from which ξ_n is obtained, cf. (2.1.4), has a penalty term.

2.1.2 Restrictions on Measures

Throughout the article we will denote the left- and right-limit of a function f at x (if it exists) by $f(x-)$ and $f(x+)$, respectively. Recalling the conventions in (2.1.1), we define inductively for $n \in \mathbb{N}$ and $y \geq 0$ the mappings

$$\begin{aligned} &c^n(\cdot, y) : (-\infty, y] \to \mathbb{R} \cup \{\infty\}, \\ &x \mapsto c^n(x, y) := c_n(x) - \left[c_{\imath_n(x;y)}(x) - (y - x)K_{\imath_n(x;y)}(y)\right]. \end{aligned} \tag{2.1.11}$$

It follows that the minimization problem in (2.1.4) is equivalent to the following minimization problem,

$$\xi_n(y) \in \arg\min_{\zeta \leq y} \frac{c^n(\zeta, y)}{y - \zeta}, \tag{2.1.12}$$

where we observe that

$$\begin{aligned} \left.\frac{c^n(\zeta, y)}{y - \zeta}\right|_{\zeta = y} &:= \lim_{\zeta \uparrow y} \frac{c^n(\zeta, y)}{y - \zeta} \\ &= \begin{cases} -c_n'(y-) + c_{\imath_n(y;y)}'(y-) + K_{\imath_n(y;y)}(y) & \text{if } c_n(y) = c_{\imath_n(y,y)}(y), \\ +\infty & \text{else.} \end{cases} \end{aligned} \tag{2.1.13}$$

Now we want to argue that there exists a minimizer the optimization problem (2.1.12). In the case $y > 0$ this can be deduced iteratively from the – a priori – piecewise continuity of $c^n(\cdot, y)$ and the fact that $c^n \geq 0$ together with the property that $\zeta \mapsto \frac{c^n(\zeta,y)}{y-\zeta} = \frac{c_n(\zeta)}{y-\zeta}$ for ζ sufficiently small, which is a non-increasing function. This is because, inductively, $\xi_1(y), \ldots, \xi_{n-1}(y)$ are finite and fixed and hence $\imath_n(\zeta; y) = 0$ for $\zeta < \min_{i<n} \xi_i(y)$.

For $y \geq 0$, we extend

$$\left.\frac{c^n(\zeta, y)}{y-\zeta}\right|_{\zeta=l_{\mu_n}} := \begin{cases} \frac{-l_{\mu_n}}{y-l_{\mu_n}}, & \text{if } l_{\mu_n} > -\infty, \\ 1 & \text{else.} \end{cases} \tag{2.1.14}$$

For later use observe

$$\min_{i \leq n} b_i^{-1}(y) \leq \xi_n(y) \leq y \tag{2.1.15}$$

which follows from the definition of ξ_n, cf. (2.1.4), and where b_i^{-1} denotes the right-continuous inverse of the barycentre function b_i, cf. (2.1.10).

Assumption $\circledast$ (Restriction on Measures)**.** *Recall definitions in* (2.1.1)–(2.1.2), (2.1.8) *and* (2.1.11). *We impose the following restrictions on the measures* $\mu_1, \ldots, \mu_n$*:*

(i) $\int |x| \mu_i(\mathrm{d}x) < \infty$ *with* $\int x \mu_i(\mathrm{d}x) = 0$ *and* $c_{i-1} \leq c_i$ *for all* $1 \leq i \leq n$*,*

(ii) for all $2 \leq i \leq n$ *and all* $0 < y < r_{\mu_i}$ *the mapping*

$$(l_{\mu_i}, y] \to \mathbb{R} \cup \{+\infty\}, \quad \zeta \mapsto \frac{c^i(\zeta, y)}{y-\zeta} \quad \textit{has a unique minimizer } \zeta^\star \tag{2.1.16}$$

and

$$c_i(y) > c_{\imath_i(y;y)}(y) \qquad \textit{whenever } \zeta^\star < y. \tag{2.1.17}$$

Remark 2.1.3 (Assumption ⊛). The condition that the call prices are non-decreasing in maturity

$$c_i \leq c_{i+1}, \qquad i = 1, \dots, n-1, \tag{2.1.18}$$

can be rephrased by saying that $\mu_1, \dots, \mu_n$ are non-decreasing in the convex order. Condition (i) in Assumption ⊛ is the necessary and sufficient condition for a uniformly integrable martingale with these marginals to exist, as shown by e.g. Strassen [92, Theorem 2] or Meyer [65, Chapter XI].

Condition (ii) in Assumption ⊛ will be discussed further in Section 2.3.

Note that if (2.1.18) holds with strict inequality then (2.1.17) is automatically satisfied.

Remark 2.1.4 (Discontinuity of ξ_1). Note that Assumption ⊛(ii) does not require that the mapping

$$\zeta \mapsto \frac{c^1(\zeta, y)}{y - \zeta} = \frac{c_1(\zeta)}{y - \zeta} \tag{2.1.19}$$

has a unique minimizer. It may happen that there is an interval of minimizers and then ξ_1 is discontinuous at such y.

2.1.3 The Main Result

Our main result shows how to iteratively define an embedding of $(\mu_1, \dots, \mu_n)$ in the spirit of Azéma and Yor [4] and Brown et al. [18] if Assumption ⊛ is in place.

Theorem 2.1.5 (Main Result). *Let* $n \in \mathbb{N}$ *and assume that the measures* $\mu_1, \dots, \mu_n$ *satisfy Assumption ⊛ from Section 2.1.2. Recall Definitions 2.1.1 and 2.1.2.*

Then $\tau_i < \infty$, $B_{\tau_i} \sim \mu_i$ for all $i = 1, \dots, n$ and $(B_{\tau_n \wedge t})_{t \geq 0}$ is a uniformly integrable martingale. In addition, we have for $y \geq 0$ and $i = 1, \dots, n$,

$$\mathbb{P}\left[\bar{B}_{\tau_i} \geq y\right] = K_i(y) \tag{2.1.20}$$

where K_i is defined in (2.1.6)*.*

Remark 2.1.6 (Inductive Nature)*.* It is important to observe that ξ_i and therefore also τ_i, only depend on $\mu_1, \dots, \mu_i$. This gives an iterative structure allowing to "add one marginal at a time" and enables us to naturally prove the Theorem by induction on n.

Remark 2.1.7 (Minimality)*.* Since all τ_i are such that $(B_{t \wedge \tau_i})_{t \geq 0}$ is a uniformly integrable martingale it follows from Monroe [66] that all τ_i are *minimal* (in the sense of Monroe [66]).

2.1.4 Examples

Examples 2.1.8 and 2.1.9, respectively, show that we recover the stopping boundaries obtained by Madan and Yor [64] and Brown et al. [18], respectively. In particular the case $n = 1$ corresponds to the solution of Azéma and Yor [4].

Example 2.1.8 (Madan and Yor [64])**.** *Recall the definition of the barycentre function b_i from* (2.1.9)*. Madan and Yor [64] consider the "increasing mean residial value" case, i.e.*

$$b_1 \leq b_2 \leq \cdots \leq b_n. \tag{2.1.21}$$

We will now show that our main result reproduces their result if Assumption ⊛ is in place. In fact, as can be seen below, our definitions of ξ_i and K_i, cf. (2.1.4) *and* (2.1.6)*, respectively, reproduce the correct stopping boundaries in the general case,* *showing that Assumption ⊛ is not necessary, cf. also Section 2.3. More precisely, we have*

$$\xi_i = b_i^{-1}, \quad K_i(y) = \frac{c_i(b_i^{-1}(y))}{y - b_i^{-1}(y)} =: \mu_i^{\mathrm{HL}}([y,\infty)), \qquad i = 1,\dots,n, \tag{2.1.22}$$

where b_i^{-1} denotes the right-continuous inverse of b_i and μ_i^{HL} is the Hardy-Littlewood transform of μ_i, cf. Carraro et al. [20].

Clearly, the claim is true for $i = 1$. Let us assume that the claim holds for all $i \leq n-1$. Now, the optimization problem for ξ_n in (2.1.4) *becomes*

$$\begin{aligned}\xi_n(y) &\in \arg\min_{\zeta\leq y}\left\{\frac{c_n(\zeta)}{y-\zeta} - \mathbb{1}_{\{\zeta > b_{n-1}^{-1}(y)\}}\left[\frac{c_{n-1}(\zeta)}{y-\zeta} - \frac{c_{n-1}(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)}\right]\right\}\\ &\in \arg\min_{\zeta\leq y}\left\{\min_{\zeta\leq b_{n-1}^{-1}(y)}\frac{c_n(\zeta)}{y-\zeta}, \min_{\zeta\geq b_{n-1}^{-1}(y)}\left(\frac{c_n(\zeta)}{y-\zeta} - \left[\frac{c_{n-1}(\zeta)}{y-\zeta} - \frac{c_{n-1}(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)}\right]\right)\right\}.\end{aligned}$$

It is clear that the first minimum is $A_1 = \frac{c_n(b_n^{-1}(y))}{y - b_n^{-1}(y)}$ since $b_n^{-1}(y) \leq b_{n-1}^{-1}(y)$.

As for the second minimum, we set

$$F(\zeta) := \frac{c_n(\zeta)}{y-\zeta} - \left[\frac{c_{n-1}(\zeta)}{y-\zeta} - \frac{c_{n-1}\left(b_{n-1}^{-1}(y)\right)}{y - b_{n-1}^{-1}(y)}\right]$$

and we see by direct calculation that for almost all $\zeta \in \mathbb{R}$

$$\begin{aligned}(y-\zeta)^2 F'(\zeta) &= (b_n(\zeta) - y)\mu_n([\zeta,\infty)) - (b_{n-1}(\zeta) - y)\mu_{n-1}([\zeta,\infty))\\ &= c_n(\zeta)\frac{b_n(\zeta) - y}{b_n(\zeta) - \zeta} - c_{n-1}(\zeta)\frac{b_{n-1}(\zeta) - y}{b_{n-1}(\zeta) - \zeta}.\end{aligned}$$

By (2.1.21)*, we conclude therefore*

$$(y-\zeta)^2 F'(\zeta) \geq (c_n(\zeta) - c_{n-1}(\zeta))\,\frac{b_{n-1}(\zeta) - y}{b_{n-1}(\zeta) - \zeta} \geq 0,$$

where the last inequality follows from the non-decrease of the μ_i's in the convex order. Hence F is non-decreasing, and it follows that it attains its minimum at *the left boundary, i.e.* $A_2 = \frac{c_n(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)} - \left[\frac{c_{n-1}(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)} - \frac{c_{n-1}(b_{n-1}^{-1}(y))}{y - b_{n-1}^{-1}(y)}\right] = \frac{c_n(b_{n-1}^{-1})(y)}{y - b_{n-1}^{-1}(y)}$. *Consequently, by* (2.1.10), $\min\{A_1, A_2\} = A_1$ *and* (2.1.22) *follows.*

Example 2.1.9 (Brown et al. [18])**.** *In the case of* $n = 2$ *our definition of* ξ_1 *and* ξ_2 *clearly recovers the stopping boundaries in the main result of Brown et al. [18]. However, our embedding is not as general as their embedding because we enforce Assumption* ⊛, *see also the discussion in Section 2.3.*

Example 2.1.10 (Locally no Constraints)**.** *In general we have*

$$K_n(y) \le \mu_n^{\mathrm{HL}}([y, \infty)), \tag{2.1.23}$$

which holds by the fact that the distribution of the maximum in the n*-marginal problem cannot be larger (in stochastic order) than in the* 1*-marginal problem where it is bounded by* μ_n^{HL}. *However, if*

$$\xi_n(y) = b_n^{-1}(y) \qquad \textit{and} \qquad \imath_n(\xi_n(y); y) = 0 \tag{2.1.24}$$

for some $y \ge 0$ *then it follows from Theorem 2.1.5 that*

$$K_n(y) = \frac{c_n(b_n^{-1}(y))}{y - b_n^{-1}(y)} = \mu_n^{\mathrm{HL}}([y, \infty)), \tag{2.1.25}$$

i.e. locally at level of maximum y *the intermediate laws have no impact on the distribution of the terminal maximum as compared with the (one marginal) Azéma-Yor embedding.*

2.1.5 Properties of ξ_n and K_n

Under Assumption ⊛ we establish the continuity of ξ_n for $n \ge 2$, cf. Lemma 2.1.11, and prove monotonicity of ξ_n for $n \ge 1$, cf. Lemma 2.1.12. In Lemma 2.1.14 we derive an ODE for K_n which will be later used to identify the distribution of the maximum of the embedding from Definition 2.1.2.

Let $n_1 < n_2$. Recalling Remark 2.1.6 it follows that the embedding of the first n_1 marginals in the n_2-marginals embedding problem coincides with the n_1-marginals embedding problem. Hence it is natural to prove the Lemma by induction over the number of marginals n.

Lemma 2.1.11 (Continuity of ξ_n)**.** *Let $n \geq 2$ and let Assumption $\circledast$ hold. Set*

$$\Delta := \{(x, y) \in \mathbb{R} \times \mathbb{R}_+ : x \leq y\} . \tag{2.1.26}$$

Then the mappings

$$c^n : \Delta \to \mathbb{R}, \qquad (x, y) \mapsto c^n(x, y), \tag{2.1.27a}$$

$$\xi_n : \mathbb{R}_+ \to \mathbb{R}, \qquad y \quad \mapsto \xi_n(y) \tag{2.1.27b}$$

are continuous.

Proof. We prove the claim by induction over n. Let us start with the induction basis $n = 1, 2$. Continuity of c^1 is the same as continuity of c_1 and continuity of c^2 is proven by Brown et al. [18], cf. Lemma 3.5 therein. As for continuity of ξ_2 we note that our Assumption $\circledast$(ii) precisely rules out discontinuities of ξ_2 as shown by Brown et al. [18, Section 3.5]. By induction hypothesis we assume continuity of $c^1, \ldots, c^{n-1}$ and $\xi_2, \ldots, \xi_{n-1}$.

Observe that for $(\tilde{x}, \tilde{y})$ where ξ_1 is continuous at $\tilde{y}$ and $\imath_n$ is constant on $(\tilde{x} - \epsilon, \tilde{x} + \epsilon) \times (\tilde{y} - \epsilon, \tilde{y} + \epsilon)$, $\epsilon > 0$, continuity of c^n follows by the observation that $y \mapsto K_{\imath_n}(y) = \frac{c^{\imath_n}(\xi_{\imath_n}(y), y)}{y - \xi_{\imath_n}(y)}$, where $\imath_n = \imath_n(x, y) < n$, is continuous at $\tilde{y}$ by induction hypothesis.

Hence, we need to investigate the case when the index $\imath_n$ changes or when discontinuities of ξ_1 occur. This only happens at $(x, y) = (\xi_k(y), y)$ for some $k < n$, or, in the case that y is a discontinuity of ξ_1, at (x, y) where $x \in [\xi_1(y-), \xi_1(y+)]$. We prove continuity at (x, y).

Consider first the following cases:

$$\text{if } x = \xi_k(y) \qquad \text{then} \quad x \neq \xi_j(y) \quad \text{for all } j \neq k,\ j < n, \tag{2.1.28a}$$

$$\text{or,} \quad \text{if } x \in [\xi_1(y-), \xi_1(y+)] \quad \text{then} \quad x \neq \xi_j(y) \quad \text{for all } j \neq 1,\ j < n. \tag{2.1.28b}$$

Note that in case (2.1.28b) we have from Remark 2.1.4

$$K_1(y) = \frac{c_1(x)}{y - x} \qquad \text{for all } x \in [\xi_1(y-), \xi_1(y+)]. \tag{2.1.29}$$

We will call a point (x, y) to be "to the right of ξ_k" if $\xi_k(y) < x$ and "to the left of ξ_k" if $\xi_k(y) \geq x$. From (2.1.28a)–(2.1.28b) it follows that there exists an $\epsilon > 0$ such that each point $(\tilde{x}, \tilde{y})$ in the ϵ-neighbourhood of (x, y) is either to the left or to the right of ξ_k and there are no other boundaries in this ϵ-neighbourhood, in particular

$$k = \imath_n(x_r; y_r), \qquad j = \imath_n(x_l; y_l) = \imath_{\imath_n(x_r;y_r)}(x_r; y_r), \tag{2.1.30}$$

where (x_r, y_r) is in the ϵ-neighbourhood of (x, y) and to the right of ξ_k and (x_l, y_l) is in the ϵ-neighbourhood of (x, y) and to the left of ξ_k.

If $x < y$, we have by induction hypothesis

$$\begin{aligned}
c^n(x_r, y_r) = \quad & c_n(x_r) - \{c_k(x_r) - (y_r - x_r)K_k(y_r)\} && (2.1.31)\\
\xrightarrow[\text{from the right}]{(x_r,y_r)\to(x,y)} \quad & c_n(x) - \{c_k(x) - (y-x)K_k(y)\} \\
\overset{(2.1.6)}{\underset{(2.1.28a)-(2.1.30)}{=}} \quad & c_n(x) - \left\{c_k(x) - \frac{y-x}{y-x}\Big(c_k(x) - [c_j(x) - (y-x)K_j(y)]\Big)\right\} \\
= \quad & c_n(x) - [c_j(x) - (y-x)K_j(y)] \\
\overset{(2.1.11)}{=} \quad & c^n(x,y) && (2.1.32)\\
= \quad & c_n(x) - [c_j(x) - (y-x)K_j(y)] \\
\xleftarrow[\text{from the left}]{(x_l,y_l)\to(x,y)} \quad & c_n(x_l) - \{c_j(x_l) - (y_l - x_l)K_j(y_l)\} = c^n(x_l, y_l). && (2.1.33)
\end{aligned}$$

From (2.1.31), (2.1.32) and (2.1.33) continuity of c^n follows for any sequence $(x_n, y_n) \to (x, y)$. We now extend the above argument to the situation when $x = y$ which establishes continuity of c^n at (y, y). In this case we have $x = \xi_k(y) = y$. For this to hold we must have $c_k(y) = c_j(y)$. By boundedness of $K_i,\ i < n$, it follows that (2.1.32) and (2.1.33) converge to each other.

To relax (2.1.28a)–(2.1.28b) we successively write out $K_k, K_j, \ldots,$ until the assumption (2.1.28a)–(2.1.28b) holds true and then apply the claim in the special case.

It remains to prove continuity of ξ_n which we prove by contradiction. Assume there exist $\epsilon > 0$ and $y > 0$ such that for all $\delta > 0$ there exists a $y' \in (y, y+\delta)$ such that $|\xi_n(y) - \xi_n(y')| > \epsilon$. By (2.1.15) the limit of $\xi_n(y')$ as $y' \downarrow y$ exists at least along some subsequence and we denote it by $\tilde{\xi}_n$. By assumption $\tilde{\xi}_n \neq \xi_n(y)$.

Consider first the case that $\xi_n(y) < y$ and $\tilde{\xi}_n < y$. Using continuity of c^n we deduce $\frac{c^n(\xi_n(y'),y')}{y'-\xi_n(y')} \to \frac{c^n(\tilde{\xi}_n,y)}{y-\tilde{\xi}_n}$ as $y' \to y$.

Now, if

$$\frac{c^n(\tilde{\xi}_n, y)}{y - \tilde{\xi}_n} \neq \frac{c^n(\xi_n(y), y)}{y - \xi_n(y)} \tag{2.1.34}$$

then we obtain a contradiction to the optimality of either $\xi_n(y)$ or some $\xi_n(y')$ for y' close enough to y by continuity of c^n. If

$$\frac{c^n(\tilde{\xi}_n, y)}{y - \tilde{\xi}_n} = \frac{c^n(\xi_n(y), y)}{y - \xi_n(y)} \tag{2.1.35}$$

we obtain a contradiction to Assumption $\circledast$(ii).

We now consider the case that either $\xi_n(y) = y$ or $\tilde{\xi}_n = y$. The case $\xi_n(y) < y$ and $\tilde{\xi}_n = y$ is ruled out by condition (2.1.17) from Assumption $\circledast$(ii): Indeed, for the sequence $\left(K_n(y') = \frac{c^n(\xi_n(y'),y')}{y'-\xi_n(y')}\right)$ to be bounded we must have $c^n(\xi_n(y'), y') \to 0$. Recalling the continuity of c^n at (y, y) implies $c_n(y) = c_{\imath_n(y;y)}(y)$.

The case $\xi_n(y) = y$ and $\tilde{\xi}_n < y$ follows as above by distinguishing the cases (2.1.34) and (2.1.35) and by recalling (2.1.13) and the continuity of c^n at (y, y). □

Lemma 2.1.12 (Monotonicity of ξ_n). *Let $n \in \mathbb{N}$ and let Assumption $\circledast$ hold.*

Then

$$\xi_n : \mathbb{R}_+ \to \mathbb{R}, \quad y \mapsto \xi_n(y) \qquad \textit{is non-decreasing.} \tag{2.1.36}$$

Proof. The claim for $n = 1, 2$ follows from Brown et al. [18]. Assume by induction hypothesis that we have proven monotonicity of $\xi_1, \ldots, \xi_{n-1}$.

We follow closely the arguments of Brown et al. [18, Lemma 3.2]. Since ξ_n is continuous it is enough to prove monotonicity at almost every $y > 0$. The set of y's which are a discontinuity of ξ_1 is a null-set, and hence we can exclude all such y's. In the following we fix a y where $\xi_1, \ldots, \xi_n$ are continuous.

We will first consider the case when $\xi_n(y) \neq \xi_j(y)$ for all $j < n$. By continuity of ξ_n it follows that there is an $\epsilon > 0$ such that

$$\xi_n(\tilde{y}) \neq \xi_j(\tilde{y}) \text{ and } \ell := \jmath_n(y) = \jmath_n(\tilde{y}) \quad \forall \tilde{y} \in (y-\epsilon, y+\epsilon) \text{ and } j < n, \tag{2.1.37}$$

and furthermore

$$(\xi_n(\tilde{y}), \tilde{y}) \in (\xi_n(y) - \epsilon, \xi_n(y) + \epsilon) \times (y - \epsilon, y + \epsilon). \tag{2.1.38}$$

Let l_1 denote a supporting tangent to $c^n(\cdot, y)$ at $\xi_n(y)$ which goes through the x-axis at y (it exists by definition of $\xi_n(y)$), i.e.

$$l_1(x) = c^n(\xi_n(y), y) + (x - \xi_n(y))(D - K_\ell(y)),$$

where D lies between the left- and right-derivatives of $c_n - c_\ell$ at $\xi_n(y)$. Using that $l_1(y) = 0$ we can write

$$D - K_\ell(y) = -\frac{c^n(\xi_n(y), y)}{y - \xi_n(y)} \overset{(2.1.11)}{=} -\frac{c_n(\xi_n(y)) - c_\ell(\xi_n(y))}{y - \xi_n(y)} - K_\ell(y)$$

and thus by (2.1.18)

$$D \leq 0. \tag{2.1.39}$$

We also have

$$l_1(y + \delta) = \delta(D - K_\ell(y)). \tag{2.1.40}$$

Choose $\delta \in (0, \epsilon)$ sufficiently small. Our goal is to prove $\xi_n(y+\delta) \geq \xi_n(y)$. Recall that $\xi_n(y+\delta)$ is determined from $y+\delta$ and $c^n(\cdot, y+\delta)$ only. Since we know that $\xi_n(y+\delta) \in (\xi_n(y) - \epsilon, \xi_n(y) + \epsilon) := I$ it will turn out to be enough to look at

$c^n(x, y+\delta)$ only for $x \in (\xi_n(y) - \epsilon, \xi_n(y) + \epsilon)$. For such an x we have

$$c^n(x, y+\delta) - c^n(x, y) \stackrel{(2.1.11)}{=} (y + \delta - x)\, K_\ell(y+\delta) - (y - x)\, K_\ell(y). \tag{2.1.41}$$

Let l_2 be the supporting tangent to $c^n(\cdot, y+\delta) - c^n(\cdot, y)$ at $\xi_n(y)$, i.e.

$$l_2(x) = c^n(\xi_n(y), y+\delta) - c^n(\xi_n(y), y) + (x - \xi_n(y))(K_\ell(y) - K_\ell(y+\delta)).$$

Hence,

$$\begin{aligned} l_1(y+\delta) + l_2(y+\delta) &\stackrel{(2.1.40)}{=} \delta(D - K_\ell(y)) \\ &\quad + c^n(\xi_n(y), y+\delta) - c^n(\xi_n(y), y) \\ &\quad + (y + \delta - \xi_n(y))(K_\ell(y) - K_\ell(y+\delta)) \\ &\stackrel{(2.1.41)}{=} \delta D \leq 0. \end{aligned} \tag{2.1.42}$$

Now, since $c^n(\cdot, y+\delta) - c^n(\cdot, y)$ is linear (and therefore convex) in the domain I, $l_1 + l_2$ is a supporting tangent to $c^n(\cdot, y+\delta)$ at $\xi_n(y)$, i.e.

$$(l_1 + l_2)(x) \leq c^n(x, y+\delta) \qquad \text{for } x \in I, \tag{2.1.43}$$

$$(l_1 + l_2)(\xi_n(y)) = c^n(\xi_n(y), y+\delta). \tag{2.1.44}$$

Recall that $\xi_n(y+\delta)$ is determined as the x-value where the supporting tangent to $c^n(\cdot, y+\delta)$ which passes the x-axis at $y+\delta$ touches $c^n(\cdot, y+\delta)$. Next we exploit the fact that $\xi_n(y+\delta) \in I$ which implies that we only need to show that $\xi_n(y+\delta) \notin (\xi_n(y) - \epsilon, \xi_n(y))$. Then this follows from (2.1.42) which yields that any supporting tangent to $c^n(\cdot, y+\delta)$ at some $\zeta \in (\xi_n(y) - \epsilon, \xi_n(y))$ must be below the x-axis when evaluated at $y + \delta$. We refer to Brown et al. [18, Fig.7] for a graphical illustration of this fact.

Now we relax the assumption (2.1.37). Assume that there exists a $\delta > 0$ such that $\xi_n(y) > \xi_n(y+\delta)$. We derive a contradiction to the special case as follows. Set $y_0 := y$ and $y_n := y + \delta$. Recall that ξ_n is continuous. Now we can choose $y_0 < y_1 < \cdots < y_{n-1} < y_n$ such that $\xi_n(y_0) > \xi_n(y_1) > \cdots > \xi_n(y_{n-1}) > \xi_n(y_n)$. Set $x_i := \xi_n(y_i)$, $i = 0, \dots, n$. Observe that by monotonicity of ξ_k, $k < n$ the graph of ξ_k intersects with at most one rectangle $(x_i, x_{i-1}) \times (y_{i-1}, y_i)$, $i = 1, \dots, n$. Consequently, there must exist at least one integer j such that the rectangle $R := (x_j, x_{j-1}) \times (y_{j-1}, y_j)$ is disjoint with the graph of every ξ_k, $k < n$. By construction and continuity of $y \mapsto \xi_n(y)$ R is not disjoint with the graph of ξ_n. Inside this rectangle R the conditions of the special case (2.1.37) are satisfied. Recalling that $\xi_n(y_j) = x_j < x_{j-1} = \xi_n(y_{j-1})$ and by continuity of $y \mapsto \xi_n(y)$, we can find two points $s_1 < s_2$ such that $z_1 = \xi_n(s_1) > \xi_n(s_2) = z_2$ and $(z_1, s_1) \in R, (z_2, s_2) \in R$. This is a contradiction. □

Remark 2.1.13 (Properties of $\imath_n$). Recalling the definition of $\imath_n$, cf. (2.1.3), we observe for later use that by the above monotonicity result we have for all $y \geq 0$:

$$\imath_n(\cdot; y) \text{ is left-continuous and has at most } n-1 \text{ jumps.} \tag{2.1.45}$$

Further, for all $x \in \mathbb{R}$:

$$\imath_n(x; \cdot) \text{ is right-continuous and has at most } n-1 \text{ jumps.} \tag{2.1.46}$$

Lemma 2.1.14 (ODE for K_n)**.** *Let $n \in \mathbb{N}$ and let Assumption ⊛ hold.*

Then

$$y \mapsto K_n(y) \qquad \textit{is absolutely continuous and non-increasing.} \tag{2.1.47}$$

If we assume in addition that the embedding property of Theorem 2.1.5 is valid for the first $n-1$ marginals then for almost all $y \geq 0$ we have:

If $\xi_n(y) < y$ then

$$K_n'(y) + \frac{K_n(y)}{y - \xi_n(y)} = K_{\jmath_n(y)}'(y) + \frac{K_{\jmath_n(y)}(y)}{y - \xi_n(y)} \tag{2.1.48}$$

where K_j' denotes the derivative of K_j which exists for almost all $y \geq 0$ and $j = 1, \dots, n$.

If $\xi_n(y) = y$ then

$$K_n(y+) = K_{\jmath_n(y)}(y+). \tag{2.1.49}$$

Finally, if $\xi_n'(y) > 0$ then for almost all $y \geq 0$,

$$K_n(y) + c_n'(\xi_n(y)+) - c_j'(\xi_n(y)+) - K_j(y) = 0 \tag{2.1.50}$$

for $j = \jmath_n(y)$ and j such that $n > j > \jmath_n(y)$ and $\xi_n(y) = \xi_j(y)$.

Proof. The proof is reported in Appendix A.1. □

2.2 Proof of the Main Result

In this section we prove the main result, Theorem 2.1.5. The key step is the identification of the distribution of the maximum, cf. Proposition 2.2.4.

Let $n \in \mathbb{N}$. For convenience we set

$$M_0 := 0, \qquad M_i := B_{\tau_i}, \qquad i = 1, \dots, n, \tag{2.2.1}$$

where τ_i is defined in Definition 2.1.2.

2.2.1 Basic Properties of the Embedding

Our first result shows that there is a "strong relation" between M and $\bar{M}$.

Lemma 2.2.1 (Relations Between M and $\bar{M}$)**.** *Let $n \in \mathbb{N}$ and let Assumption $\circledast$ hold.*

Then the following implications hold.

$$M_n > \xi_n(y) \quad \Longrightarrow \quad \bar{M}_n \geq y, \tag{2.2.2}$$

$$M_n \geq \xi_n(y) \quad \Longrightarrow \quad \bar{M}_n \geq y \quad \textit{if } \xi_n \textit{ is strictly increasing at } y. \tag{2.2.3}$$

For $y \geq 0$ such that $\jmath_n(y) \neq 0$ we have

$$M_{\jmath_n(y)} \geq \xi_n(y) > \xi_{\jmath_n(y)}(y) \quad \Longrightarrow \quad M_n \geq \xi_n(y), \tag{2.2.4}$$

$$\bar{M}_{\jmath_n(y)} < y, \;\; \bar{M}_n \geq y \quad \Longrightarrow \quad M_n \geq \xi_n(y), \tag{2.2.5}$$

$$\bar{M}_{\jmath_n(y)} \geq y, \;\; M_{\jmath_n(y)} < \xi_n(y) \quad \Longrightarrow \quad M_n < \xi_n(y). \tag{2.2.6}$$

If ξ_n is strictly increasing at $y \geq 0$ and $\jmath_n(y) = 0$ then the following holds.

$$M_n \geq \xi_n(y) \quad \Longleftrightarrow \quad \bar{M}_n \geq y. \tag{2.2.7}$$

Proof. Write $\jmath = \jmath_n$. We have

$$\xi_{\jmath(y)}(y) < \xi_n(y) \leq \xi_i(y), \quad i = \jmath(y) + 1, \ldots, n.$$

In the following we are using continuity and monotonicity of $\xi_1, \ldots, \xi_n$, cf. Lemma 2.1.11 and 2.1.12.

Case $\jmath(y) \neq 0$. As for implication (2.2.2) assume that $M_n > \xi_n(y)$ and $\bar{M}_n < y$ holds. In this case M_n cannot be at the boundary ξ_n. There has to be a $j < n$ such that $M_n = M_j, \bar{M}_n = \bar{M}_j$ and $M_j = \xi_j(\bar{M}_j) = \xi_j(y')$ for some $y' < y$. However, this cannot be true because $\xi_n(y') \leq \xi_n(y) < \xi_j(y') = M_n$ and hence case (2.1.7a) of the definition of $\tau_1, \dots, \tau_n$ would have been triggered.

Implication (2.2.3) follows by the same arguments as for implication (2.2.2).

Implication (2.2.4) now follows from implication (2.2.2) applied for $\jmath(y)$ and the fact that either $M_n = M_{\jmath(y)}$ (case (2.1.7b)) or M moves to a point at the boundary $\xi_i(y') \geq \xi_n(y)$ for some $i = \jmath(y) + 1, \dots, n$, $y' \geq y$ (case (2.1.7a)).

Implication (2.2.5) holds because if M increases its maximum at time $\jmath(y)$, which is $< y$, to some $y' \geq y$ at time n, it will hit a boundary point $\xi_i(y') \geq \xi_n(y)$ for some $i = \jmath(y) + 1, \dots, n$.

Implication (2.2.6) holds because from $\bar{M}_{\jmath(y)} \geq y$ and $M_{\jmath(y)} < \xi_n(y)$ it follows that $M_{\jmath(y)} = \xi_i(y') < \xi_n(y) \leq \xi_j(y')$ for some $i \leq \jmath(y)$, $y' \geq y$, $j > \jmath(y)$. From this it follows that M will stay where it is until time n, cf. case (2.1.7b).

Case $\jmath(y) = 0$. Assume that ξ_n is strictly at y and that $\bar{M}_n \geq y$ holds. In this case M_n must be at a boundary point $\xi_i(y') \geq \xi_n(y)$ for some $i = 1, \dots, n$, $y' \geq y$. The converse direction is just (2.2.3), together giving (2.2.7). □

As an application of Lemma 2.2.1 we obtain the following result.

Lemma 2.2.2 (Contributions to the Maximum)**.** *Let $n \in \mathbb{N}$ and let Assumption $\circledast$ hold. Assume ξ_n is strictly increasing at $y \geq 0$.*

Then, if $\jmath_n(y) \neq 0$,

$$\mathbb{P}\left[\bar{M}_n \geq y\right] = \mathbb{P}\left[M_n \geq \xi_n(y)\right] - \mathbb{P}\left[M_{\jmath_n(y)} \geq \xi_n(y)\right] + \mathbb{P}\left[\bar{M}_{\jmath_n(y)} \geq y\right] \tag{2.2.8}$$

and if $\jmath_n(y) = 0$,

$$\mathbb{P}\left[\bar{M}_n \geq y\right] = \mathbb{P}\left[M_n \geq \xi_n(y)\right]. \tag{2.2.9}$$

Proof. Write $\jmath = \jmath_n$.

Case $\jmath(y) \neq 0$. Firstly, let us compute

$$\begin{aligned}
&\mathbb{P}\left[\bar{M}_n \geq y\right] - \mathbb{P}\left[M_n \geq \xi_n(y)\right] \\
\overset{(2.2.3)}{=}\; &\mathbb{P}\left[\bar{M}_n \geq y\right] - \mathbb{P}\left[M_n \geq \xi_n(y), \bar{M}_n \geq y\right] = \mathbb{P}\left[\bar{M}_n \geq y, M_n < \xi_n(y)\right] \\
=\; &\mathbb{P}\left[\bar{M}_n \geq y, M_n < \xi_n(y), \bar{M}_{\jmath(y)} \geq y\right] + \mathbb{P}\left[\bar{M}_n \geq y, M_n < \xi_n(y), \bar{M}_{\jmath(y)} < y\right] \\
\underset{(2.2.5)}{\overset{(2.2.6)}{=}}\; &\mathbb{P}\left[M_n < \xi_n(y), \bar{M}_{\jmath(y)} \geq y, M_{\jmath(y)} < \xi_n(y)\right].
\end{aligned}$$

Secondly, let us compute

$$\begin{aligned}
&\mathbb{P}\left[\bar{M}_{\jmath(y)} \geq y\right] - \mathbb{P}\left[M_{\jmath(y)} \geq \xi_n(y)\right] \\
=\; &\mathbb{P}\left[\bar{M}_{\jmath(y)} \geq y, M_{\jmath(y)} \geq \xi_n(y)\right] + \mathbb{P}\left[\bar{M}_{\jmath(y)} \geq y, M_{\jmath(y)} < \xi_n(y)\right] - \mathbb{P}\left[M_{\jmath(y)} \geq \xi_n(y)\right] \\
\overset{(2.2.2)}{=}\; &\mathbb{P}\left[\bar{M}_{\jmath(y)} \geq y, M_{\jmath(y)} < \xi_n(y)\right] \overset{(2.2.6)}{=} \mathbb{P}\left[M_n < \xi_n(y), \bar{M}_{\jmath(y)} \geq y, M_{\jmath(y)} < \xi_n(y)\right].
\end{aligned}$$

Comparing these two equations yields the claim.

Case $\jmath(y) = 0$. The claim follows directly from (2.2.7). □

2.2.2 Law of the Maximum

Our next goal is to identify the distribution of M_n. We will achieve this by deriving an ODE for $\mathbb{P}\left[\bar{M}_n \geq \cdot\right]$ using excursion theoretical results, cf. Lemma 2.2.3, and link it to the ODE satisfied by K_n, cf. Lemma 2.1.14.

Lemma 2.2.3 (ODE for the Maximum)**.** *Let $n \in \mathbb{N}$ and let Assumption ⊛ hold.*

Then the mapping

$$y \mapsto \mathbb{P}\left[\bar{M}_n \geq y\right]$$

is absolutely continuous and for almost all $y \geq 0$ we have:

If $\xi_n(y) < y$ *then*

$$\frac{\partial \mathbb{P}\left[\bar{M}_n \geq y\right]}{\partial y} + \frac{\mathbb{P}\left[\bar{M}_n \geq y\right]}{y - \xi_n(y)} = \frac{\mathbb{P}\left[\bar{M}_{\jmath_n(y)} \geq y\right]}{y - \xi_n(y)} + \left.\frac{\partial \mathbb{P}\left[\bar{M}_j \geq y\right]}{\partial y}\right|_{j=\jmath_n(y)}. \quad (2.2.10)$$

If $\xi_n(y) = y$ *then*

$$\mathbb{P}\left[\bar{M}_n > y\right] = \mathbb{P}\left[\bar{M}_{\jmath_n(y)} > y\right]. \quad (2.2.11)$$

Proof. Write $\jmath = \jmath_n$. We exclude all $y > 0$ which are a discontinuity of ξ_1. This is clearly a null-set.

The cases $n = 1, 2$ are true by Brown et al. [18]. Assume by induction hypothesis that we have proven the claim for $i = 1, \dots, n-1$.

If $\xi_n(y) = y$ then it is clear from the definition of the embedding, cf. Definition 2.1.2, that

$$\bar{M}_n > y \qquad \Longleftrightarrow \qquad \bar{M}_{\jmath(y)} > y. \quad (2.2.12)$$

Case $\jmath(y) \neq 0$. For $\delta > 0$ we have

$$\begin{aligned} & \mathbb{P}\left[\bar{M}_n \geq y+\delta, \bar{M}_{\jmath(y)} < y+\delta\right] - \mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y\right] \\ = \; & \mathbb{P}\left[\bar{M}_n \geq y+\delta, \bar{M}_{\jmath(y)} < y\right] - \mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y\right] \\ + \; & \underbrace{\mathbb{P}\left[\bar{M}_n \geq y+\delta, y \leq \bar{M}_{\jmath(y)} < y+\delta\right]}_{=0 \text{ for } \delta>0 \text{ small enough by definition of } \jmath(y) \text{ and continuity of } \xi_i}. \end{aligned} \quad (2.2.13)$$

For $r > 0$ we define

$$\bar{\xi}_j(r) := \max_{k:j\leq k\leq n} \{\xi_k(r) : \xi_k(y) = \xi_n(y)\},$$

$$\underline{\xi}_j(r) := \min_{k:j\leq k\leq n} \{\xi_k(r) : \xi_k(y) = \xi_n(y)\}$$

and note that

$$\bar{\xi}_{\jmath(y)}(r) \to \xi_n(y), \qquad \underline{\xi}_{\jmath(y)}(r) \to \xi_n(y) \qquad \text{as } r \to y \tag{2.2.14}$$

by continuity of ξ_i at y for $i = 1, \ldots, n$.

Let $\delta > 0$. We have by excursion theoretical results, cf. e.g. Rogers [83],

$$\begin{aligned} &\mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y\right] \exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \bar{\xi}_{\jmath(y)}(r)}\right) \\ \leq\, &\mathbb{P}\left[\bar{M}_n \geq y + \delta, \bar{M}_{\jmath(y)} < y\right] \\ \leq\, &\mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y\right] \exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \underline{\xi}_{\jmath(y)}(r)}\right). \end{aligned} \tag{2.2.15}$$

Now we compute for y such that $\xi_n(y) < y$,

$$\begin{aligned} &\frac{\mathbb{P}\left[\bar{M}_n \geq y + \delta, \bar{M}_{\jmath(y)} < y + \delta\right] - \mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y\right]}{\delta} \\ \overset{(2.2.13),(2.2.15)}{\leq}\; &\mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y\right] \frac{\exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \underline{\xi}_{\jmath(y)}(r)}\right) - 1}{\delta} \\ \xrightarrow[\text{as } \delta \downarrow 0]{\text{by (2.2.14)}}\; &-\frac{\mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y\right]}{y - \xi_n(y)} \end{aligned} \tag{2.2.16}$$

and analogously

$$\begin{aligned} &\frac{\mathbb{P}\left[\bar{M}_n \geq y + \delta, \bar{M}_{\jmath(y)} < y + \delta\right] - \mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y\right]}{\delta} \\ \overset{(2.2.13),(2.2.15)}{\geq}\; &\mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y\right] \frac{\exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \bar{\xi}_{\jmath(y)}(r)}\right) - 1}{\delta} \\ \xrightarrow[\text{as } \delta \downarrow 0]{\text{by (2.2.14)}}\; &-\frac{\mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_{\jmath(y)} < y\right]}{y - \xi_n(y)}. \end{aligned} \tag{2.2.17}$$

Hence, from (2.2.16) and (2.2.17) it follows that the right-derivative of

$$y \mapsto \mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_j < y\right]\Big|_{j=\jmath(y)} \tag{2.2.18}$$

exists. Similar arguments for $\delta < 0$ show that the left-derivative exists and is the same as the right-derivative. Local Lipschitz continuity of (2.2.18) then follows from (2.2.16) and (2.2.17).

Observe the obvious equality

$$\mathbb{P}\left[\bar{M}_n \geq y\right] = \mathbb{P}\left[\bar{M}_j \geq y\right] + \mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_j < y\right]. \tag{2.2.19}$$

Taking $j = \jmath(y)$ in (2.2.19) and fixing it, we conclude by induction hypothesis that $y \mapsto \mathbb{P}\left[\bar{M}_n \geq y\right]$ is locally Lipschitz continuous and hence absolutely continuous and its derivative reads

$$\frac{\partial \mathbb{P}\left[\bar{M}_n \geq y\right]}{\partial y} = \left.\frac{\partial \mathbb{P}\left[\bar{M}_j \geq y\right]}{\partial y}\right|_{j=\jmath_n(y)} + \frac{\mathbb{P}\left[\bar{M}_{\jmath_n(y)} \geq y\right] - \mathbb{P}\left[\bar{M}_n \geq y\right]}{y - \xi_n(y)}.$$

Case $\jmath(y) = 0$. For $\delta > 0$ we have by excursion theoretical results

$$\begin{aligned}
& \mathbb{P}\left[\bar{M}_n \geq y+\delta, \bar{M}_1 < y+\delta\right] - \mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_1 < y\right] \\
= \; & \mathbb{P}\left[\bar{M}_n \geq y+\delta, \bar{M}_1 < y+\delta\right] - \mathbb{P}\left[\bar{M}_n \geq y+\delta, \bar{M}_1 < y\right] \\
& + \mathbb{P}\left[\bar{M}_n \geq y+\delta, \bar{M}_1 < y\right] - \mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_1 < y\right] \\
\leq \; & \int_y^{y+\delta} \mathbb{P}\left[\bar{M}_1 \in \mathrm{d}s\right] \frac{\left(\xi_1(s) - \underline{\xi}_1(s)\right)^+}{s - \underline{\xi}_1(s)} \exp\left(-\int_s^{y+\delta} \frac{\mathrm{d}r}{r - \underline{\xi}_1(r)}\right) \\
& + \mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_1 < y\right]\left[\exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \underline{\xi}_1(r)}\right) - 1\right].
\end{aligned} \tag{2.2.20}$$

Similarly, we have

$$\begin{aligned}
& \mathbb{P}\left[\bar{M}_n \geq y+\delta, \bar{M}_1 < y+\delta\right] - \mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_1 < y\right] \\
\geq \; & \int_y^{y+\delta} \mathbb{P}\left[\bar{M}_1 \in \mathrm{d}s\right] \frac{\left(\xi_1(s) - \bar{\xi}_1(s)\right)^+}{s - \bar{\xi}_1(s)} \exp\left(-\int_s^{y+\delta} \frac{\mathrm{d}r}{r - \bar{\xi}_1(r)}\right) \\
& + \mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_1 < y\right]\left[\exp\left(-\int_y^{y+\delta} \frac{\mathrm{d}r}{r - \bar{\xi}_1(r)}\right) - 1\right].
\end{aligned} \tag{2.2.21}$$

From (2.2.20) and (2.2.21) it follows that the right-derivative of

$$y \mapsto \mathbb{P}\left[\bar{M}_n \geq y, \bar{M}_1 < y\right] \tag{2.2.22}$$

exists. Similar arguments for $\delta < 0$ show that the left-derivative exists and is the same as the right-derivative. Local Lipschitz continuity of (2.2.22) then follows from (2.2.20) and (2.2.21). Now we can conclude from (2.2.19)–(2.2.21) applied with $j = 1$ that $y \mapsto \mathbb{P}\left[\bar{M}_n \geq y\right]$ is locally Lipschitz continuous and hence absolutely continuous and its derivative reads

$$\frac{\partial \mathbb{P}\left[\bar{M}_n \geq y\right]}{\partial y} \overset{(2.2.14)}{=} \frac{\partial \mathbb{P}\left[\bar{M}_1 \geq y\right]}{\partial y} - \frac{\partial \mathbb{P}\left[\bar{M}_1 \geq y\right]}{\partial y} \frac{(\xi_1(y) - \xi_n(y))^+}{y - \xi_n(y)} - \frac{\mathbb{P}\left[\bar{M}_n \geq y\right] - \mathbb{P}\left[\bar{M}_1 \geq y\right]}{y - \xi_n(y)},$$

which implies by induction hypothesis

$$\frac{\mathbb{P}\left[\bar{M}_n \geq y\right]}{y - \xi_n(y)} + \frac{\partial \mathbb{P}\left[\bar{M}_n \geq y\right]}{\partial y} = 0.$$

This finishes the proof. □

Finally, we argue that $\mathbb{P}\left[\bar{M}_n \geq y\right] = K_n(y)$ holds for all $y \geq 0$.

Proposition 2.2.4 (Law of the Maximum)**.** *Let $n \in \mathbb{N}$ and let Assumption ⊛ hold. Assume that the embedding property of Theorem 2.1.5 is valid for the first $n-1$ marginals.*

Then for all $y \geq 0$ we have

$$\mathbb{P}\left[\bar{M}_n \geq y\right] = K_n(y). \tag{2.2.23}$$

Proof. The case $n = 1$ holds by the Azéma-Yor embedding. Assume by induction hypothesis that

$$K_i = \mathbb{P}\left[\bar{M}_i \geq \cdot\right], \qquad i = 1, \dots, n-1.$$

In Lemma 2.1.14 and 2.2.3 we derived an ODE for K_n and $\mathbb{P}\left[\bar{M}_n \geq \cdot\right]$, respectively, in terms of $K_1, \dots, K_{n-1}$ and $\mathbb{P}\left[\bar{M}_1 \geq \cdot\right], \dots, \mathbb{P}\left[\bar{M}_{n-1} \geq \cdot\right]$, respectively. These ODEs are valid for a.e. $y \geq 0$. By induction hypothesis both drivers of these ODEs coincide everywhere and hence the claim follows from the boundary conditions

$$\begin{aligned} K_n(y) &\to 0 \quad \text{as } y \to \infty, & K_n(y) &\to 1 \quad \text{as } y \to 0,\\ \mathbb{P}\left[\bar{M}_n \geq y\right] &\to 0 \quad \text{as } y \to \infty, & \mathbb{P}\left[\bar{M}_n \geq y\right] &\to 1 \quad \text{as } y \to 0, \end{aligned}$$

absolute continuity of K_n and $\mathbb{P}\left[\bar{M}_n \geq \cdot\right]$ and the fact that the ODE

$$\Big(\mathbb{P}\left[\bar{M}_n \geq y\right] - K_n(y)\Big)' = -\frac{\mathbb{P}\left[\bar{M}_n \geq y\right] - K_n(y)}{y - \xi_n(y)}, \qquad \mathbb{P}\left[\bar{M}_n \geq 0\right] - K_n(0) = 0,$$

has unique solution given by 0. □

2.2.3 Embedding Property

In this section we prove that the stopping times $\tau_1, \dots, \tau_n$ from Definition 2.1.2 embed the laws $\mu_1, \dots, \mu_n$ if Assumption ⊛ is in place. More precisely, given Proposition 2.2.4 above and by inductive reasoning, to complete the proof of Theorem 2.1.5 we only need to show the following:

Proposition 2.2.5 (Embedding)**.** *In the setup of Theorem 2.1.5 we have*

$$B_{\tau_n} \sim \mu_n \tag{2.2.24}$$

and $(B_{\tau_n \wedge t})_{t \geq 0}$ *is a uniformly integrable martingale.*

Proof. The case $n = 1$ is just the Azéma-Yor embedding. By induction hypothesis, assume that the claim holds for all $i \leq n-1$.

We claim that ξ_n ranges continuously over the full support of μ_n. This is because, firstly, we know from Lemma 2.1.12 that $\xi_2, \dots, \xi_n$ are continuous. Secondly, we have by using $l_{\mu_n} \leq l_{\mu_i}$ that

$$\inf_{\zeta \leq 0} \frac{c^n(\zeta, 0)}{0-\zeta} \geq \inf_{\zeta \leq 0} \min_{1 \leq i < n} \left\{ \underbrace{\frac{c_n(\zeta) - c_i(\zeta)}{0-\zeta}}_{\geq 0} + \underbrace{K_i(0)}_{=1} \right\} \wedge \underbrace{\frac{c_n(l_{\mu_n})}{0 - l_{\mu_n}}}_{=1} = 1$$

which shows that $\xi_n(0) = l_{\mu_n}$ (note that by Assumption $\circledast$ we cannot have $c_n \equiv c_i$ on open intervals). Furthermore, by using $r_{\mu_n} \geq r_{\mu_i}$ we have from (2.1.9) and (2.1.15) that

$$\xi_n(r_{\mu_n}) = r_{\mu_n}.$$

Let $y > 0$ be such that $\xi_n'(y) > 0$ and $\xi_n(y)$ is not an atom of any $\mu_1, \dots, \mu_n$ and y is not a discontinuity of ξ_1. Applying previous results we obtain for almost all such y,

$$\begin{aligned} & \mathbb{P}\left[M_n \geq \xi_n(y)\right] - \mathbb{P}\left[M_{\jmath_n(y)} \geq \xi_n(y)\right] + \mathbb{P}\left[\bar{M}_{\jmath_n(y)} \geq y\right] \\ \overset{\text{Lemma 2.2.2}}{=} & \mathbb{P}\left[\bar{M}_n \geq y\right] \\ \overset{\text{Prop. 2.2.4}}{=} & K_n(y) \\ \overset{(2.1.50)}{=} & -c_n'(\xi_n(y)) + c_{\jmath_n(y)}'(\xi_n(y)) + K_{\jmath_n(y)}(y), \end{aligned}$$

where we used in the last equality that $\xi_n(y)$ is not an atom of any $\mu_1, \dots, \mu_n$. This implies by the induction hypothesis that

$$\mathbb{P}\left[M_n \geq \xi_n(y)\right] = -c_n'(\xi_n(y)) = \mu_n([\xi_n(y), \infty)),$$

that is, we have matched the distribution of M_n to μ_n at almost all points inside the support. The embedding property follows.

Now we prove uniform integrability by applying a result from Azéma et al. [5] which states that if

$$\lim_{x\to\infty} x\mathbb{P}\left[|\bar{B}|_{\tau_n} \geq x\right] = 0 \tag{2.2.25}$$

then $(B_{\tau_n \wedge t})_{t\geq 0}$ is uniformly integrable.

Let us verify (2.2.25). Set $H_x = \inf\{t > 0 : B_t = x\}$. We have (here ξ_i^{-1} denotes the left-continuous inverse of ξ_i)

$$\begin{aligned}\mathbb{P}\left[|\bar{B}|_{\tau_n} \geq x\right] &\leq \mathbb{P}\left[H_{-x} < H_{\max_{i\leq n}\xi_i^{-1}(-x)}\right] + \mathbb{P}\left[\bar{B}_{\tau_n} \geq x\right] \\ &= \frac{\max_{i\leq n}\xi_i^{-1}(-x)}{x + \max_{i\leq n}\xi_i^{-1}(-x)} + K_n(x).\end{aligned}$$

From the definition of ξ_n, cf. (2.1.4), and the properties of b_i, cf. (2.1.9) we have

$$0 \leq \max_{i\leq n}\xi_i^{-1}(-x) \leq \max_{i\leq n} b_i(-x) \underset{x\to\infty}{\longrightarrow} 0$$

and hence, recalling the definition of μ_n^{HL} in (2.1.22),

$$\lim_{x\to\infty} x\mathbb{P}\left[|\bar{B}|_{\tau_n} \geq x\right] \leq \lim_{x\to\infty} xK_n(x) \leq \lim_{x\to\infty} x\frac{c_n(b_n^{-1}(x))}{x - b_n^{-1}(x)} = \lim_{x\to\infty} x\mu_n^{\mathrm{HL}}([x,\infty)) = 0.$$

This finishes the proof. □

2.3 Discussion of Assumption $\circledast$ and Extensions

In this section we focus on our main technical assumption so far: the condition (ii) in Assumption $\circledast$. We construct a simple example of probability measures μ_1, μ_2, μ_3 which violate the condition and where the stopping boundaries ξ_1, ξ_2, ξ_3, obtained via (2.1.4), fail to embed (μ_1, μ_2, μ_3). It follows that the assumption is not merely technical but does rule out certain type of interdependence between the marginals. If it is not satisfied then it may not be enough to perturb the measures slightly to satisfy it.

We then present an extension of our embedding, in the case $n = 3$, which works in all generality. More precisely, we show how to modify the optimisation problem from which ξ_3 is determined in order to obtain the embedding property. The general embedding, as compared to the embedding in the presence of Assumption $\circledast$(ii), gains an important degree of freedom and becomes less explicit. In consequence it is also much harder to implement in practice, to the point that we do not believe this is worth pursuing for $n > 3$. This is also why, as well as for the sake of brevity, we keep the discussion in the section rather formal.

2.3.1 Counterexample for Assumption $\circledast$(ii)

In Figure 2.3.1 we define measures via their potentials

$$U\mu : \mathbb{R} \to \mathbb{R}, \qquad x \mapsto U\mu(x) := -\int_{\mathbb{R}} |u - x| \, \mu(\mathrm{d}u). \tag{2.3.1}$$

We refer to Obłój [73, Proposition 2.3] for useful properties of $U\mu$.

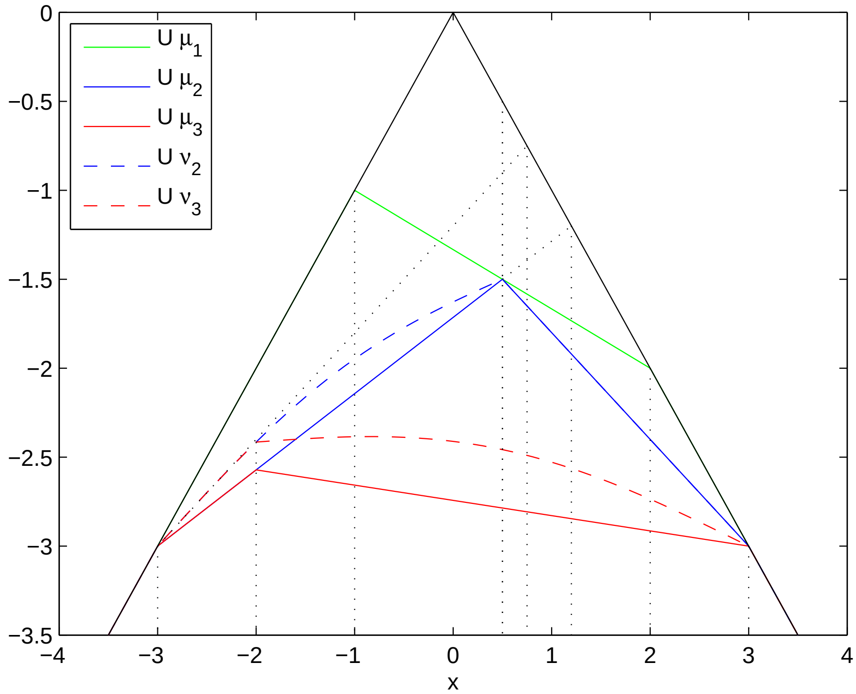

FIGURE 2.3.1: Potentials of $\mu_1, \mu_2, \mu_3, \nu_2$ and ν_3.

The measures with potentials illustrated in Figure 2.3.1 are given as

$$\mu_1(\{-1\}) = \frac{2}{3}, \quad \mu_1(\{2\}) = \frac{1}{3}, \tag{2.3.2a}$$

$$\mu_2(\{-3\}) = \frac{2}{7}, \quad \mu_2\left(\left\{\frac{1}{2}\right\}\right) = \frac{18}{35}, \quad \mu_2(\{3\}) = \frac{1}{5}, \tag{2.3.2b}$$

$$\mu_3(\{-3\}) = \frac{2}{7}, \quad \mu_3(\{-2\}) = \frac{9}{35}, \quad \mu_3(\{3\}) = \frac{16}{35}. \tag{2.3.2c}$$

Observe that the embedding for (μ_1, μ_2, μ_3) is unique: We write $H_{a,b}$ for the exit time of $[a, b]$ and denote $H_{a,b} \circ \theta_\tau := \inf\{t > \tau : B_t \notin (a, b)\}$. Then the embedding $(\tau_1, \tau_2', \tau_3)$ can be written as

$$\tau_1 = H_{-1,2}, \quad \tau_2' = H_{-3,\frac{1}{2}} \circ \theta_{\tau_1} \mathbb{1}_{\{B_{\tau_1}=-1\}} + H_{\frac{1}{2},3} \circ \theta_{\tau_1} \mathbb{1}_{\{B_{\tau_1}=2\}}, \quad \tau_3 = H_{-2,3} \circ \theta_{\tau_2'}. \tag{2.3.3}$$

As mentioned earlier, our construction yields the same first two stopping boundaries as the method of Brown et al. [18]. In this case, cf. Figure 2.3.2,

$$\xi_1(y) := \begin{cases} -1 & \text{if } y \in [0,2), \\ y & \text{else}, \end{cases} \qquad \xi_2(y) := \begin{cases} -3 & \text{if } y \in [0,\frac{1}{2}), \\ \frac{1}{2} & \text{if } y \in [\frac{1}{2},3), \\ y & \text{else}. \end{cases}$$

This already shows that our embedding fails to embed μ_2. To see this one just has to compare the stopping boundary ξ_2 in the Definition of τ_2 with (2.3.3). In Section 2.3.2 we will recall from Brown et al. [18] how the stopping time τ_2 has to be modified into τ_2', giving the stopping time above.

More importantly, the embedding for μ_3 fails because the optimization problem (2.1.4) does not return the third (unique) stopping boundary which is required for the embedding of (μ_1, μ_2, μ_3). Indeed, for sufficiently small $y > \frac{1}{2}$, in the region $\zeta < \min(\xi_1(y), \xi_2(y)) = -1$ we are looking at the minimization of $\zeta \mapsto \frac{c_3(\zeta)}{y-\zeta}$ which is attained by $\xi_3(y) = -3 < -2$ since μ_3 has an atom at -3. Consequently, there will be a positive probability to hit -3 after τ_2. This contradicts (2.3.2c). This, together with the correct third boundary η_3, is illustrated in Figure 2.3.2. Note also that $b_1 \leq b_3 = \xi_3^{-1}$ and hence, in some sense, it is the embedding of μ_2 which causes problems.

This example does not contradict our main result because Assumption $\circledast$(ii)(a) is not satisfied for $i = 2$ and $y = \frac{1}{2}$, where $\zeta = -3$ minimizes the objective function but $c_2\left(\frac{1}{2}\right) = c_1\left(\frac{1}{2}\right)$ holds. Our counterexample also shows that a "small perturbation" to (μ_1, μ_2, μ_3) does not remove the problem. Indeed, similar reasoning to the one above holds for measures (μ_1, ν_2, ν_3) defined by their potentials in Figure 2.3.1. Assumption $\circledast$ rules out a certain type of subtle structure between the marginals and not only some "isolated" or "singular" configurations of measures.

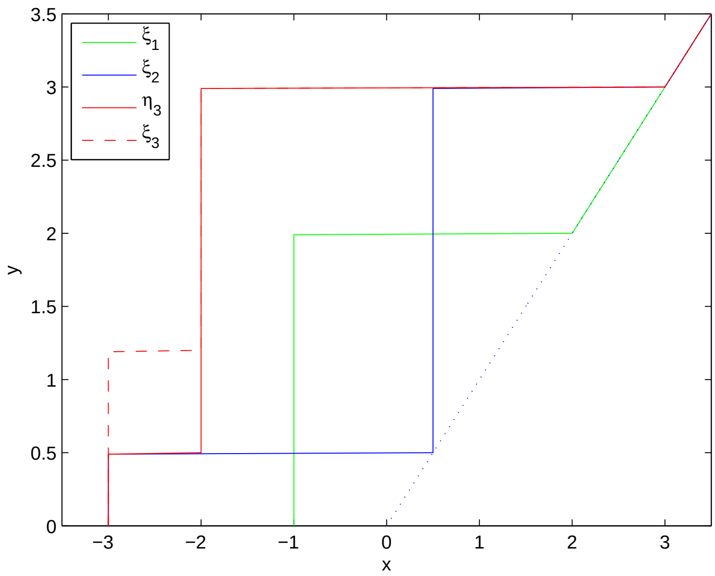

FIGURE 2.3.2: We illustrate the (unique) boundaries ξ_1, ξ_2, η_3 required for the embedding of (μ_1, μ_2, μ_3) from (2.3.2a)–(2.3.2c) and the stopping boundary ξ_3 obtained from (2.1.4). In order to ensure the embedding for μ_2, the mass stopped at τ_2 in -1 on the event $\{\bar{B}_{\tau_2} \in (1/2, 2)\}$ is diffused to -3 or to $1/2$ at τ_2', without affecting the maximum: $\bar{B}_{\tau_2} = \bar{B}_{\tau_2'}$. Note that the case $\xi_2(y) = y$, here for $y = 1/2$, is possible and required to define the embedding. After τ_2' we need to define τ_3 which embeds μ_3 which here is implied directly by (2.3.3). In Section 2.3.2 we develop arguments which generalise this.

2.3.2 Sketch for General Embedding in the Case $n = 3$

In the example of the measures (μ_1, μ_2, μ_3) from (2.3.2a)–(2.3.2c) the (unique) embedding could still be seen as a type of "iterated Azéma-Yor type embedding" although it does not satisfy the relations from Lemma 2.2.1. Consequently, one might conjecture that a modification of the optimization problem (2.1.4) and a relaxation of Lemma 2.2.1 might lead to a generally applicable embedding. We now explain in which sense this is true. Our aim is to outline new ideas and arguments which are needed. The technical details quickly become very involved and lengthy. In the sake of brevity, but also to better illustrate the main points, we restrict ourselves to a formal discussion and the case $n = 3$.

In order to understand the problem in more detail, we need to recall from Brown et al. [18] how the embedding for μ_2 looks like in general. It reads

$$\tau_2^{\mathrm{BHR}} := \begin{cases} \tau_2' & \text{if } \bar{B}_{\tau_1} \in \left(\frac{1}{2}, 2\right], \\ \tau_2 & \text{else}, \end{cases} \tag{2.3.4}$$

where τ_2' is some stopping time with $\bar{B}_{\tau_2} = \bar{B}_{\tau_2'}$. Its existence is established by Brown et al. [18] by showing that the relative parts of the mass which are further diffused have the same mass, mean and are in convex order. In general there will be infinitely many such stopping times τ_2'. Although this is not true for (μ_1, μ_2, μ_3) in (2.3.2a)–(2.3.2c) because their embedding was unique, it is true for measures (μ_1, ν_2, ν_3) which are defined via their potentials in Figure 2.3.1.

Let ξ_1 and ξ_2 be defined as in (2.1.4) and let $M_2 = B_{\tau_2^{\mathrm{BHR}}}$. Now our goal is to define an embedding $\tilde{\tau}_3$ for the third marginal on top of the embedding of Brown et al. [18] in a situation as in (μ_1, ν_2, ν_3). We still want to define our iterated Azéma-Yor type embedding through a stopping rule based on some stopping boundary $\tilde{\xi}_3$ as a first exit time,

$$\tilde{\tau}_3 := \begin{cases} \inf\left\{t \geq \tau_2^{\mathrm{BHR}} : B_t \leq \tilde{\xi}_3(\bar{B}_t)\right\} & \text{if } B_{\tau_2^{\mathrm{BHR}}} > \tilde{\xi}_3(\bar{B}_{\tau_2^{\mathrm{BHR}}}), \\ \tau_2^{\mathrm{BHR}} & \text{else}, \end{cases} \tag{2.3.5}$$

and prove that this is a valid embedding of μ_3. We observe that now the choice of τ_2' in the definition of τ_2^{BHR} may matter for the subsequent embedding. Similarly as in the embedding of Brown et al. [18] we expect that this will be only possible if the procedure which produces $\tilde{\xi}_3$ yields a continuous $\tilde{\xi}_3$. Otherwise an additional step, producing a stopping time $\tau_3' \geq \tilde{\tau}_3$ would be required and further complicate the presentation.

With this, a more canonical approach in the context of Lemma 2.2.1 is to write

$$\mathbb{P}\left[\bar{M}_3 \geq y\right] = \mathbb{P}\left[M_3 \geq \tilde{\xi}_3(y)\right] + \text{“error-term”}, \tag{2.3.6}$$

which we formalize in (2.3.24). As it will turn out, this “error-term” provides a suitable “book-keeping procedure” to keep track of the masses in the embedding. We proceed along the lines of the proof of our main result. For simplicity, we further assume that ξ_2 has only one discontinuity, i.e. $\underline{z} := \xi_2(\underline{y}-) < \xi_2(\underline{y}+) := \overline{z}$ for some $\underline{y} \geq 0$ and we let $\bar{y} := \xi_1^{-1}(\overline{z})$. As explained below, this is not restrictive since our procedure is localised. If $\bar{y} \leq \underline{y}$ then μ_1 can be “ignored” and the results of Brown et al. [18] apply. Hence we assume $\bar{y} > \underline{y}$.

2.3.2.1 Redefining ξ_3 and K_3

Define the following auxiliary terms,

$$F(\zeta, y; \tau_2') \quad := \mathbb{1}_{\{\bar{M}_1 \geq y\}}\left(\zeta - M_2\right)^+, \tag{2.3.7}$$

$$f^{\mathrm{iAY}}(\zeta, y; \tau_2') := \mathbb{E}\left[F(\zeta, y; \tau_2')\right]. \tag{2.3.8}$$

As the notation underlines, these quantities may depend on the additional choice of stopping time τ_2' between τ_2 and τ_3. Note that for $\zeta \in [\underline{z}, \overline{z}]$ and $y \in [\underline{y}, \bar{y}]$,

$$\frac{\partial f^{\mathrm{iAY}}}{\partial \zeta}(\zeta, y; \tau_2') = \mathbb{P}\left[\bar{M}_1 \geq y, M_2 < \zeta\right], \tag{2.3.9}$$

and

$$\begin{aligned} \frac{\partial f^{\mathrm{iAY}}}{\partial y}(\zeta, y; \tau_2') &= -\mathbb{E}\left[\frac{\mathbb{1}_{\{\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta\}}}{\mathrm{d}y}\right]\zeta + \mathbb{E}\left[\frac{\mathbb{1}_{\{\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta\}}}{\mathrm{d}y} M_2\right] \\ &= -\Big(\zeta - \alpha(\zeta, y; \tau_2')\Big)\frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta\right]}{\mathrm{d}y} \end{aligned} \tag{2.3.10}$$

where

$$\alpha(\zeta, y; \tau_2') := \mathbb{E}\left[M_2 \middle| \bar{M}_1 = y, M_2 < \zeta\right], \tag{2.3.11a}$$

$$\beta(\zeta, y; \tau_2') := \mathbb{E}\left[M_2 \middle| \bar{M}_1 = y, M_2 \geq \zeta\right]. \tag{2.3.11b}$$

With these definitions we have by the properties of τ_2',

$$\begin{aligned} \alpha(\zeta, y; \tau_2') \frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta\right]}{\mathrm{d}y} &+ \beta(\zeta, y; \tau_2') \frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y, M_2 \geq \zeta\right]}{\mathrm{d}y} \\ = \xi_1(y) \frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y\right]}{\mathrm{d}y}. \end{aligned} \tag{2.3.12}$$

We now redefine ξ_3 and K_3 from (2.1.4) and (2.1.6), respectively, and denote the new definition by $\tilde{\xi}_3$ and $\tilde{K}_3$. To this end, introduce the function

$$\tilde{c}^3(\zeta, y) := \begin{cases} c_3(\zeta) - f^{\mathrm{iAY}}(\zeta, y; \tau_2') & \text{if } \underline{z} \leq \zeta \leq \overline{z}, \underline{y} \leq y \leq \bar{y}, \quad (2.3.13a) \\ c^3(\zeta, y) & \text{else.} \quad (2.3.13b) \end{cases}$$

We have that $\tilde{c}^3$ is continuous and $\tilde{c}^3 \leq c^3$. Using the properties of τ_2' this can be seen from the following:

$$\begin{aligned} f^{\mathrm{iAY}}(\zeta, y; \tau_2') &= \mathbb{E}\left[(\zeta - M_2)^+ \mathbb{1}_{\{\bar{M}_1 \geq y\}}\right] = \mathbb{E}\left[\left\{(M_2 - \zeta)^+ - (y - \zeta)\right\} \mathbb{1}_{\{\bar{M}_1 \geq y\}}\right] \\ &= \mathbb{E}\left[(M_2 - \zeta)^+ \mathbb{1}_{\{\bar{M}_1 \geq y\}}\right] - (y - \zeta) K_1(y) \\ &\geq \max\left(c_1(\zeta) - (y - \zeta) K_1(y), 0\right) \end{aligned} \tag{2.3.14}$$

for $\zeta \in [\underline{z}, \overline{z}]$ and $y \in [\underline{y}, \bar{y}]$, with equality for $\zeta = \overline{z}$. Continuity at $\zeta = \underline{z}$ holds by the properties of τ_2' which yield $f^{\mathrm{iAY}}(\underline{z}, y; \tau_2') = 0$. For $y = \bar{y}$ we also have $f^{\mathrm{iAY}}(\zeta, \bar{y}; \tau_2') = 0$. As for continuity at $y = \underline{y}$ we take ξ_2 left-continuous, i.e. in particular $\xi_2(\underline{y}-) = \xi_2(\underline{y})$. Then, it is enough to observe by direct computation that

$$\mathbb{E}\left[(M_2 - \zeta)^+ \mathbb{1}_{\{\bar{M}_1 \geq \underline{y}\}}\right] = c_2(\zeta) - (\underline{y} - \zeta)(K_2(\underline{y}) - K_1(\underline{y})).$$

As before, let

$$\tilde{\xi}_3(y) := \arg\min_{\zeta \leq y} \frac{\tilde{c}^3(\zeta, y)}{y - \zeta} \tag{2.3.15}$$

and

$$\tilde{K}_3(y) := \frac{\tilde{c}^3(\tilde{\xi}_3(y), y)}{y - \tilde{\xi}_3(y)}. \tag{2.3.16}$$

It is clear that a discontinuity of ξ_2 results in a local perturbation of c^3 into $\tilde{c}^3$ and in consequence of ξ_3 into $\tilde{\xi}_3$. If ξ_2 has multiple discontinuities the construction above applies to each of them giving a global definition of $\tilde{c}^3$. Then $\tilde{K}_3$ and $\tilde{\xi}_3$ are defined as above.

2.3.2.2 Law of the Maximum

In the following we assume that $y \in [\underline{y}, \bar{y}]$ and $\tilde{\xi}_3(y), \zeta \in [\underline{z}, \overline{z}]$. Otherwise $\tilde{c}^3 = c^3$ and the arguments from Sections 2.1 and 2.2 apply. We have $\tilde{\xi}_3(y) < \xi_2(y)$ and $\bar{M}_1 = \bar{M}_2$ on $\{\bar{M}_1 \in [\underline{y}, \bar{y}]\}$.

Note the obvious decomposition

$$\mathbb{P}\left[\bar{M}_3 \geq y\right] = \mathbb{P}\left[\bar{M}_1 < y, \bar{M}_3 \geq y\right] + \mathbb{P}\left[\bar{M}_1 \geq y\right].$$

We compute by similar excursion theoretical arguments as in the proof of Lemma 2.2.3,

$$\begin{aligned} \left.\frac{\partial \mathbb{P}\left[\bar{M}_1 < y, \bar{M}_3 \geq m\right]}{\partial y}\right|_{m=y} &=: p(\tilde{\xi}_3(y), y; \tau_2') \\ =& \frac{\mathbb{P}\left[M_2 \geq \tilde{\xi}_3(y), \bar{M}_1 \in \mathrm{d}y\right]}{\mathrm{d}y} \frac{\beta(\tilde{\xi}_3(y), y; \tau_2') - \tilde{\xi}_3(y)}{y - \tilde{\xi}_3(y)}. \end{aligned} \tag{2.3.17}$$

In analogy to (2.2.13), and because $\tilde{\xi}_3(y) < \xi_2(y)$,

$$\left.\frac{\partial \mathbb{P}\left[\bar{M}_1 < m, \bar{M}_3 \geq y\right]}{\partial y}\right|_{m=y} = -\frac{\mathbb{P}\left[\bar{M}_3 \geq y\right] - \mathbb{P}\left[\bar{M}_1 \geq y\right]}{y - \tilde{\xi}_3(y)}.$$

Hence, combining the above

$$\begin{aligned}
\frac{\partial}{\partial y}\mathbb{P}\left[\bar{M}_3 \geq y\right] &= p(\tilde{\xi}_3(y), y; \tau_2') - \frac{\mathbb{P}\left[\bar{M}_3 \geq y\right] - \mathbb{P}\left[\bar{M}_1 \geq y\right]}{y - \tilde{\xi}_3(y)} + \frac{\partial \mathbb{P}\left[\bar{M}_1 \geq y\right]}{\partial y} \\
&\overset{(2.2.10)}{=} -\frac{\mathbb{P}\left[\bar{M}_3 \geq y\right]}{y - \tilde{\xi}_3(y)} - \frac{\tilde{\xi}_3(y) - \xi_1(y)}{y - \tilde{\xi}_3(y)}\frac{\partial \mathbb{P}\left[\bar{M}_1 \geq y\right]}{\partial y} + p(\tilde{\xi}_3(y), y; \tau_2').
\end{aligned} \tag{2.3.18}$$

In the redefined domain the first order condition for optimality of $\tilde{\xi}_3(y)$ reads

$$\tilde{K}_3(y) + c_3'(\tilde{\xi}_3(y)) - \frac{\partial f^{\mathrm{iAY}}}{\partial \zeta}(\tilde{\xi}_3(y), y; \tau_2') = 0. \tag{2.3.19}$$

By similar calculations as in (A.1.17) below we have

$$\begin{aligned}
\tilde{K}_3'(y) &\overset{(2.3.19)}{=} -\frac{\tilde{K}_3(y)}{y - \tilde{\xi}_3(y)} - \frac{\frac{\partial f^{\mathrm{iAY}}}{\partial y}(\tilde{\xi}_3(y), y; \tau_2')}{y - \tilde{\xi}_3(y)} \\
&\overset{(2.3.10)}{=} -\frac{\tilde{K}_3(y)}{y - \tilde{\xi}_3(y)} + \frac{\tilde{\xi}_3(y) - \alpha(\tilde{\xi}_3(y), y; \tau_2')}{y - \tilde{\xi}_3(y)}\frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y, M_2 < \tilde{\xi}_3(y)\right]}{\mathrm{d}y} \\
&\overset{(2.3.12)}{=} -\frac{\tilde{K}_3(y)}{y - \tilde{\xi}_3(y)} + \frac{\tilde{\xi}_3(y) - \xi_1(y)}{y - \tilde{\xi}_3(y)}\frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y\right]}{\mathrm{d}y} \\
&\quad + \frac{\beta(\tilde{\xi}_3(y), y; \tau_2') - \tilde{\xi}_3(y)}{y - \tilde{\xi}_3(y)}\frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y, M_2 \geq \tilde{\xi}_3(y)\right]}{\mathrm{d}y} \\
&\overset{(2.3.17)}{=} -\frac{\tilde{K}_3(y)}{y - \tilde{\xi}_3(y)} - \frac{\tilde{\xi}_3(y) - \xi_1(y)}{y - \tilde{\xi}_3(y)}\frac{\partial \mathbb{P}\left[\bar{M}_1 \geq y\right]}{\partial y} + p(\tilde{\xi}_3(y), y; \tau_2').
\end{aligned} \tag{2.3.20}$$

Consequently, by comparing (2.3.18) and (2.3.20), and in conjunction with Proposition 2.2.4, we obtain

$$\tilde{K}_3(y) = \mathbb{P}\left[\bar{M}_3 \geq y\right] \qquad \forall y \geq 0. \tag{2.3.21}$$

2.3.2.3 Embedding Property

After having found the distribution of the maximum, the final step is to prove the embedding property. To achieve this we will need that $\tilde{\xi}_3$ is non-decreasing.

Recall the first order condition of optimality of $\tilde{\xi}_3$ in (2.3.19) and then the second order condition for optimality of $\tilde{\xi}_3(y)$ reads

$$c_3''(\tilde{\xi}_3(y)) - \frac{\partial^2 f^{\text{iAY}}}{\partial \zeta^2}(\tilde{\xi}_3(y), y; \tau_2') \geq 0. \tag{2.3.22}$$

Now, differentiating (2.3.19) in y yields

$$\tilde{K}_3'(y) + c_3''(\tilde{\xi}_3(y))\tilde{\xi}_3'(y) - \frac{\partial^2 f^{\text{iAY}}}{\partial \zeta^2}(\tilde{\xi}_3(y), y; \tau_2')\tilde{\xi}_3'(y) - \frac{\partial^2 f^{\text{iAY}}}{\partial \zeta \partial y}(\tilde{\xi}_3(y), y; \tau_2') = 0$$

or equivalently,

$$\tilde{\xi}_3'(y) \underbrace{\left(c_3''(\tilde{\xi}_3(y)) - \frac{\partial^2 f^{\text{iAY}}}{\partial \zeta^2}(\tilde{\xi}_3(y), y; \tau_2') \right)}_{\geq 0 \text{ by } (2.3.22)} = -\tilde{K}_3'(y) + \frac{\partial^2 f^{\text{iAY}}}{\partial \zeta \partial y}(\tilde{\xi}_3(y), y; \tau_2').$$

In order to formally infer

$$\tilde{\xi}_3'(y) \geq 0$$

we require

$$-\tilde{K}_3'(y) + \frac{\partial^2 f^{\text{iAY}}}{\partial \zeta \partial y}(\tilde{\xi}_3(y), y; \tau_2') \geq 0. \tag{2.3.23}$$

Direct computation shows that

$$\frac{\partial^2 f^{\text{iAY}}}{\partial \zeta \partial y}(\zeta, y; \tau_2') = -\frac{\mathbb{P}\left[\bar{M}_1 \in \mathrm{d}y, M_2 < \zeta\right]}{\mathrm{d}y}$$

and by (2.3.21),

$$-\tilde{K}_3'(y) = \frac{\mathbb{P}\left[\bar{M}_3 \in \mathrm{d}y\right]}{\mathrm{d}y}$$

which implies (2.3.23) and hence that $\tilde{\xi}_3$ is non-decreasing.

By definition of the embedding in (2.3.5), and since $\tilde{\xi}_3$ is non-decreasing, we have

$$\begin{aligned}\mathbb{P}\left[\bar{M}_3 \geq y\right] &= \mathbb{P}\left[M_3 \geq \tilde{\xi}_3(y)\right] + \mathbb{P}\left[\bar{M}_3 \geq y, M_3 < \tilde{\xi}_3(y)\right] \\ &= \mathbb{P}\left[M_3 \geq \tilde{\xi}_3(y)\right] + \mathbb{P}\left[\bar{M}_1 \geq y, M_2 < \tilde{\xi}_3(y)\right] \\ &\overset{(2.3.9)}{=} \mathbb{P}\left[M_3 \geq \tilde{\xi}_3(y)\right] + \frac{\partial f^{\mathrm{iAY}}}{\partial \zeta}(\tilde{\xi}_3(y), y; \tau_2'). \end{aligned} \tag{2.3.24}$$

and then, by (2.3.21), (2.3.19) and (2.3.24),

$$-c_3'(\tilde{\xi}_3(y)) = \mathbb{P}\left[M_3 \geq \tilde{\xi}_3(y)\right]$$

which is the desired embedding property.

The above construction hinged on the appropriate choice of the auxiliary term F in (2.3.7) whose expectation allows for the error book keeping, as suggested in (2.3.6). We identified the correct F by analysing the "error terms" which cause strict inequality for $(B_u : u \leq \tau_2')$ in the pathwise inequality (4.1) of Henry-Labordère et al. [49]. This is natural since this inequality is used to prove optimality of our embedding. It gives an upper bound but fails to be sharp if condition (ii) in Assumption ⊛ does not hold. In order to recover a sharp bound one has to look at the error terms causing strict inequality when Assumption ⊛ fails. The same principle applies for $n > 3$. However then interactions between discontinuities of boundaries $\xi_2, \tilde{\xi}_3$ etc come into play and the relevant terms become very involved. The construction would become increasingly technical and implicit and we decided to stop at this point.

2.4 Numerical Example

In our numerical example we take an iterated Azéma-Yor type embedding based on the stopping boundaries $\xi_1, \ldots, \xi_4$ displayed in Figure 2.4.1. These boundaries have been chosen arbitrarily for the purpose of this example. Our goal is to illustrate the laws $\mu_1, \ldots, \mu_4$ and the law of the maximum, K_4, of this iterated Azéma-Yor type embedding.

Approaching this task numerically, in the first step we obtain estimates $\hat{c}_1, \ldots, \hat{c}_4$ of the true call prices $c_1, \ldots, c_4$. The estimated call prices are shown in terms of their Black-Scholes implied volatilities in Figure 2.4.2.

Implicitly assuming that the laws $\mu_1, \ldots, \mu_4$ of the iterated Azéma-Yor type embedding based on our (continuous and increasing) boundaries $\xi_1, \ldots, \xi_4$ satisfy Assumption $\circledast$, in the second step we use Definition 2.1.1 to compute the stopping boundaries $\hat{\xi}_1, \ldots, \hat{\xi}_4$ and the tail-distribution of the maximum $\hat{K}_4$ based on $\hat{c}_1, \ldots, \hat{c}_4$. In Figure 2.4.1 we visualize that the true and computed boundaries are numerically indistinguishable. The computed tail-distribution of the maximum at time $t = 4$, $\hat{K}_4$, is displayed in Figure 2.4.3.

In order to compare how much impact the intermediate laws μ_1, μ_2, μ_3 have on the law of the maximum at time $t = 4$ we compare with the one-marginal case, i.e. we discard c_1, c_2, c_3 and then compute the maximal law of the maximum. The result is displayed in Figure 2.4.3 and suggests that in this example the intermediate laws do not have a lot of impact.

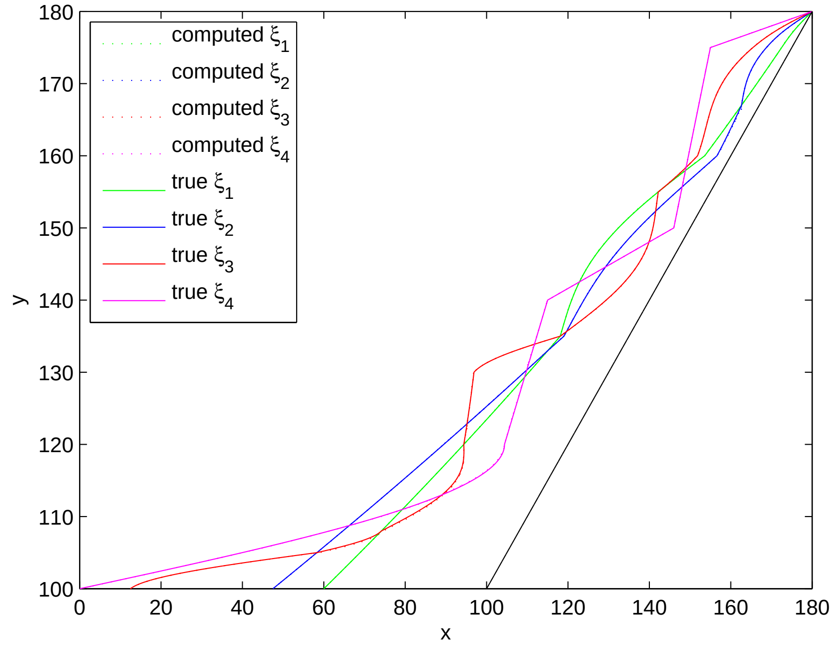

FIGURE 2.4.1: Stopping boundaries $\xi_1, \ldots, \xi_4$ used for the iterated Azéma-Yor type embedding.

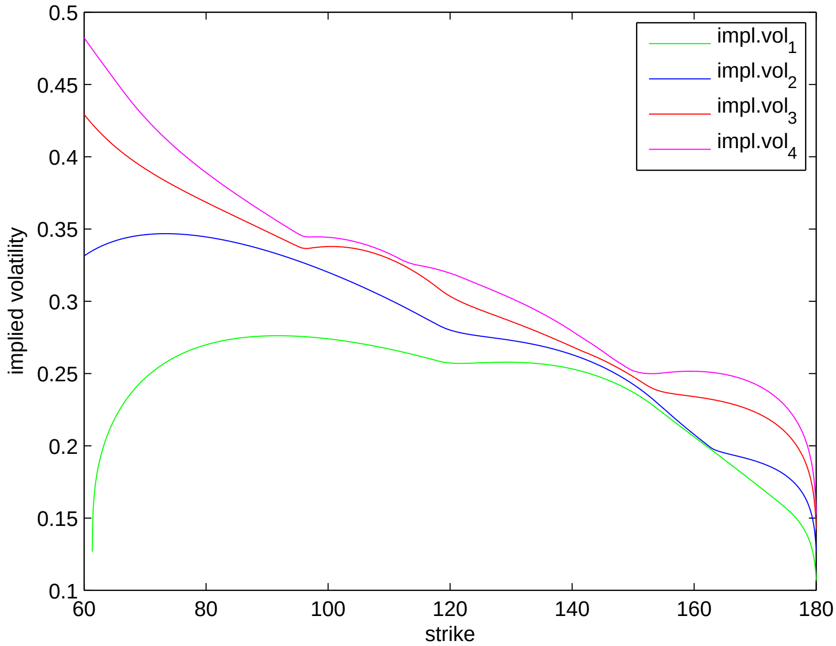

FIGURE 2.4.2: Black-Scholes implied volatilities of the marginal laws of the iterated Azéma-Yor type embedding based on $\xi_1, \ldots, \xi_4$ from Figure 2.4.1.

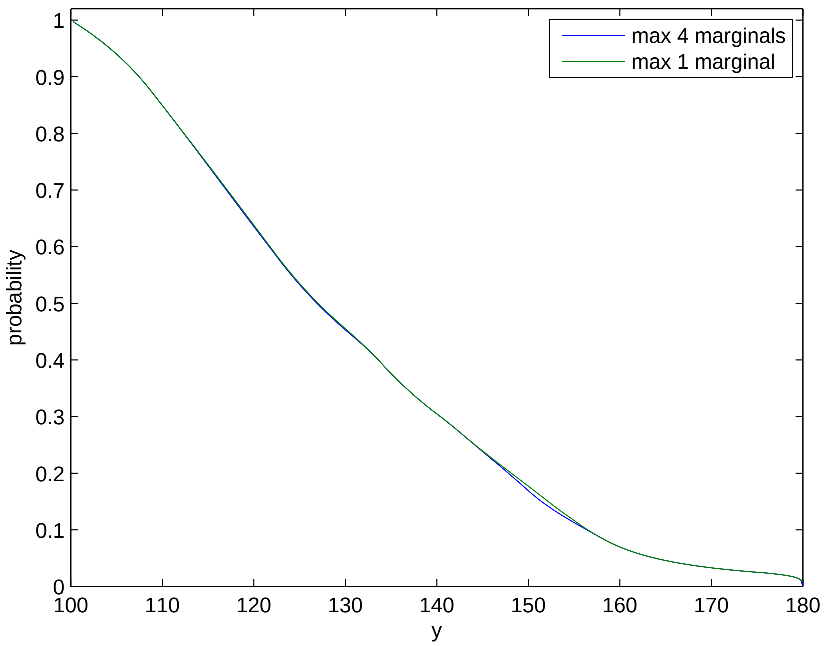

FIGURE 2.4.3: We illustrate the impact of the intermediate marginal laws μ_1, μ_2, μ_3 on the law of the terminal maximum $\mathbb{P}\left[\bar{M}_4 \geq y\right]$. In this example these intermediate laws do not have a lot of impact.

Chapter 3

Optimality of the Iterated Azéma-Yor Type Embedding

Joint Publication Certain results of this chapter are submitted as part of a joint publication with Pierre Henry-Labordère, Jan Obłój and Nizar Touzi, see Henry-Labordère et al. [49].

The problem of controlling the maximum of a continuous martingale X using its terminal distribution has a long and rich history, starting with Doob's maximal and L^p inequalities. In seminal contributions, Blackwell and Dubins [14] and Azéma and Yor [4] established that the distribution of the maximum $\bar{X}_T := \sup_{t\leq T} X_t$ of a martingale X is bounded above, in stochastic order, by the so called Hardy-Littlewood transform of the distribution of X_T, and the bound is attained. This led to series of studies on the possible distributions of $(X_T, \bar{X}_T)$ including Gilat and Meilijson [45], Kertz and Rösler [57, 58]; Kertz and Rösler [59], Rogers [84], Vallois [94], see also Carraro et al. [20]. More recently, such problems appeared very naturally within the field of mathematical finance, as we explain below, which motivated further developments. The original result was generalised to the case of a non trivial starting law by Hobson [51] and to the case of a fixed intermediate law by Brown et al. [18].

Considerable effort has been put in the previous chapter to obtain the iterated Azéma-Yor type embedding. In this chapter we argue that this embedding is special in many ways. We prove that it attains the *maximum maximum* of a martingale constrained by finitely many intermediate marginal laws. Further, after developing some financial interpretation, we identify it as the extremal model in the robust superhedging problem of the Lookback option. This will allow us to derive explicitly the optimal superhedging strategy.

The key ingredient for these results is a certain trajectorial inequality. Once postulated, this inequality is elementary to verify but it may be hard to obtain it "barehanded". Therefore, in Section 3.1 we give some details on how this inequality was found in first place by a mix of arguments coming from stochastic control and intuition from Mathematical Finance.

3.1 Stochastic Control Methodology

A stochastic control methodology was the starting point for both the pathwise arguments in Section 3.2 and for the construction of the embedding in Chapter 2. In the context of robust superhedging of simple barrier options, it allowed us to single out the static position of the optimal superhedging strategy. Intuition from the two-marginal ($n = 2$) case, see Brown et al. [18], allowed to complement this with a dynamic part.

In the following, we briefly sketch the main ideas behind the stochastic control ansatz, for a full account we refer to Henry-Labordère et al. [49].

Writing $\phi^{\boldsymbol{\lambda}}(\boldsymbol{x}, m) = \phi(m) - \sum_{i=1}^{n} \lambda_i(x_{t_i})$, $X_{\boldsymbol{t}} = (X_{t_1}, \ldots, X_{t_n})$ and $\mu_i(\lambda_i) = \int \lambda_i \mathrm{d}\mu_i, t_n = T$, we proved the duality

$$
\begin{aligned}
\text{q.s.-superhedging cost of } \phi(\bar{X}_T) \quad &= \quad \inf_{\boldsymbol{\lambda} \in \Lambda_n^{\boldsymbol{\mu}}} \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}} \left[\phi^{\boldsymbol{\lambda}}(X_{\boldsymbol{t}}, \bar{X}_T) \right] + \sum_{i=1}^{n} \mu_i(\lambda_i) \\
&= \quad \inf_{\boldsymbol{\lambda} \in \Lambda_n^{\boldsymbol{\mu}}} \sup_{\boldsymbol{\tau} \in \mathcal{T}_\infty^n} \mathbb{E}_{\mathbb{P}_0} \left[\phi^{\boldsymbol{\lambda}}(B_{\boldsymbol{\tau}}, \bar{B}_{\tau_n}) \right] + \sum_{i=1}^{n} \mu_i(\lambda_i) \\
&=: \quad \inf_{\boldsymbol{\lambda} \in \Lambda_n^{\boldsymbol{\mu}}} \left(u^{\boldsymbol{\lambda}}(X_0, X_0) + \boldsymbol{\mu}(\boldsymbol{\lambda}) \right),
\end{aligned}
$$

where $\Lambda_n^{\boldsymbol{\mu}}$ is a suitable set of Lagrange multipliers $\boldsymbol{\lambda}$, $\mathcal{P}$ is the set of martingale measures $\mathbb{P}$ for the canonical process X, $\mathcal{T}_\infty^n$ is the set of ordered stopping times $\tau_1 \leq \cdots \leq \tau_n$ and B is a $\mathbb{P}_0$-Brownian motion. Following Galichon et al. [44] it turned out to be possible to obtain by means of dynamic programming that $u^{\boldsymbol{\lambda}} \leq v^{\boldsymbol{\psi}}$ holds for some explicitly given function $v^{\boldsymbol{\psi}}$ with a free boundary $\boldsymbol{\psi}$ which is in one-to-one correspondence with $\boldsymbol{\lambda}$. A computation gave (essentially)

$$
u^{\boldsymbol{\lambda}}(X_0, X_0) + \boldsymbol{\mu}(\boldsymbol{\lambda}) \leq \phi(X_0) + \int_{X_0}^{\infty} \left(\frac{c_i(\psi_i(m))}{m - \psi_i(m)} - \frac{c_i(\psi_{i+1}(m))}{m - \psi_{i+1}(m)} \right) \mathrm{d}\phi(m).
$$

Pointwise minimization over $\boldsymbol{\psi}$ yield $\boldsymbol{\psi}^\star$ which in turn yield $\boldsymbol{\lambda}^\star$ by the one-to-one correspondence of $\boldsymbol{\psi}$ and $\boldsymbol{\lambda}$.

As we will explain in more detail in Section 3.4, the Lagrange multipliers $\boldsymbol{\lambda}$ encode the intermediate marginal constraints and have the financial interpretation as the static positions in Vanilla options. Hence, this method identified a candidate optimal static position given by $\boldsymbol{\lambda}^\star$ and the dynamic part could be guessed, giving the key Proposition 3.2.1 below.

3.2 A Trajectorial Inequality

The following trajectorial inequality is the building block for optimality of the iterated Azéma-Yor type embedding, robust superhedging and pathwise duality of the Lookback option in the n-marginal case.

Proposition 3.2.1 (Trajectorial Inequality). *Let ω be a càdlàg path and denote $\bar{\omega}_t := \sup_{0\le s\le t}\omega_s$.*

Then, for $m \ge \omega_0$, $t_1 \le \cdots \le t_n = T$ and $\zeta_1 \le \cdots \le \zeta_n < m$:

$$\mathbb{1}_{\{\bar{\omega}_{t_n}\ge m\}} \le \Upsilon_n(\omega, m, \boldsymbol{\zeta}) := \sum_{i=1}^{n}\left(\frac{(\omega_{t_i}-\zeta_i)^+}{m-\zeta_i} + \mathbb{1}_{\{\bar{\omega}_{t_{i-1}}<m\le\bar{\omega}_{t_i}\}}\frac{m-\omega_{t_i}}{m-\zeta_i}\right) - \sum_{i=1}^{n-1}\left(\frac{(\omega_{t_i}-\zeta_{i+1})^+}{m-\zeta_{i+1}} + \mathbb{1}_{\{m\le\bar{\omega}_{t_i},\zeta_{i+1}\le\omega_{t_i}\}}\frac{\omega_{t_{i+1}}-\omega_{t_i}}{m-\zeta_{i+1}}\right). \tag{3.2.1}$$

Proof. Denote by $\Upsilon_n = \Upsilon_n(\omega, m, \boldsymbol{\zeta})$ the right hand-side of (3.2.1), and let us prove the claim by induction. First, in the case $n = 1$, the required inequality is the same as that of Brown et al. [17, Lemma 2.1]:

$$\Upsilon_1 = \frac{(\omega_{t_1}-\zeta_1)^+ + \mathbb{1}_{\{\bar{\omega}_{t_0}<m\le\bar{\omega}_{t_1}\}}(m-\omega_{t_1})}{m-\zeta_1} \ge \frac{\omega_{t_1}-\zeta_1+m-\omega_{t_1}}{m-\zeta_1}\mathbb{1}_{\{m\le\bar{\omega}_{t_1}\}} \ge \mathbb{1}_{\{m\le\bar{\omega}_{t_1}\}}. \tag{3.2.2}$$

We next assume that $\Upsilon_{n-1} \ge \mathbb{1}_{\{\bar{\omega}_{t_{n-1}}\ge m\}}$ and show that $\Upsilon_n \ge \mathbb{1}_{\{\bar{\omega}_{t_n}\ge m\}}$. We consider two cases.

Case 1: $\bar{\omega}_{t_{n-1}} \ge m$. Then $\bar{\omega}_{t_n} \ge m$, and it follows from the induction hypothesis that $1 = \mathbb{1}_{\{\bar{\omega}_{t_n}\ge m\}} = \mathbb{1}_{\{\bar{\omega}_{t_{n-1}}\ge m\}} \le \Upsilon_{n-1}$. In order to see that $\Upsilon_{n-1} \le \Upsilon_n$, we compute directly that, in the present case,

$$\Upsilon_n - \Upsilon_{n-1} = \frac{\omega_{t_n}-\zeta_n}{m-\zeta_n}\left(\mathbb{1}_{\{\omega_{t_n}\ge\zeta_n\}} - \mathbb{1}_{\{\omega_{t_{n-1}}\ge\zeta_n\}}\right) \ge 0. \tag{3.2.3}$$

Case 2: $\bar{\omega}_{t_{n-1}} < m$. As $\bar{\omega}$ is non-decreasing, it follows that $\bar{\omega}_{t_i} < m$ for all $i \leq n-1$. With a direct computation we obtain

$$\Upsilon_n = \Upsilon_n^0 + \frac{(\omega_{t_n} - \zeta_n)^+}{m - \zeta_n} + \mathbb{1}_{\{m \leq \bar{\omega}_{t_n}\}} \frac{m - \omega_{t_n}}{m - \zeta_n},$$

where

$$\Upsilon_n^0 := \sum_{i=1}^{n-1} \left(\frac{(\omega_{t_i} - \zeta_i)^+}{m - \zeta_i} - \frac{(\omega_{t_i} - \zeta_{i+1})^+}{m - \zeta_{i+1}} \right). \tag{3.2.4}$$

Since $m > \bar{\omega}_{t_i} \geq \omega_{t_i}$ for $i \leq n-1$, the functions $\zeta \longmapsto (\omega_{t_i} - \zeta)^+/(m - \zeta)$ are non-increasing. This implies that $\Upsilon_n^0 \geq 0$ by the fact that $\zeta_i \leq \zeta_{i+1}$ for all $i \leq n$. Then,

$$\begin{aligned} \Upsilon_n \geq & \frac{(\omega_{t_n} - \zeta_n)^+ + \mathbb{1}_{\{m \leq \bar{\omega}_{t_n}\}}(m - \omega_{t_n})}{m - \zeta_n} \geq \frac{(\omega_{t_n} - \zeta_n)^+ + m - \omega_{t_n}}{m - \zeta_n} \mathbb{1}_{\{m \leq \bar{\omega}_{t_n}\}} \\ \geq & \frac{\omega_{t_n} - \zeta_n + m - \omega_{t_n}}{m - \zeta_n} \mathbb{1}_{\{m \leq \bar{\omega}_{t_n}\}} = \mathbb{1}_{\{m \leq \bar{\omega}_{t_n}\}}. \end{aligned} \tag{3.2.5}$$

The claim follows. □

The next Proposition shows that the iterated Azéma-Yor type embedding has a special structure as it attains pathwise equality in (3.2.1).

Proposition 3.2.2 (Pathwise Equality)**.** *Let* $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$ *be non-decreasing and right-continuous and assume that* $(B_{t \wedge \tau_n} : t \geq 0)$ *is uniformly integrable, where* τ_n *is given by* (2.1.7a)–(2.1.7b)*. Set* $M = (M_{t_1}, \ldots, M_{t_n})$ *where* $M_{t_i} = B_{\tau_i}$*.*

Then, for any $m > M_0 := B_0$ *with* $\xi_n(m) < m$*,* M *achieves equality in* (3.2.1)*, i.e.*

$$\mathbb{1}_{\{\bar{M}_{t_n}(\omega) \geq m\}} = \Upsilon_n\big(M(\omega), m, \boldsymbol{\zeta}(m)\big), \qquad \forall \omega \in \Omega, \tag{3.2.6}$$

where

$$\zeta_i(m) = \min_{j \geq i} \xi_j(m), \qquad i = 1, \dots, n. \tag{3.2.7}$$

Proof. Fix $m \geq M_0$ and write for notational convenience $\zeta_i = \zeta_i(m)$, $\xi_i = \xi_i(m)$. Let $n_1 < \cdots < n_k = n$ be such that

$$\zeta_1 = \cdots = \zeta_{n_1} < \zeta_{n_1+1} = \cdots = \zeta_{n_2} < \cdots < \zeta_{n_{k-1}+1} = \cdots = \zeta_{n_k} = \zeta_n = \xi_n.$$

Then by

$$M_{t_j} \geq \zeta_j \qquad \Longrightarrow \qquad M_{t_l} \geq \zeta_j \qquad \forall l \geq j,$$

for all $j \leq n$, we obtain

$$\begin{aligned} \Upsilon_n(M, m, \boldsymbol{\zeta}) = & \sum_{j=1}^{k} \left(\frac{\left(M_{t_{n_j}} - \zeta_i\right)^+}{m - \zeta_{n_j}} + \mathbb{1}_{\left\{\bar{M}_{t_{n_{j-1}}} < m \leq \bar{M}_{t_{n_j}}\right\}} \frac{m - M_{t_{n_j}}}{m - \zeta_{n_j}} \right) \\ & - \sum_{j=1}^{k-1} \left(\frac{\left(M_{t_{n_j}} - \zeta_{n_{j+1}}\right)^+}{m - \zeta_{n_{j+1}}} + \mathbb{1}_{\left\{m \leq \bar{M}_{t_{n_j}}, \zeta_{n_{j+1}} \leq M_{t_{n_j}}\right\}} \frac{M_{t_{n_{j+1}}} - M_{t_{n_j}}}{m - \zeta_{n_{j+1}}} \right). \end{aligned}$$

Therefore, it is enough to prove the claim for the case

$$\zeta_1 = \xi_1 < \zeta_2 = \xi_2 < \cdots < \zeta_n = \xi_n.$$

By the same induction as in the proof of Proposition 3.2.1 it remains to prove that $(M, \bar{M})$ achieves equality in (3.2.2), (3.2.3) and (3.2.5). As for equality in (3.2.2) we note that

$$\bar{M}_{t_1} \geq m \quad \Longrightarrow \quad M_{t_1} \geq \xi_1 \qquad \text{and} \qquad \bar{M}_{t_1} < m \quad \Longrightarrow \quad M_{t_1} \leq \xi_1.$$

Equality in (3.2.3) holds by

$$\bar{M}_{t_{n-1}} \geq m,\ M_{t_{n-1}} \geq \xi_n \quad \Longrightarrow \quad M_{t_n} \geq \xi_n,$$
$$\bar{M}_{t_{n-1}} \geq m,\ M_{t_{n-1}} < \xi_n \quad \Longrightarrow \quad M_{t_n} < \xi_n$$

which one verifies using the definition of the iterated Azéma-Yor type embedding. Similarly, we now argue equality in (3.2.5). This equation corresponds to the case $\bar{M}_{t_{n-1}} < m$. By Lemma 2.2.1 we then have $\bar{M}_{t_{n-1}} \leq \zeta_{n-1}$ and therefore, together with the induction hypothesis, $\Upsilon_n^0 = 0$ (Υ_n^0 is defined in (3.2.4)). Then the conclusion holds by

$$\bar{M}_{t_{n-1}} < m,\ \bar{M}_{t_n} \geq m \quad \Longrightarrow \quad M_{t_n} \geq \xi_n,$$
$$\bar{M}_{t_{n-1}} < m,\ \bar{M}_{t_n} < m \quad \Longrightarrow \quad M_{t_n} \leq \xi_n.$$

The claim follows. □

3.3 Maximum Maximum

Having obtained Propositions 3.2.1 and 3.2.2 we now give a short proof the the fact that the iterated Azéma-Yor type embedding has the *maximum maximum* of a martingale constrained by finitely many intermediate laws.

Theorem 3.3.1 (Maximum Maximum). *Let $\mu_1, \ldots, \mu_n$ satisfy Assumption ⊛.*

Then for any stopping times $\tau_1 \leq \cdots \leq \tau_n$ such that $(B_{t\wedge\tau_n})_{t\geq 0}$ is a uniformly integrable martingale and $B_{\tau_i} \sim \mu_i$ for $i = 1, \ldots, n$ we have

$$\mathbb{P}\left[\bar{M}_n \geq m\right] \geq \mathbb{P}\left[\bar{B}_{\tau_n} \geq m\right] \qquad \forall m \geq 0 \tag{3.3.1}$$

where M is the iterated Azéma-Yor type embedding.

Proof. Under our assumptions, Lemma 2.1.11 implies that $\xi_1, \ldots, \xi_n$ from Definition 2.1.1 are non-decreasing.

Firstly, let $\xi_n(m) = \zeta_n(m) < m$. Then by Proposition 3.2.2 and taking expectations,

$$\mathbb{P}\left[\bar{M}_n \geq m\right] = \mathbb{E}\left[\Upsilon_n(M, m, \boldsymbol{\zeta})\right].$$

Denote $X_i = B_{\tau_i}$, $i = 1, \ldots, n$. Then, by the fact that $X_i \sim \mu_i$ and Proposition 3.2.1,

$$\mathbb{E}\left[\Upsilon_n(M, m, \boldsymbol{\zeta})\right] = \mathbb{E}\left[\Upsilon_n(X, m, \boldsymbol{\zeta})\right] \geq \mathbb{P}\left[\bar{X}_n \geq m\right].$$

Secondly, let $\xi_n(m) = \zeta_n(m) = m$. Assume there exists a sequence (m_k) such that $m_k \uparrow m$ and $\xi_n(m_k) = \zeta_n(m_k) < m_k$. Then, by the previous arguments and left-continuity of the tail-distribution function,

$$\mathbb{P}\left[\bar{M}_n \geq m\right] = \lim_{k\to\infty} \mathbb{P}\left[\bar{M}_n \geq m_k\right] \geq \lim_{k\to\infty} \mathbb{P}\left[\bar{X}_n \geq m_k\right] = \mathbb{P}\left[\bar{X}_n \geq m\right].$$

If no such sequence (m_k) exists, it follows (inductively) that $c_n \equiv c_j$ on $(m-\epsilon, m]$ for some $j = \jmath_n(m), \ldots, n-1$, and then Assumption $\circledast$(ii) is violated. $\square$

Building on the intuition from the general embedding in the case $n = 3$, cf. Section 2.3.2, to obtain the general *maximum maximum* it seems plausible that the choice of τ_2' becomes important. We have the following conjecture.

Conjecture 3.3.2 (General Maximum Maximum, $n = 3$)**.** *Let* τ_1, τ_2, τ_3 *solve the* 3*-marginal SEP of* μ_1, μ_2, μ_3*.*

Then we have

$$\mathbb{P}\left[\bar{M}_3 \geq m\right] \geq \mathbb{P}\left[\bar{B}_{\tau_3} \geq m\right] \qquad \forall m \geq 0$$

where M is the extended iterated Azéma-Yor type embedding from Section 2.3.2 with

$$\tau_2' := \inf\left\{u > \tau_1 : B_u \notin \left(\gamma_l(\bar{B}_u), \gamma_r(\bar{B}_u)\right)\right\}$$

where γ_l and γ_r are non-decreasing and γ_l is (pointwise) the smallest valid choice.

3.4 Financial Interpretation

We develop now a financial interpretation of the right hand side of (3.2.1) as a pathwise superhedging strategy for a simple knock-in digital barrier option with payoff $\mathbb{1}_{\left\{\bar{X}_T \geq m\right\}}$. It consists of three elements: a static position in call options, a forward transaction (with the shortest available maturity) when the barrier m is hit and rebalancing thereafter at times t_i. More precisely:

(a) Static position in calls:

$$\boldsymbol{\lambda}^{\boldsymbol{\zeta}}(X_{\boldsymbol{t}}) := \sum_{i=1}^{n}\left(\frac{(X_{t_i} - \zeta_i)^+}{m - \zeta_i} - \frac{(X_{t_i} - \zeta_{i+1})^+}{m - \zeta_{i+1}}\mathbb{1}_{\{i<n\}}\right).$$

For $1 \leq i < n$, we hold a portfolio long and short calls with maturity t_i and strikes ζ_i and ζ_{i+1} respectively. This yields a "tent like" payoff which becomes negative only if the underlying exceeds level m. Note that by setting $\zeta_i = \zeta_{i+1}$ we may avoid trading the t_i-maturity calls. For maturity t_n we are only long in a call with strike ζ_n.

(b) Forward transaction if the barrier m is hit:

$$\mathbb{1}_{\{\bar{X}_{t_{i-1}}<m\leq\bar{X}_{t_i}\}}\frac{m-X_{t_i}}{m-\zeta_i}$$

At the moment when the barrier m is hit, say between maturities t_{i^*-1} and t_{i^*}, we enter into forward contracts with maturity t_{i^*}. Note that the long call position with maturity t_{i^*} together with the forward then superhedge the knock-in digital barrier option, cf. (3.2.5). This resembles the robust semi-static hedge in the one-marginal case, cf. Brown et al. [18, Lemma 2.4]. All the "tent like" payoffs up to maturity t_{i^*-1} are non-negative.

If the trajectory jumps over the level m then a profit from the forward transaction will be realized and the superhedging property remains preserved.

(c) Rebalancing of portfolio to hedge calendar spreads:

$$-\sum_{i=1}^{n-1}\mathbb{1}_{\{m\leq\bar{X}_{t_i},\zeta_{i+1}\leq X_{t_i}\}}\frac{X_{t_{i+1}}-X_{t_i}}{m-\zeta_{i+1}}$$

After the barrier m was hit, we start trading at times t_i in such a way that a potential negative payoff of the calendar spreads $\frac{(X_{t_{i+1}}-\zeta_{i+1})^+}{m-\zeta_{i+1}}-\frac{(X_{t_i}-\zeta_{i+1})^+}{m-\zeta_{i+1}}$, $i^*\leq i<n$, is offset, cf. (3.2.3).

In the above, (b) and (c) are instances of dynamic trading which is done in a self-financing way. Their combined payoff may be written as $\int_0^T H_s^{\boldsymbol{\zeta}}(X)\mathrm{d}X_s$ for a suitable choice of (simple) integrand $H^{\boldsymbol{\zeta}}(X)$. The initial cost entering into the static position in (a) is $\boldsymbol{\mu}(\boldsymbol{\lambda}^{\boldsymbol{\zeta}})=\sum_{i=1}^n\left(\frac{c_i(\zeta_i)}{m-\zeta_i}-\frac{c_i(\zeta_{i+1})}{m-\zeta_{i+1}}\mathbb{1}_{\{i<n\}}\right)$. Then

$$\int_0^T H_s^{\boldsymbol{\zeta}}(X)\mathrm{d}X_s+\boldsymbol{\lambda}^{\boldsymbol{\zeta}}(X_{\boldsymbol{t}}) \tag{3.4.1}$$

is an example of a semi-static trading strategy and the inequality (3.2.1) now simply says that for any choice of $\zeta_1 \leq \cdots \leq \zeta_n < m$, our strategy in (3.4.1) superreplicates $\mathbb{1}_{\{\bar{X}_T \geq m\}}$.

3.5 Duality

We showed in Section 3.4 that (3.4.1) constitutes a pathwise superreplication for the Lookback option $\mathbb{1}_{\{\bar{X}_T \geq m\}}$. Obviously this strategy directly yields a pathwise superhedging strategy for piecewise constant, non-decreasing and right-continuous payoffs. In this section we prove that this strategy is optimal in the sense of (3.5.6) if Assumption ⊛ holds and that the iterated Azéma-Yor type embedding is the extremal model in the robust superhedging problem for $\phi(X_T)$ where ϕ is non-decreasing and right-continuous.

Let Ω_{X_0} denote the set of càdlàg paths ω such that $\omega_0 = X_0$. A mapping $\psi : [0,T] \times \Omega_{X_0} \to \mathbb{R}$ is called progressively measurable if for any $\omega, \tilde{\omega} \in \Omega_{X_0}$

$$\omega_s = \tilde{\omega}_s, \quad \forall s \leq [0,t] \qquad \Longrightarrow \qquad \psi(t,\omega) = \psi(t,\tilde{\omega}),$$

where we assume that $\psi(\cdot \wedge t, \cdot)$ is $B([0,t]) \otimes \sigma(X_s, s \leq t)$-measurable for every $t \leq T$. In general, we define a pathwise superreplication strategy as a triplet $(v, H, \boldsymbol{\lambda})$ where $v \in \mathbb{R}$, $\lambda_i : \mathbb{R} \to \mathbb{R}$ is measurable and H is progressively measurable and of the form

$$H_t(\omega) = \sum_{i=0}^{\infty} h_i\left(\gamma_i(\omega), \omega\right) \mathbb{1}_{\left[\gamma_i(\omega), \gamma_{i+1}(\omega) \wedge T\right)}(t) \tag{3.5.1}$$

where $0 =: \gamma_0 \leq \gamma_1 \leq \gamma_2 \leq \ldots$ is an increasing sequence of stopping times such that $\lim_{i \to \infty} \gamma_i(\omega) = \infty$ for all $\omega \in \Omega_{X_0}$ and $h_i : [0,T] \times \Omega_{X_0} \to \mathbb{R}$ is progressively measurable. For these strategies we can define $\int_0^{\cdot} H_t(\omega) \mathrm{d}\omega_t$ pathwise and then the superhedging property reads

$$\phi(\bar{\omega}_T) \leq v + \int_0^T H_t(\omega)\mathrm{d}\omega_t + \sum_{i=1}^n \lambda_i(\omega_{t_i}) \qquad \forall \omega \in \Omega_{X_0}. \tag{3.5.2}$$

Clearly, the strategy $H^{\boldsymbol{\zeta}}$ from Section 3.4 is of the form (3.5.1).

The *pathwise superhedging cost* of the payoff $\phi(\bar{X}_T)$ is defined as

$$\mathbb{A}^{\boldsymbol{\mu}}\left(\phi(\bar{X}_T)\right) := \inf \Big\{ v + \boldsymbol{\mu}(\boldsymbol{\lambda}) : (v, H, \boldsymbol{\lambda}) \text{ is a pathwise superreplication strategy for } \phi(\bar{X}_T) \Big\}. \tag{3.5.3}$$

Theorem 3.5.1 (Duality)**.** *Let $\mu_1, \ldots, \mu_n$ satisfy Assumption $\circledast$. Suppose ϕ is a right-continuous, non-decreasing function. Then*

$$\mathbb{A}^{\boldsymbol{\mu}}\left(\phi(\bar{X}_T)\right) = \sup_{\mathbb{P}} \mathbb{E}_{\mathbb{P}}\left[\phi(\bar{X}_T)\right] = \phi(X_0) + \int_{(X_0,\infty)} K_n(m)\mathrm{d}\phi(m) \tag{3.5.4}$$

where the supremum is over all martingale measures $\mathbb{P}$ such that the canonical process X on Ω_{X_0} satisfies $X_{t_i} \sim \mu_i$, and we recall the definition of K_n from (2.1.6)*.*

Moreover, recall the definition of $\boldsymbol{\xi}$ in (2.1.4) *and set*

$$\zeta_i(m) := \min_{j \geq i} \xi_j(m), \qquad i = 1, \ldots, n, \quad m > X_0. \tag{3.5.5}$$

Let $M = (M_{t_1}, \ldots, M_{t_n})$ be the iterated Azéma-Yor type embedding of $\mu_1, \ldots, \mu_n$. Then for every $\epsilon > 0$ there exist $m_k > X_0$ and $a_k > 0$, $k \in \mathbb{N}$, such that

$$\begin{aligned} &\phi\big(\bar{M}_T\big) - \phi(X_0) \\ \leq\ &\sum_{k=1}^\infty a_k \left\{ \int_0^T H_s^{\boldsymbol{\zeta}(m_k)}(M)\mathrm{d}M_s + \boldsymbol{\lambda}^{\boldsymbol{\zeta}(m_k)}(M_{\boldsymbol{t}}) \right\} \\ \leq\ &\phi\big(\bar{M}_T\big) - \phi(X_0) + \epsilon \end{aligned} \tag{3.5.6}$$

where the above inequality is understood pathwise and where we recall the definitions of $\boldsymbol{\lambda}^{\zeta}$ and H^{ζ} from Section 3.4.

Proof. First we prove the claim for the payoff $\phi(\bar{X}_T) = \mathbb{1}_{\{\bar{X}_T \geq m\}}$ for m where $\zeta_n(m) < m$. If Assumption ⊛ holds, then by Proposition 3.2.2 the iterated Azéma-Yor type embedding M achieves pathwise equality in (3.2.1) simultaneously for all $m > X_0$ satisfying $\zeta_n(m) < m$. This yields (3.5.4) and (3.5.6) with $\epsilon = 0$.

From the proof of Theorem 3.3.1 we know that for every m such that $\zeta_n(m) = m$ there exists an arbitrarily close $\tilde{m} < m$ such that $\zeta_n(\tilde{m}) < \tilde{m}$. Hence, we can find a piecewise constant, right-continuous, non-decreasing function

$$\phi^{\epsilon}(m) = \phi(X_0) + \sum_{k=1}^{\infty} a_k \mathbb{1}_{\{m \geq m_k\}}$$

with discontinuities at $\tilde{m}_k$, $\zeta_n(\tilde{m}_k) < \tilde{m}_k$ such that $0 \leq \phi^{\epsilon}(\bar{X}_T) - \phi(\bar{X}_T) < \epsilon$. By the above argument, this implies that (3.5.4) and (3.5.6) with $\epsilon = 0$ hold for the payoff $\phi^{\epsilon}(\bar{X}_T)$ as $\phi^{\epsilon}(\bar{X}_T)$ is just a positive linear combination of simple barrier options. By considering $\tilde{\phi}^{\epsilon}$ such that $0 \leq \phi(X_T) - \tilde{\phi}^{\epsilon}(X_T) < \epsilon$ this further establishes (3.5.6) for general right-continuous, non-decreasing ϕ.

Let $\phi^{\epsilon,-1}$ and ϕ^{-1} denote the right-continuous inverses of ϕ^{ϵ} and ϕ, respectively. Recall from Lemma 2.1.14 that K_n is non-increasing and continuous. Hence, applying the previous case to ϕ^{ϵ}, we have

$$\begin{aligned}
\mathbb{E}\left[\phi\left(\bar{M}_{t_n}\right)\right] &= \lim_{\epsilon \to 0} \mathbb{E}\left[\phi^{\epsilon}\left(\bar{M}_{t_n}\right)\right] = \lim_{\epsilon \to 0}\left\{\phi(X_0) + \int_{(X_0,\infty)} K_n(m)\mathrm{d}\phi^{\epsilon}(m)\right\} \\
&= \lim_{\epsilon \to 0}\left\{\phi(X_0) + \int_{\left(\phi^{\epsilon}(X_0),\phi^{\epsilon}(\infty)\right)} K_n\left(\phi^{\epsilon,-1}(x)\right)\mathrm{d}x\right\} \\
&= \phi(X_0) + \int_{(X_0,\infty)} K_n(m)\mathrm{d}\phi(m),
\end{aligned}$$

where we used monotone convergence since $K_n\left(\phi^{\epsilon,-1}(x)\right) \geq K_n\left(\phi^{-1}(x)\right)$. With this (3.5.4) follows from the Dambis-Dubins-Schwarz time change and

$$\begin{aligned}\mathbb{E}\left[\phi\left(\bar{M}_{t_n}\right)\right] &= \sup_{\mathbb{P}} \mathbb{E}_{\mathbb{P}}\left[\phi(\bar{X}_T)\right] \leq \mathbb{A}^{\boldsymbol{\mu}}\left(\phi(\bar{X}_T)\right) \leq \lim_{\epsilon\downarrow 0} \mathbb{A}^{\boldsymbol{\mu}}\left(\phi^{\epsilon}(\bar{X}_T)\right) \\ &= \lim_{\epsilon\downarrow 0}\sup_{\mathbb{P}} \mathbb{E}_{\mathbb{P}}\left[\phi^{\epsilon}(\bar{X}_T)\right] = \mathbb{E}\left[\phi\left(\bar{M}_{t_n}\right)\right].\end{aligned}$$

□

Remark 3.5.2 (Upper Bound). In the context of Theorem 3.5.1, if Assumption ⊛(ii) is not assumed, we have

$$\sup_{\mathbb{P}} \mathbb{E}_{\mathbb{P}}\left[\phi(\bar{X}_T)\right] \leq \phi(X_0) + \int_{(X_0,\infty)} K_n(m)\mathrm{d}\phi(m) \tag{3.5.7}$$

which follows from the arguments of Section 3.4.

Remark 3.5.3 (Piecewise Constant Lookback Options). It follows from the proof of Theorem 3.5.1 that for piecewise constant, non-decreasing and right-continuous payoffs ϕ (3.5.6) holds with $\epsilon = 0$.

Remark 3.5.4 (Continuous Trading). It can be shown that under certain integrability conditions and by allowing dynamic trading strategies of bounded variation, one can obtain explicitly a superreplication strategy which is optimal in the sense that it is a perfect hedging strategy for the iterated Azéma-Yor type embedding. This is formalized by Henry-Labordère et al. [49].

Remark 3.5.5 (Continuity of Extremal Model). Note that the extremal model for (3.5.4) can be chosen continuous although the duality is valid for càdlàg paths.

Remark 3.5.6 (Assumption ⊛). It follows from the discussion of Section 2.3 that if Assumption ⊛ fails then (3.5.6) is not true anymore. Also the duality statement (3.5.4) remains open in that case.

Chapter 4

Martingale Inequalities for the Maximum via Pathwise Arguments

Joint Publication Large parts of this chapter will be submitted as part of a joint publication with Jan Obłój and Nizar Touzi, see Obłój et al. [78].

In this chapter we look at certain martingale inequalities for the terminal maximum of a càdlàg submartingale. The difference to many existing martingale inequalities in the literature is that they feature information about the process at intermediate time points.

The starting point is the simple observation that Proposition 3.2.1 yields the following martingale inequality for càdlàg submartingales X (X_0 deterministic) and right-continuous, non-decreasing functions ϕ on $\mathbb{R}$,

$$\begin{aligned}
&\mathbb{E}\left[\phi(\bar{X}_{t_n})\right] \qquad (4.0.1)\\
&\leq \phi(X_0) + \int_{(X_0,\infty)} \sum_{i=1}^{n} \left(\frac{\mathbb{E}\left[(X_{t_i} - \zeta_i(m))^+\right]}{m - \zeta_i(m)} - \frac{\mathbb{E}\left[(X_{t_i} - \zeta_{i+1}(m))^+\right]}{m - \zeta_{i+1}(m)} \mathbb{1}_{\{i<n\}} \right) \mathrm{d}\phi(m)
\end{aligned}$$

for functions $\zeta_1(m) \leq \cdots \leq \zeta_n(m) < m$.

This is because for a càdlàg submartingale X, it follows by taking expectation in inequality 3.2.1 that

$$\mathbb{E}\left[\mathbb{1}_{\left\{\bar{X}_{t_n} \geq m\right\}}\right] \leq \sum_{i=1}^{n}\left(\frac{\mathbb{E}\left[(X_{t_i}-\zeta_i)^+\right]}{m-\zeta_i(m)} - \frac{\mathbb{E}\left[(X_{t_i}-\zeta_{i+1})^+\right]}{m-\zeta_{i+1}(m)}\mathbb{1}_{\{i<n\}}\right),$$

for any $\zeta_1 \leq \cdots \leq \zeta_n < m$. Then, for a right-continuous non-decreasing function ϕ, we write

$$\phi(\bar{X}_{t_n}) = \phi(X_0) + \int_{(X_0,\infty)} \mathbb{1}_{\left\{\bar{X}_{t_n} \geq m\right\}} \mathrm{d}\phi(m),$$

and we plug the previous inequality after taking expectations.

One goal of this chapter is to understand how the bound induced by these more elaborate inequalities compares to simpler inequalities which do not use information about the process at intermediate time points. We show that in our context these bounds can be both better or worse.

Throughout, we emphasize the simplicity of our arguments, which are all elementary. This is illustrated in Section 4.2 where we obtain amongst others the sharp versions of Doob's inequalities.

The idea of deriving martingale inequalities from pathwise inequalities is already present in the work on robust pricing and hedging by Hobson [50]. Other authors have used pathwise arguments to derive martingale inequalities, e.g. Doob's inequalities are considered by Acciaio et al. [2] and Obłój and Yor [77]. The Burkholder-Davis-Gundy inequality is rediscovered with pathwise arguments by Beiglböck and Siorpaes [9]. In this context we also refer to Cox and Wang [24] and Cox and Peskir [23] whose pathwise inequalities relate a process and time. In a similar spirit, bounds for local time are obtained by Cox et al. [25]. Beiglböck and Nutz [8] look at general martingale inequalities and explain how they can be obtained from deterministic inequalities.

In a a discrete time and quasi-sure setup, the results of Nutz and Bouchard [71] can be seen as general theoretical underpinning of many ideas we present here in the special case of martingale inequalities for the terminal maximum.

Organization of the Chapter In the next section we state and prove our main results. In Section 4.2 we specialize our inequalities and demonstrate how they can be used to derive, amongst others, Doob's inequalities. We also investigate in which sense our martingale inequalities can provide sharper versions of Doob's inequalities.

4.1 Main Result

In our main results, we obtain and compare inequalities for càdlàg submartingales which are directly implied by Proposition 3.2.1.

4.1.1 Main Result – Part 1

In the first part of our main result we devise a general martingale inequality for $\mathbb{E}\left[\phi(\bar{X}_T)\right]$ and prove that it is attained under some conditions.

Define

$$\begin{aligned}\mathscr{Z} := \Big\{ \boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_n) \ : \zeta_i : (X_0, \infty) \to \mathbb{R} \text{ is right-continuous}, \\ \zeta_1(m) \le \cdots \le \zeta_n(m) < m, \quad n \in \mathbb{N} \Big\}.\end{aligned} \tag{4.1.1}$$

In order to ensure that the expectations we consider are finite we will occasionally need the technical condition that there exists $\alpha \in (0,1)$ such that

$$\liminf_{m\to\infty} \frac{\zeta_1(m)}{\alpha m} \geq 1, \qquad \limsup_{m\to\infty} \frac{\phi(m)}{m^\gamma} = 0 \quad \text{for some } \gamma < \frac{1}{1-\alpha}. \tag{4.1.2}$$

Theorem 4.1.1 (Main Result – Part 1)**.** *Let X be a càdlàg submartingale, ϕ a right-continuous, non-decreasing function and $t_1 \leq \cdots \leq t_n = T$. Then, for all $\boldsymbol{\zeta} = (\zeta_1, \ldots, \zeta_n) \in \mathscr{Z}$ we have*

$$\mathbb{E}\left[\phi(\bar{X}_T)\right] \leq \mathrm{UB}\left(X, \phi, \boldsymbol{\zeta}\right) := \phi(X_0) + \int_{(X_0,\infty)} \sum_{i=1}^{n} \mathbb{E}\left[\lambda_i^{\boldsymbol{\zeta},m}(X_{t_i})\right] \mathrm{d}\phi(m) \tag{4.1.3}$$

where

$$\lambda_i^{\boldsymbol{\zeta},m}(x) := \frac{(x-\zeta_i(m))^+}{m-\zeta_i(m)} - \frac{(x-\zeta_{i+1}(m))^+}{m-\zeta_{i+1}(m)} \mathbb{1}_{\{i<n\}}. \tag{4.1.4}$$

Now fix $\boldsymbol{\zeta} \in \mathscr{Z}$ and consider (4.1.3) *as being valid for every càdlàg submartingale. If ζ_1 is non-decreasing and satisfies together with ϕ the condition* (4.1.2)*, then there exists a continuous martingale which achieves equality in* (4.1.3)*.*

Remark 4.1.2 (Optimization over $\boldsymbol{\zeta}$)*.* If X and $t_1, \ldots, t_n$ are fixed we can optimize (4.1.3) over $\boldsymbol{\zeta} \in \mathscr{Z}$ to obtain a minimizer $\boldsymbol{\zeta}^\star$. Clearly, more intermediate points t_i in (4.1.3) can only improve the bound for this particular process X. However, only for very special processes (e.g. the iterated Azéma-Yor type embedding) is there hope that (4.1.3) will hold with equality. This is, loosely speaking, because a finite number of intermediate marginal law constraints will in general not uniquely determine the law of the maximum at terminal time t_n.

Proof of Theorem 4.1.1. Equation (4.1.3) has been proven in the introduction of this chapter.

If ζ_1 is non-decreasing and $\zeta_1(m) \geq \alpha m$ for m large, $\alpha \in (0,1)$, we define X by

$$X_t = \begin{cases} B_{\frac{t}{t_1 - t} \wedge \tau_{\zeta_1}} & \text{if } t < t_1, \\ B_{\tau_{\zeta_1}} & \text{if } t \geq t_1. \end{cases}$$

where B is a Brownian motion, $B_0 = X_0$, and $\tau_{\zeta_1} := \inf\left\{u > 0 \ : \ B_u \leq \zeta_1(\bar{B}_u)\right\}$. X is a uniformly integrable martingale by similar arguments as in the proof of Proposition 2.2.5. Then, one readily verifies together with Proposition 3.2.2 that

$$\Upsilon_n(X, m, \boldsymbol{\zeta}) = \Upsilon_1(X, m, \boldsymbol{\zeta}) = \mathbb{1}_{\left\{\bar{X}_{t_1} \geq m\right\}} = \mathbb{1}_{\left\{\bar{X}_{t_n} \geq m\right\}}.$$

Condition (4.1.2) ensures that $\mathbb{E}\left[\phi(\bar{X}_{t_n})\right] < \infty$ because then by excursion theoretical results, cf. e.g. Rogers [83], we compute

$$\begin{aligned} \mathbb{P}\left[\bar{X}_{t_n} \geq y\right] &= \exp\left(-\int_{(X_0, y]} \frac{1}{z - \zeta_1(z)} \mathrm{d}z\right) \leq \text{const} \cdot \exp\left(-\int_{(1, y]} \frac{1}{z - \alpha z} \mathrm{d}z\right) \\ &= \text{const} \cdot y^{-\frac{1}{1-\alpha}} \end{aligned}$$

for large y. Set $\phi^m(y) := \mathbb{1}_{\{y \geq m\}}$. Now the claim follows from

$$\begin{aligned} \mathbb{E}\left[\phi(\bar{X}_{t_n})\right] &= \phi(X_0) + \int_{(X_0, \infty)} \mathbb{E}\left[\mathbb{1}_{\left\{\bar{X}_{t_n} \geq m\right\}}\right] \mathrm{d}\phi(m) \\ &= \phi(X_0) + \int_{(X_0, \infty)} \mathrm{UB}\left(X, \phi^m, \boldsymbol{\zeta}\right) \mathrm{d}\phi(m) \\ &= \phi(X_0) + \mathrm{UB}\left(X, \phi, \boldsymbol{\zeta}\right) \end{aligned}$$

where we applied Fubini's theorem. □

4.1.2 Main Result – Part 2

As mentioned in the introduction, the speciality about our martingale inequality from Theorem 4.1.1 is that it uses information about the process at intermediate times. The second part of our main result sheds light on the question whether this information gives more accurate bounds than e.g. in the case when no information about the process at intermediate times is used. In short, the answer is negative, i.e. we demonstrate that for a large class of $\tilde{\boldsymbol{\zeta}}$'s there is no "universally better" choice of $\boldsymbol{\zeta}$ in the sense that it yields a tighter bound in the class of inequalities for $\mathbb{E}\left[\phi(\bar{X}_T)\right]$ from Theorem 4.1.1.

To avoid overly technicalities in the second part of our main result, we impose additional conditions on $\boldsymbol{\zeta} \in \mathscr{Z}$ and ϕ below. Many of these conditions could be relaxed to obtain a stronger statement in Theorem 4.1.3. We define

$$\mathscr{Z}^{\text{cts}} := \left\{ \boldsymbol{\zeta} \in \mathscr{Z} \ : \ \boldsymbol{\zeta} \text{ is continuous} \right\} \tag{4.1.5}$$

and

$$\begin{aligned} \tilde{\mathscr{Z}} := \Big\{ \boldsymbol{\zeta} \in \mathscr{Z} \ : \ & \zeta_i \text{ is strictly increasing and continuous,} \\ & \exists \alpha > 0 \text{ s.t. } \liminf_{m \to \infty} \zeta_1(m)/\alpha m \geq 1, \\ & \exists \epsilon > 0 \ \text{ s.t. } \ \zeta_1(m) = \cdots = \zeta_n(m) \ \ \forall m \in (X_0, X_0 + \epsilon) \Big\}. \end{aligned} \tag{4.1.6}$$

Before we proceed, we want to argue that the set $\tilde{\mathscr{Z}}$ arises quite naturally. In the setting of Remark 4.1.2, if X is a martingale such that its marginal laws

$$\mu_1 := \mathcal{L}\left(X_{t_1}\right), \quad \ldots, \quad \mu_n := \mathcal{L}\left(X_{t_n}\right)$$

satisfy Assumption ⊛, $\int (x - \zeta)^+ \mu_i(\mathrm{d}x) < \int (x - \zeta)^+ \mu_{i+1}(\mathrm{d}x)$ for all ζ in the interior of the support of μ_{i+1} and their barycenter functions satisfy the mean residual value property of Madan and Yor [64] close to X_0 and have no atoms at the left end of support, then the optimization over $\boldsymbol{\zeta}$ as described in Remark 4.1.2 yields a unique $\tilde{\boldsymbol{\zeta}}^\star \in \tilde{\mathscr{Z}}$. Hence, the set of these $\tilde{\boldsymbol{\zeta}}^\star$ seems a "good candidate set" for $\boldsymbol{\zeta}$'s to be used in Theorem 4.1.1.

The statement of the Theorem 4.1.3 concerns the negative orthant of $\mathscr{Z}^{\text{cts}}$,

$$\mathscr{Z}^{\text{cts}}_{-}(\phi,\tilde{\boldsymbol{\zeta}}) := \Big\{\boldsymbol{\zeta} \in \mathscr{Z}^{\text{cts}} \ : \text{UB}\left(X,\phi,\boldsymbol{\zeta}\right) \leq \text{UB}\left(X,\phi,\tilde{\boldsymbol{\zeta}}\right) \ \ \forall \text{ càdlàg submartingales } X \text{ and } < \text{for one } X\Big\}. \tag{4.1.7}$$

Theorem 4.1.3 (Main Result – Part 2)**.** *Let ϕ be a right-continuous, strictly increasing function. Suppose $\tilde{\boldsymbol{\zeta}} \in \tilde{\mathscr{Z}}$ and assume* (4.1.2) *holds. Then,*

$$\mathscr{Z}^{\text{cts}}_{-}(\phi,\tilde{\boldsymbol{\zeta}}) = \emptyset. \tag{4.1.8}$$

The key ingredient for the proof of Theorem 4.1.3 is isolated in the following Proposition.

Proposition 4.1.4 (Positive Error)**.** *Let $\tilde{\boldsymbol{\zeta}} \in \tilde{\mathscr{Z}}$ and $\boldsymbol{\zeta} \in \mathscr{Z}^{\text{cts}}$ satisfy $\tilde{\boldsymbol{\zeta}} \neq \boldsymbol{\zeta}$. Set $\phi^m(y) := \mathbb{1}_{\{y\geq m\}}$. Then there exists a non-empty interval $(m_1,m_2) \subseteq (X_0,\infty)$ such that*

$$\forall m \in (m_1,m_2): \quad \text{UB}\left(X,\phi^m,\tilde{\boldsymbol{\zeta}}\right) < \text{UB}\left(X,\phi^m,\boldsymbol{\zeta}\right) \tag{4.1.9}$$

where X is an iterated Azéma-Yor type embedding based on some $\tilde{\boldsymbol{\xi}}$.

Proof. To each $\tilde{\boldsymbol{\zeta}} \in \tilde{\mathscr{Z}}$ we can associate a non-decreasing and continuous stopping boundary $\tilde{\boldsymbol{\xi}}$ which satisfies

$$\tilde{\xi}_n(m) < \cdots < \tilde{\xi}_1(m) < m \qquad \forall m \in (X_0, X_0+\epsilon), \tag{4.1.10a}$$

$$\tilde{\boldsymbol{\xi}}(m) = \tilde{\boldsymbol{\zeta}}(m) \qquad \forall m \geq X_0+\epsilon, \tag{4.1.10b}$$

for some $\epsilon > 0$, and hence

$$\tilde{\zeta}_i(m) = \min_{j\geq i} \tilde{\xi}_j(m) \qquad \forall m > X_0. \tag{4.1.11}$$

Fix such a $\tilde{\boldsymbol{\xi}}$ and let X be an iterated Azéma-Yor type embedding based on this $\tilde{\boldsymbol{\xi}}$. Let $j \geq 1$. Using the notation of Definition 2.1.2, it follows by monotonicity of $\tilde{\boldsymbol{\xi}}$, (4.1.10b) and (4.1.11) that on the set $\{B_{\tau_j} = \tilde{\xi}_j(\bar{B}_{\tau_j}),\ \bar{B}_{\tau_j} \geq X_0 + \epsilon\}$ we have $B_{\tau_j} = \tilde{\xi}_j(\bar{B}_{\tau_j}) \leq \tilde{\xi}_{j+1}(\bar{B}_{\tau_j})$. Therefore, the condition of (2.1.7a) in the definition of the iterated Azéma-Yor type embedding is not satisfied and hence $\tau_{j+1} = \tau_j$. Consequently,

$$\begin{aligned} &X_{t_j} = X_{t_{j+1}} = \cdots = X_{t_n} \quad \text{and} \quad \bar{X}_{t_j} = \bar{X}_{t_{j+1}} = \cdots = \bar{X}_{t_n} \\ &\text{on the set } \left\{X_{t_j} = \tilde{\xi}_j(\bar{X}_{t_j}),\ \bar{X}_{t_j} \geq X_0 + \epsilon\right\} \end{aligned} \tag{4.1.12}$$

for all $j \geq 1$.

Take $1 \leq j \leq n$. Denote $\chi := \max\{k \leq n : \exists t \leq H_{X_0+\epsilon} \text{ s.t. } B_t \leq \tilde{\xi}_k(\bar{B}_t)\} \vee 0$, where $H_x := \inf\{u > 0 : B_u = x\}$ and $\mathcal{H} := \{\chi = j-1,\ H_{X_0+\epsilon} < \infty\}$. By (4.1.10a) we have $\mathbb{P}[\mathcal{H}] > 0$. Further, by using $\tilde{\zeta}_1(m) \leq \cdots \leq \tilde{\zeta}_n(m) < m$ we conclude by the properties of Brownian motion that $\mathbb{P}\left[\mathcal{H} \cap \{\bar{B}_{\tau_j} \in \mathcal{O}\}\right] > 0$ for $\mathcal{O} \subseteq (X_0 + \epsilon, \infty)$ an open set. Relabelling and using (4.1.10b) yields

$$\mathbb{P}\left[X_{t_j} = \tilde{\zeta}_j(\bar{X}_{t_j}), \bar{X}_{t_j} \in \mathcal{O}, \bar{X}_{t_{j-1}} < X_0 + \epsilon\right] > 0 \tag{4.1.13}$$

for all open $\mathcal{O} \subseteq (X_0 + \epsilon, \infty)$.

By $\tilde{\boldsymbol{\zeta}} \neq \boldsymbol{\zeta}$ either Case A or Case B below holds (possibly by changing ϵ above). In our arguments we refer to the proof of the pathwise inequality of Proposition 3.2.1 and argue that certain inequalities in this proof become strict.

Case A: $\exists \mathcal{O} = (m_1, m_2) \subseteq (X_0 + \epsilon, \infty)$ and $j \leq n$ such that $\tilde{\zeta}_j(m_1) > \zeta_j(m_2)$. Take $m > m_2$. Then on $\left\{X_{t_j} = \tilde{\zeta}_j(\bar{X}_{t_j}),\ \bar{X}_{t_j} \in \mathcal{O}\right\}$ we have almost surely

$$\Upsilon_n(X, m, \boldsymbol{\zeta}) \overset{(4.1.12)}{=} \Upsilon_j(X, m, \boldsymbol{\zeta}) > 0 = \mathbb{1}_{\{m \leq \bar{X}_{t_j}\}} \overset{(4.1.12)}{=} \mathbb{1}_{\{m \leq \bar{X}_{t_n}\}} \overset{\text{Prop. 3.2.2}}{=} \Upsilon_n(X, m, \tilde{\boldsymbol{\zeta}})$$

where the strict inequality holds by noting that $(X_{t_j} - \zeta_j(m))^+ > 0$ for all $m \in (m_1, m_2)$ on the above set and then directly verifying that the second inequality of (3.2.5) applied with $\boldsymbol{\zeta}$ and X is strict.

Case B: $\exists \mathcal{O} = (m_1, m_2) \subseteq (X_0 + \epsilon, \infty)$ and $j \leq n$ such that $\tilde{\zeta}_j(m_2) < \zeta_j(m_1)$. Take $m \in \mathcal{O}$. Then on $\left\{X_{t_j} = \tilde{\zeta}_j(\bar{X}_{t_j}),\ \bar{X}_{t_j} \in \mathcal{O} \cap (m, \infty),\ \bar{X}_{t_{j-1}} < X_0 + \epsilon\right\}$ we have almost surely

$$\Upsilon_n(X, m, \boldsymbol{\zeta}) \overset{(4.1.12)}{=} \Upsilon_j(X, m, \boldsymbol{\zeta}) > 1 = \mathbb{1}_{\{m \leq \bar{X}_{t_j}\}} = \mathbb{1}_{\{m \leq \bar{X}_{t_n}\}} \overset{\text{Prop. 3.2.2}}{=} \Upsilon_n(X, m, \tilde{\boldsymbol{\zeta}})$$

where the strict inequality holds by observing that the last inequality in (3.2.5) applied with $\boldsymbol{\zeta}$ and X is strict because $(X_j - \zeta_j(m))^+ = 0 > X_j - \zeta_j(m)$ for all $m \in \mathcal{O}$ on the above set.

Combining, in both cases A and B the claim (4.1.9) follows from (4.1.13). □

Proof of Theorem 4.1.3. Take $\boldsymbol{\zeta} \in \mathscr{X}^{\text{cts}}$ such that strict inequality holds for one submartingale in the definition of $\mathscr{X}_{-}^{\text{cts}}$, see (4.1.7). Then we must have $\boldsymbol{\zeta} \neq \tilde{\boldsymbol{\zeta}}$.

As in the proof of Proposition 4.1.4 we choose a $\tilde{\boldsymbol{\xi}}$ such that (4.1.10a)–(4.1.10b), (4.1.11) hold and let X be an iterated Azéma-Yor type embedding based on this $\tilde{\boldsymbol{\xi}}$. Propositions 3.2.1 and 3.2.2 yield

$$\mathbb{E}\left[\phi^m(\bar{X}_{t_n})\right] = \text{UB}\left(X, \phi^m, \tilde{\boldsymbol{\xi}}\right) \leq \text{UB}\left(X, \phi^m, \boldsymbol{\zeta}\right) \qquad \forall m > X_0$$

and by Proposition 4.1.4

$$\text{UB}\left(X, \phi^m, \tilde{\boldsymbol{\xi}}\right) < \text{UB}\left(X, \phi^m, \boldsymbol{\zeta}\right)$$

for all $m \in \mathcal{O}$ where $\mathcal{O} \subseteq (X_0, \infty)$ is some open set. Now the claim follows as in the proof of Theorem 4.1.1. □

Remark 4.1.5. In the setting of Theorem 4.1.3 let $\tilde{\boldsymbol{\zeta}}^1, \tilde{\boldsymbol{\zeta}}^2 \in \tilde{\mathscr{Z}}$, $\tilde{\boldsymbol{\zeta}}^1 \neq \tilde{\boldsymbol{\zeta}}^2$, and assume that (4.1.2) holds for $(\phi, \tilde{\boldsymbol{\zeta}}^1)$ and $(\phi, \tilde{\boldsymbol{\zeta}}^2)$. Then there exist martingales X^1 and X^2 such that

$$\mathrm{UB}\left(X^1, \phi, \tilde{\boldsymbol{\zeta}}^1\right) < \mathrm{UB}\left(X^1, \phi, \tilde{\boldsymbol{\zeta}}^2\right),$$
$$\mathrm{UB}\left(X^2, \phi, \tilde{\boldsymbol{\zeta}}^1\right) > \mathrm{UB}\left(X^2, \phi, \tilde{\boldsymbol{\zeta}}^2\right).$$

This follows by essentially reversing the roles of $\tilde{\boldsymbol{\zeta}}^1$ and $\tilde{\boldsymbol{\zeta}}^2$ in the proof of Theorem 4.1.3.

4.2 Doob's Inequalities

In this section we demonstrate how Theorem 4.1.1 can be used to derive Doob's inequalities. Further, we investigate in which sense there is an improvement to Doob's inequalities.

Related work on pathwise interpretations of Doob's inequalities can be found in Acciaio et al. [2] and Obłój and Yor [77]. Peskir [80, Section 4] derives Doob's inequalities and shows that the constants he obtains are optimal. We now give an alternative proof of these statements and we provide new sharp inequalities for the case $0 \leq p < 1$.

4.2.1 Doob's L^p-Inequalities, $p > 1$

Using a special case of Theorem 4.1.1 we obtain an improvement to Doob's inequalities.

Denote $\mathrm{pow}^p(m) = m^p$, $\zeta^\alpha(m) := \alpha m$.

Proposition 4.2.1 (Doob's L^p-Inequalities, $p > 1$)**.** *Let* $(X_t)_{t \leq T}$ *be a non-negative càdlàg submartingale (here* T *is assumed to be a deterministic constant). Then:*

(i) *for* $p > 1$,

$$\begin{aligned} \mathbb{E}\left[\bar{X}_T^p\right] &\leq \mathrm{UB}\left(X, \mathrm{pow}^p, \zeta^{\frac{p-1}{p}}\right) && (4.2.1a)\\ &\leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[X_T^p\right] - \frac{p}{p-1} X_0^p. && (4.2.1b)\end{aligned}$$

(ii) *for every* $\epsilon > 0$ *there exists a martingale* X *such that*

$$0 \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[X_T^p\right] - \frac{p}{p-1} X_0^p - \mathbb{E}\left[\bar{X}_T^p\right] < \epsilon. \tag{4.2.2}$$

(iii) *The inequality in* (4.2.1b) *is strict if and only if either:*

$$\mathbb{E}\left[\bar{X}_T^p\right] < \infty \text{ and } X_T < \frac{p-1}{p} X_0 \text{ with positive probability, or} \tag{4.2.3a}$$

$$\mathbb{E}\left[\bar{X}_T^p\right] < \infty \text{ and } X \text{ is a strict submartingale.} \tag{4.2.3b}$$

Proof. Let us first prove (4.2.1a) and (4.2.1b). If $\mathbb{E}\left[X_T^p\right] = \infty$ there is nothing to show. In the other case, equation (4.2.1a) follows from Theorem 4.1.1 applied with $n = 1$, $\phi(y) = \mathrm{pow}^p(y) = y^p$ and $\zeta_1 = \zeta^\alpha$, $\alpha < 1$. We compute

$$\begin{aligned} \mathbb{E}\left[\bar{X}_T^p\right] - X_0^p \;&\leq\; \mathrm{UB}\left(X, \mathrm{pow}^p, \zeta^\alpha\right) - X_0^p \;=\; \mathbb{E}\left[\int_{X_0}^{\infty} p y^{p-1} \frac{(X_T - \alpha y)^+}{y - \alpha y} \mathrm{d}y\right] \\ &= \mathbb{E}\left[\int_{X_0}^{\frac{X_T}{\alpha} \vee X_0} p y^{p-1} \frac{X_T - \alpha y}{y - \alpha y} \mathrm{d}y\right] \;\leq\; \mathbb{E}\left[\int_{X_0}^{\frac{X_T}{\alpha}} p y^{p-1} \frac{X_T - \alpha y}{y - \alpha y} \mathrm{d}y\right] \\ &= \frac{p}{p-1} \frac{1}{1-\alpha} \mathbb{E}\left[\left\{\left(\frac{X_T}{\alpha}\right)^{p-1} - X_0^{p-1}\right\} X_T\right] - \frac{\alpha}{1-\alpha} \mathbb{E}\left[\left(\frac{X_T}{\alpha}\right)^p - X_0^p\right] \\ &\leq \frac{1}{p-1} \frac{1}{(1-\alpha)\alpha^{p-1}} \mathbb{E}\left[X_T^p\right] - \frac{p - \alpha(p-1)}{(p-1)(1-\alpha)} X_0^p \end{aligned} \tag{4.2.4}$$

where we used at the last inequality that X is a submartingale.

Note that the function $\alpha \mapsto \frac{1}{(1-\alpha)\alpha^{p-1}}$ attains its minimum at $\alpha^\star = \frac{p-1}{p}$. Plugging $\alpha = \alpha^\star$ into the above yields (4.2.1b).

Now we prove that Doob's L^p-inequality is attained asymptotically in the sense of (4.2.2). This fact was also proven by Peskir [80, Section 4].

Let $X_0 > 0$, otherwise the claim is trivial. Set $\alpha^\star = \frac{p-1}{p}$ and take $\alpha^\star < \alpha := \frac{p+\epsilon-1}{p+\epsilon} < 1$. Let $X_T = B_{\tau_\alpha}$ where B is a Brownian motion stared at X_0 and $\tau_\alpha := \inf\{u > 0 : B_u \le \alpha \bar{B}_u\}$. Then by using exursion theoretical results, cf. e.g. Rogers [83],

$$\bar{F}(y) := \mathbb{P}\left[\bar{X}_T \ge y\right] = \exp\left(-\int_{X_0}^{y} \frac{1}{z-\alpha z}\mathrm{d}z\right) = \left(\frac{y}{X_0}\right)^{-\frac{1}{1-\alpha}}$$

and then direct computation shows

$$\mathbb{E}\left[\bar{X}_T^p\right] = \frac{p+\epsilon}{\epsilon}X_0^p.$$

By Doob's L^p-inequality,

$$\mathbb{E}\left[\bar{X}_T^p\right] \le \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[X_T^p\right] - \frac{p}{p-1}X_0^p = \left(\frac{\alpha}{\alpha^\star}\right)^p \mathbb{E}\left[\bar{X}_T^p\right] - \frac{p}{p-1}X_0^p$$

and one verifies

$$\left\{\left(\frac{p}{p-1}\right)^p \cdot \left[\frac{p+\epsilon-1}{p+\epsilon}\right]^p - 1\right\} \cdot \frac{p+\epsilon}{\epsilon}X_0^p \quad \xrightarrow[\epsilon\downarrow 0]{} \quad \frac{p}{p-1}X_0^p.$$

This establishes the claim in (4.2.2).

Finally, we note that in the calculations (4.2.4) which led to (4.2.1b) there are three inequalities: the first one comes from Theorem 4.1.1 and does not concern the claim regarding (4.2.3a)–(4.2.3b). The second one is clearly strict if and only if (4.2.3a) holds. The third one is clearly strict if and only if (4.2.3b) holds. □

Remark 4.2.2 (Asymptotic Attainability). For the martingales in (ii) of Proposition 4.2.1 we have

$$\mathrm{UB}\left(X, \mathrm{pow}^p, \zeta^{\frac{p-1}{p}}\right) = \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[X_T^p\right] - \frac{p}{p-1} X_0^p$$

and $\mathbb{E}\left[X_T^p\right] \to \infty$ as $\epsilon \to 0$.

4.2.2 Doob's L^1-Inequality

Using a special case of Theorem 4.1.1 we obtain an improvement to Doob's inequalities.

Denote $\mathrm{id}(m) = m$, and

$$\underline{\zeta}^{\alpha}(m) := \begin{cases} -\infty & \text{if } m < 1, \\ \alpha m & \text{if } m \geq 1. \end{cases} \tag{4.2.5}$$

Proposition 4.2.3 (Doob's L^1-Inequality). *Let $(X_t)_{t \leq T}$ be a non-negative càdlàg submartingale. Then:*

(i) *with* $0 \log(0) := 0$ *and* $V(x) := x - x\log(x)$,

$$\mathbb{E}\left[\bar{X}_T\right] \leq \mathrm{UB}\left(X, \mathrm{id}, \underline{\zeta}^{\frac{1}{e}}\right) \tag{4.2.6a}$$

$$\leq \frac{e}{e-1}\Big(\mathbb{E}\left[X_T \log\left(X_T\right)\right] + V(1 \vee X_0)\Big). \tag{4.2.6b}$$

(ii) *in the case $X_0 \geq 1$ there exists a martingale which achieves equality in both* (4.2.6a) *and* (4.2.6b) *and in the case $X_0 < 1$ there exists a submartingale which achieves equality in both* (4.2.6a) *and* (4.2.6b).

(iii) *the inequality in* (4.2.6b) *is strict if and only if either:*

$$\mathbb{E}\left[\bar{X}_T\right] < \infty \text{ and } \bar{X}_T \geq 1, \quad X_T < \frac{1}{e}X_0 \text{ with positive probability,} \tag{4.2.7a}$$

$$\mathbb{E}\left[\bar{X}_T\right] < \infty \text{ and } \bar{X}_T \geq 1, \quad \mathbb{E}\left[X_T\right] > X_0 \vee 1, \quad \text{or} \tag{4.2.7b}$$

$$\mathbb{E}\left[\bar{X}_T\right] < \infty \text{ and } \bar{X}_T < 1 \text{ with positive probability.} \tag{4.2.7c}$$

Proof. Let us first prove (4.2.6a) and (4.2.6b). If $\mathbb{E}\left[\bar{X}_T\right] = \infty$ there is nothing to show. In the other case, equation (4.2.6a) follows from Theorem 4.1.1 applied with $n = 1$, $\phi(y) = \mathrm{id}(y) = y$ and $\zeta_1 = \underline{\zeta}^{\frac{1}{e}}$.

In the case $X_0 \geq 1$ we further compute using $\zeta_1 = \underline{\zeta}^{\alpha}$, $\alpha < 1$,

$$\begin{aligned}
\mathbb{E}\left[\bar{X}_T\right] - X_0 \quad &\leq \quad \mathrm{UB}\left(X, \mathrm{id}, \underline{\zeta}^{\alpha}\right) - X_0 \\
= \quad & \int_{X_0}^{\frac{X_T}{\alpha} \vee X_0} \frac{\mathbb{E}\left[X_T - \alpha y\right]}{y - \alpha y} \mathrm{d}y \quad \leq \quad \mathbb{E}\left[\int_{X_0}^{\frac{X_T}{\alpha}} \frac{X_T - \alpha y}{(1-\alpha) y} \mathrm{d}y\right] \\
= \quad & \frac{\alpha}{1-\alpha}\mathbb{E}\left[\frac{X_T}{\alpha}\left\{\log\left(\frac{X_T}{\alpha}\right) - \log(X_0)\right\}\right] - \frac{\alpha}{1-\alpha}\mathbb{E}\left[\frac{X_T}{\alpha} - X_0\right] \\
\overset{\alpha = 1/e}{=} \quad & \frac{e}{e-1}\mathbb{E}\left[X_T \log\left(X_T\right)\right] - \frac{e}{e-1}\mathbb{E}\left[X_T\right]\log(X_0) + \frac{1}{e-1}X_0 \\
\leq \quad & \frac{e}{e-1}\mathbb{E}\left[(X_T) \log\left(X_T\right)\right] - \frac{e}{e-1}X_0\log(X_0) + \frac{1}{e-1}X_0.
\end{aligned} \tag{4.2.8}$$

where the choice $\alpha = \frac{1}{e}$ gives a convenient cancellation and we used again that X is a submartingale. This is (4.2.6b) in the case $X_0 \geq 1$.

For the case $0 < X_0 < 1$ we obtain from Proposition 3.2.1 for $n = 1$,

$$\mathbb{P}\left[\bar{X}_T \geq y\right] \;\leq\; \inf_{\zeta < y} \frac{\mathbb{E}\left[(X_T - \zeta)^+\right]}{y - \zeta} \leq \frac{\mathbb{E}\left[(X_T - \alpha y)^+\right]}{y - \alpha y}$$

for $\alpha < 1$ and therefore

$$\begin{aligned}
\mathbb{E}\left[\bar{X}_T\right] - X_0 \;&=\; \int_{X_0}^{\infty} \mathbb{P}\left[\bar{X}_T \geq y\right] \mathrm{d}y \;\leq\; (1 - X_0) + \int_1^{\infty} \mathbb{P}\left[\bar{X}_T \geq y\right] \mathrm{d}y \\
&\leq (1 - X_0) + \frac{e}{e-1}\mathbb{E}\left[(X_T)\log\left(X_T\right)\right] + \frac{1}{e-1}
\end{aligned} \tag{4.2.9}$$

by (4.2.8). This is (4.2.6b) in the case $X_0 < 1$.

Now we prove that Doob's L^1-inequality is attained. This was also proven by Peskir [80, Section 4].

Firstly, let $X_0 \geq 1$. Then the martingale

$$X = \left(B_{\frac{t}{T-t} \wedge \tau_{\frac{1}{e}}} \right)_{t \leq T} \tag{4.2.10}$$

where B is a Brownian motion, $B_0 = X_0$, achieves equality in both (4.2.6a) and (4.2.6b). Secondly, let $X_0 < 1$. Then the submartingale X defined by

$$\begin{cases} X_0 & \text{if } t < T/2, \\ B_{\frac{t-T/2}{T/2-(t-T/2)} \wedge \tau_{\frac{1}{e}}} & \text{if } t \geq T/2, \end{cases} \tag{4.2.11}$$

where B is a Brownian motion, $B_0 = 1$, achieves equality in both (4.2.6a) and (4.2.6b).

Finally, we note that in the calculations (4.2.8) which led to (4.2.1b) there are three inequalities: the first one comes from Theorem 4.1.1 and does not concern the claim regarding (4.2.7a)–(4.2.7c). The second one is clearly strict if and only if (4.2.7a) holds. The third one is clearly strict if and only if (4.2.7b) holds. In addition, in the case $X_0 < 1$ there is an additional error coming from (4.2.9). Note that

$$\left. \frac{\mathbb{E}\left[(X_T - \zeta)^+\right]}{y - \zeta} \right|_{\zeta = \infty} := \lim_{\zeta \to -\infty} \frac{\mathbb{E}\left[(X_T - \zeta)^+\right]}{y - \zeta} = 1$$

in the case when $\mathbb{E}\left[\bar{X}_T\right] < \infty$. Hence, the first inequality in (4.2.9) is strict if and only if (4.2.7c) holds. The second inequality in (4.2.9) is strict if and only if (4.2.7a) or (4.2.7b) holds. $\square$

4.2.3 Doob Type Inequalities, $0 \le p < 1$

It is well known that if X is a positive continuous supermartingale converging a.s. to zero, then

$$\bar{X}_\infty \sim \frac{X_0}{U} \tag{4.2.12}$$

where U is a uniform random variable on $[0,1]$. Further, if X does not converge to zero but a non-negative limit X_∞, we can define a process Y by means of time-changing X and extension with a positive continuous supermartingale converging a.s. to zero. Clearly, Y is then a positive continuous supermartingale converging a.s. to zero and $\bar{X}_\infty \le \bar{Y}_\infty$. Hence, for any positive continuous supermartingale X, and for $p \in [0,1)$,

$$\mathbb{E}\left[\bar{X}_T^p\right] \le \mathbb{E}\left[\left(\frac{X_0}{U}\right)^p\right] = \int_0^1 \left(\frac{X_0}{u}\right)^p \mathrm{d}u = \frac{X_0^p}{1-p} \tag{4.2.13}$$

and (4.2.13) is attained.

We now generalize (4.2.13) to a non-negative submartingale.

Proposition 4.2.4 (Doob Type Inequalities, $0 \le p < 1$)**.** *Let X be a positive càdlàg submartingale, $0 \le p < 1$, denote $m_r := X_0^{-r}\mathbb{E}\left[X_T^r\right]$ for $r \le 1$. Then:*

(i) *there is a unique $\hat{\alpha} \in (0,1]$ which solves*

$$m_p \hat{\alpha}^{-p} = \frac{1-p+pm_1}{1-p+p\hat{\alpha}} \tag{4.2.14}$$

and for which we have

$$\mathbb{E}\left[\bar{X}_T^p\right] \le X_0^p m_p \hat{\alpha}^{-p} = \frac{X_0^p}{1-p+p\hat{\alpha}} + X_0^{p-1}\frac{p}{1-p+p\hat{\alpha}}\Big(\mathbb{E}\left[X_T\right] - X_0\Big) \tag{4.2.15a}$$

$$\le \frac{X_0^p}{1-p} + X_0^{p-1}\frac{p}{1-p}\Big(\mathbb{E}\left[X_T\right] - X_0\Big). \tag{4.2.15b}$$

(ii) *there exists a martingale which attains equality in* (4.2.15a). *Further, for every* $\epsilon > 0$ *there exists a martingale such that*

$$0 \le \frac{X_0^p}{1-p} + X_0^{p-1}\frac{p}{1-p}\Big(\mathbb{E}\left[X_T\right] - X_0\Big) - \mathbb{E}\left[\bar{X}_T^p\right] < \epsilon. \tag{4.2.16}$$

(iii) *the inequality in* (4.2.15b) *is strict if and only if* $\mathbb{E}\left[X_T\right] < \infty$ *and* $\bar{X}_T \neq X_0$ *a.s.*

Proof. Following the calculations in (4.2.4), we see that

$$\mathbb{E}\left[\bar{X}_T^p\right] \le \frac{1}{1-\alpha}X_0^p + \frac{1}{(1-\alpha)(1-p)}\mathbb{E}\left[-\alpha^{1-p}X_T^p + pX_0^{p-1}X_T\right] = X_0^p f(\alpha),$$

where, with the notation m_r introduced in the statement of the Proposition,

$$f(\alpha) := \frac{1}{1-\alpha} + \frac{-\alpha^{1-p}m_p + pm_1}{(1-\alpha)(1-p)}, \qquad \alpha \in [0,1].$$

Next we prove the existence of a unique $\hat{\alpha} \in (0,1]$ such that $f(\hat{\alpha}) = \min_{\alpha\in[0,1]} f(\alpha)$. To do this, we first compute that

$$f'(\alpha) = \frac{h(\alpha)}{(1-p)(1-\alpha)^2}, \quad \text{where} \quad h(\alpha) := 1-p+pm_1 - (1-p+p\alpha)m_p\alpha^{-p}.$$

By direct calculation, we see that h is continuous and strictly increasing on $(0,1]$, with $h(0+) = -\infty$ and $h(1) = 1-p+pm_1-m_p$. Moreover, it follows from the Jensen inequality and the submartingale property of X that $m_p \le m_1^p$ and $m_1 \ge 1$.

This implies that $h(1) \geq 0$, and therefore there exists $\hat{\alpha} \in (0,1]$ such that $h \leq 0$ on $(0,\hat{\alpha}]$ and $h \geq 0$ on $[\hat{\alpha},1]$. Consequently, f is decreasing on $[0,\hat{\alpha}]$ and increasing on $[\hat{\alpha},1]$, proving that $\hat{\alpha}$ is the unique minimizer of f.

Now the first inequality (4.2.15a) follows by plugging the equation $h(\hat{\alpha}) = 0$ into the expression for f. Choosing $\alpha = 0$ in the expression for f yields the second inequality (4.2.15b) and hence together the claim in (i).

As for (ii), the claim regarding a martingale attaining equality in (4.2.15a) follows precisely as in the proof of Proposition 4.2.1. Let $\alpha \in (0,1)$. Then one computes as in the proof of Proposition 4.2.1 that $\mathbb{E}\left[\bar{B}^p_{\tau_\alpha}\right] = \frac{X_0^p}{1-p+p\alpha}$ and $\mathbb{E}\left[B_{\tau_\alpha}\right] = X_0$.

Assertion (iii) is trivial. □

4.2.4 No Improvements

Next, we prove that beyond the improvement stated in Proposition 4.2.1 no sharper bounds can be obtained from the inequalities of Theorem 4.1.1.

Proposition 4.2.5 (No Improvement of Doob's L^p-Inequality from Theorem 4.1.1)**.** *Let $p > 1$ and $\tilde{\boldsymbol{\zeta}} \in \tilde{\mathscr{Z}}$ be such that $\tilde{\zeta}_j(m) \neq \zeta^{\frac{p-1}{p}}(m) = \frac{p-1}{p}m$ for some $m > X_0$ and some j. Then there exists a martingale X such that*

$$\left(\frac{p}{p-1}\right)^p \mathbb{E}\left[X_T^p\right] - \frac{p}{p-1}X_0^p < \mathrm{UB}\left(X, \mathrm{pow}^p, \tilde{\boldsymbol{\zeta}}\right). \tag{4.2.17}$$

Proof. Let $\alpha > \frac{p-1}{p} =: \alpha^\star$ and take X^α satisfying

$$0 = X^\alpha_{t_1} = \ldots, X^\alpha_{t_{j-1}}, \qquad B_{\tau_\alpha} = X^\alpha_{t_j} = \cdots = X^\alpha_{t_n}$$

where B is Brownian motion started at X_0 and $\tau_\alpha = \inf\{u > 0 : B_u \leq \zeta^\alpha(\bar{B}_u)\}$. It follows easily that for this process X^α,

$$\mathrm{UB}\left(X^\alpha, \mathrm{pow}^p, \tilde{\zeta}_j\right) \leq \mathrm{UB}\left(X^\alpha, \mathrm{pow}^p, \tilde{\boldsymbol{\zeta}}\right)$$

and hence it is enough to prove the claim for $n = 1$ and $\tilde{\boldsymbol{\zeta}} = \tilde{\zeta}_j$.

For all $\alpha \in (\alpha^\star, \alpha^\star + \epsilon)$, $\epsilon > 0$, Proposition 4.1.4 yields existence of a non-empty, open interval $\mathcal{I}_\alpha$ such that

$$\forall m \in \mathcal{I}_\alpha : \quad \mathrm{UB}\left(X^\alpha, \phi^m, \zeta^\alpha\right) < \mathrm{UB}\left(X^\alpha, \phi^m, \tilde{\zeta}_j\right). \tag{4.2.18}$$

In fact, taking $\epsilon > 0$ small enough, $\mathcal{I}_\alpha$ can be chosen such that

$$\bigcap_{\alpha \in (\alpha^\star, \alpha^\star + \epsilon)} \mathcal{I}_\alpha \quad \supseteq \quad (m_1, m_2), \qquad\qquad X_0 < m_1 < m_2. \tag{4.2.19}$$

We can further (recalling the arguments in Case A and Case B in the proof of Proposition 4.1.4) assume that for all $\alpha \in (\alpha^\star, \alpha^\star + \epsilon)$:

$$\forall m \in (m_1, m_2) : \quad \mathrm{UB}\left(X^\alpha, \phi^m, \tilde{\zeta}_j\right) - \mathrm{UB}\left(X^\alpha, \phi^m, \zeta^\alpha\right) \geq \delta > 0. \tag{4.2.20}$$

The claim follows by letting $\alpha \downarrow \alpha^\star$ and using the asymptotic optimality of X^α, see (4.2.2). □

In addition to the result of Proposition 4.2.5 we prove that there is no "intermediate moment refinement of Doob's L^p-inequalities" in the sense formalized in the next Proposition. Intuitively, this could be explained by the fact that the p^{th} moment of a continuous martingale is continuously non-decreasing and hence does not add relevant information about the p^{th} moment of the maximum. Only the final p^{th} moment matters in this context.

Proposition 4.2.6 (No Intermediate Moment Refinement of Doob's L^p-Inequality)**.** *If $a_1, \ldots, a_n$ are such that for every continuous submartingale X, $X_0 = 0$, we have*

$$\mathbb{E}\left[\bar{X}_T^p\right] \leq \sum_{i=1}^{n} a_i \mathbb{E}\left[X_{t_i}^p\right] \tag{4.2.21}$$

or

$$\mathbb{E}\left[\bar{X}_T^p\right] \leq \sum_{i=1}^{n} a_i \mathbb{E}\left[|X_{t_i} - X_{t_{i-1}}|^p\right], \tag{4.2.22}$$

then

$$\left(\frac{p}{p-1}\right)^p \mathbb{E}\left[X_T^p\right] \leq \sum_{i=1}^{n} a_i \mathbb{E}\left[X_{t_i}^p\right] \tag{4.2.23}$$

or

$$\left(\frac{p}{p-1}\right)^p \mathbb{E}\left[X_T^p\right] \leq \sum_{i=1}^{n} a_i \mathbb{E}\left[|X_{t_i} - X_{t_{i-1}}|^p\right], \tag{4.2.24}$$

respectively.

Proof. From Peskir [80, Example 4.1] or our Proposition 4.2.1 we know that Doob's L^p-inequality given in (4.2.1b) is enforced by a sequence of continuous martingales (Y^ϵ) in the sense of (4.2.2). We will write $\left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|Y_T^\epsilon|^p\right] \simeq \mathbb{E}\left[\max_{t\leq T} |Y_t^\epsilon|^p\right]$.

Firstly consider the case of (4.2.21) and (4.2.23). By scalability of the asymptotically optimal martingales (Y^ϵ) we can assume

$$\mathbb{E}\left[|X_{t_n}|^p\right] = \mathbb{E}\left[|Y_{t_n}^\epsilon|^p\right].$$

In addition we can find times $u_1 \leq \cdots \leq u_{n-1}$ such that

$$\mathbb{E}\left[|X_{t_i}|^p\right] = \mathbb{E}\left[|Y_{u_i}^\epsilon|^p\right].$$

Therefore, writing $u_n = t_n = T$ and using asymptotic optimality of (Y^ϵ),

$$\begin{aligned}\left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|X_{t_n}|^p\right] = & \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|Y^\epsilon_{u_n}|^p\right] \simeq \mathbb{E}\left[\max_{t\leq T} |Y^\epsilon_t|^p\right] \\ & \overset{(4.2.21)}{\leq} \sum_{i=1}^n a_i \mathbb{E}\left[|Y^\epsilon_{u_i}|^p\right] = \sum_{i=1}^n a_i \mathbb{E}\left[|X_{t_i}|^p\right].\end{aligned}$$

Equation (4.2.23) follows.

Secondly consider the case of (4.2.22) and (4.2.24). Taking a martingale which is constant until time t_{i-1} and after t_i and using the the fact that Doob's L^p inequality is sharp yields

$$\left(\frac{p}{p-1}\right)^p \leq a_i \qquad \text{for all } i = 1, \dots, n.$$

Equation (4.2.24) follows. □

Remark 4.2.7. Analogous statements holds for Doob's L^1 inequality. This can be argued in the same way by using that Doob's L^1 inequality is attained (cf. e.g. Peskir [80, Example 4.2] or our Proposition 4.2.3) and observing that the function $x \mapsto x\log(x)$ is convex.

Chapter 5

Characterization of Market Models in the Presence of Traded Vanilla and Barrier Options

Calibration of models to market data is one major challenge in Mathematical Finance. Typically, one uses call option prices to incorporate information about univariate distributional properties. In some markets there are in addition options which are informative about joint distributional properties. Here we are interested in the case when there are traded Vanilla and Barrier options. We obtain a characterization of these joint distributions for the stock and its running maximum in the case of traded options with multiple maturities. Our characterization requires a kind of decomposition of certain call price functions and once it is obtained, we have an explicit expression for certain joint probabilities in the models characterized by this decomposition. We discuss interpolations of these joint probabilities which yield a fully specified marginal joint distribution which is consistent with the market.

Once one is given marginal joint distributions, we note that there are methods for defining diffusion-type models consistent with these marginals, cf. research by Carr [19], Cox et al. [30], Forde [41] and Forde et al. [42].

Notation The underlying asset will be denoted by S. For its maximum we write $M_T = \sup_{t\leq T} S_t$. The standard Markovian time-shift operator is denoted by $\theta_t(\omega) := (\omega_u)_{u\geq t}$. The first hitting time of B is denoted by H_B.

5.1 Characterization of Market Models

In this section we present our main results. We characterize the existence of a market model, both in the case of a single and multiple maturities.

5.1.1 Market Data and Market Models

Suppose $c(K_1), \ldots, c(K_n)$ are the prices for call options with strikes $0 < K_1 < \cdots < K_n$, respectively. Further, let $\boldsymbol{b} = \big(b(B_1), \ldots, b(B_m)\big)$ be the prices for simple barrier options $\mathbb{1}_{\{S_T\geq B_j\}}$ with barrier levels $S_0 =: B_0 < B_1 \leq \cdots \leq B_m$, respectively.

Definition 5.1.1 (Market Model). *A market model is a filtered probability space* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ *where the filtration* $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ *satisfies the usual hypothesis and there is an* $\mathbb{F}$*-adapted martingale* S *defined on this space satisfying*

$$\mathbb{E}\left[(S_T - K_i)^+\right] = c(K_i), \qquad i = 1, \ldots, n, \tag{5.1.1a}$$

$$\mathbb{E}\left[\mathbb{1}_{\{M_T\geq B_j\}}\right] = b(B_j), \qquad j = 1, \ldots, m. \tag{5.1.1b}$$

Using this definition, it is clear how to extend the notion of market model to a setting where there are options with multiple maturities.

For our main result we require the following assumption.

Assumption (i) (Asset). *Assume that the asset S*

(a) has continuous trajectories,

(b) has zero cost of carry (e.g. when interest rates are zero),

(c) is (strictly) positive.

5.1.2 One Maturity

We now state and prove our main result in the case of one maturity. This result will be extended to multiple maturities in the next section. The proof of this extension will be an induction over the number of maturities and hence will rely on the one maturity statement.

Theorem 5.1.2 (Characterization Market Model – One Maturity). *Let Assumption* (i) *hold. Then there exists a market model if and only if*

1. *there exists a call price function*[1] $c_\mu(\cdot) = \int (x - \cdot)^+ \mu(\mathrm{d}x)$ *which interpolates the given call prices* $c(K_1), \dots, c(K_n)$, $-c'_\mu(0+) = 1$.

2. *there exist* $\mathfrak{c}^{B_1}, \dots, \mathfrak{c}^{B_m}$ *such that for all* $j = 1, \dots, m$,

$$\mathfrak{c}^{B_j} : \mathbb{R}_{\geq 0} \to [0, S_0] \qquad \textit{is convex,} \tag{5.1.2a}$$

$$0 \leq \mathfrak{c}^{B_1} \leq \dots \leq \mathfrak{c}^{B_m} \leq \mathfrak{c}^{B_{m+1}} := c_\mu, \tag{5.1.2b}$$

$$-\frac{\mathrm{d}\mathfrak{c}^{B_j}}{\mathrm{d}x}(0+) = 1, \quad \mathfrak{c}^{B_j}(0) = S_0, \quad \mathfrak{c}^{B_j}(x) = 0, \qquad \forall x \geq B_j, \tag{5.1.2c}$$

$$-\frac{\mathrm{d}\mathfrak{c}^{B_j}}{\mathrm{d}x}(B_j -) = b(B_j), \tag{5.1.2d}$$

$$x \mapsto \mathfrak{c}^{B_{j+1}}(x) - \mathfrak{c}^{B_j}(x) + b(B_j)(B_j - x)^+ \qquad \textit{is convex.} \tag{5.1.2e}$$

[1] i.e. a non-negative convex function c such that $-\frac{\partial c}{\partial x}(0+) \leq 1$ and $c(x) \to 0$ as $x \to \infty$, cf. Davis and Hobson [33].

This market model can be chosen with bounded support (but does not have to).

Furthermore, in the market models characterized by (5.1.2a)–(5.1.2e) *we have for* $0 \leq x < B_j$,

$$\mathbb{P}\left[S_T > x, M_T < B_j\right] = -\frac{\mathrm{d}\mathfrak{c}^{B_j}}{\mathrm{d}x}(x+) - b(B_j). \tag{5.1.3}$$

Proof. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a market model, i.e.

$$\begin{aligned} \mathbb{E}\left[(S_T - K_i)^+\right] &= c(K_i) \qquad \forall i = 1, \dots, n, \\ \mathbb{P}\left[M_T \geq B_j\right] &= b(B_j) \qquad \forall j = 1, \dots, m. \end{aligned}$$

By continuity of S the Dambis-Dubins-Schwarz time change yields

$$(S_t)_{t \leq T} = (X_{\rho_t})_{t \leq T}$$

where X is Geometric Brownian motion started at S_0 and $\rho_t = \langle \log(S) \rangle_t$.

Define

$$\mathfrak{c}^{B_j}(x) := \mathbb{E}\left[\left(X_{\rho_T \wedge H_{B_j}} - x\right)^+\right] \qquad j = 1, \dots, m, \quad x \in \mathbb{R}. \tag{5.1.4}$$

Clearly $\mathfrak{c}^{B_1}, \dots, \mathfrak{c}^{B_m}$ satisfy (5.1.2a)–(5.1.2c).

Note

$$b(B_j) = \mathbb{P}\left[M_T \geq B_j\right] = \mathbb{P}\left[X_{\rho_T \wedge H_{B_j}} \geq B_j\right] = -\frac{\mathrm{d}\mathfrak{c}^{B_j}}{\mathrm{d}x}(B_j-)$$

which is condition (5.1.2d). Note also

$$b(B_j) = \mathbb{P}\left[\rho_T \geq H_{B_j}\right].$$

Let X^{B_j} denote a Geometric Brownian motion started at B_j. We have for $j < m$,

$$
\begin{aligned}
& \mathfrak{c}^{B_{j+1}}(x) - \mathfrak{c}^{B_j}(x) \\
= \ & \mathbb{E}\left[\mathbb{1}_{\{\rho_T < H_{B_j}\}} \left(X_{\rho_T \wedge H_{B_j}} - x\right)^+ + \mathbb{1}_{\{\rho_T \geq H_{B_j}\}} \left(X_{\rho_T \wedge H_{B_j}} - x\right)^+\right] \\
& -\mathbb{P}\left[\rho_T \geq H_{B_j}\right] \cdot (B_j - x)^+ - \mathbb{E}\left[\left(X_{\rho_T \wedge H_{B_j}} - x\right)^+\right] \\
& +\mathbb{E}\left[\left(X_{\rho_T \wedge H_{B_{j+1}}} - x\right)^+ \mathbb{1}_{\{\rho_T \geq H_{B_j}\}}\right] \\
= \ & \mathbb{P}\left[\rho_T \geq H_{B_j}\right] \cdot \left(-(B_j - x)^+ + \mathbb{E}\left[\left(X_\psi^{B_j} - x\right)^+\right]\right)
\end{aligned}
\tag{5.1.5}
$$

where ψ is an independent stopping time (independent of X) such that

$$
\mathbb{P}[\psi \leq x] = \mathbb{P}\left[(\rho_T \wedge H_{B_{j+1}}) - H_{B_j} \leq x \middle| \rho_T \geq H_{B_j}\right].
$$

These two properties of ψ together with the Markov property of X ensure that ψ embeds the same distribution as required by the above equation. Then, since $x \mapsto \mathbb{E}\left[\left(X_\psi^{B_j} - x\right)^+\right]$ is convex condition (5.1.2e) follows.

Now assume that conditions (5.1.2a)–(5.1.2e) hold. By the Skorokhod embedding theorem the claim is true for $m = 0$. Inductively, let us assume that the claim is true until $m - 1$, i.e. there exist stopping times $\gamma_1 \leq \cdots \leq \gamma_{m-1}$, $\gamma_j \leq H_{B_j}$, such that for $j = 1, \ldots, m-1$,

$$
\begin{aligned}
\mathfrak{c}^{B_j}(x) &= \mathbb{E}\left[\left(X_{\gamma_j} - x\right)^+\right] \qquad \forall x \in \mathbb{R}, \\
b(B_j) &= \mathbb{P}\left[\gamma_j = H_{B_j}\right].
\end{aligned}
$$

Define

$$
\varphi_m(x) := \frac{\mathfrak{c}^{B_m}(x) - \mathbb{E}\left[\left(X_{\gamma_{m-1}} - x\right)^+\right]}{b(B_{m-1})} + (B_{m-1} - x)^+. \tag{5.1.6}
$$

It follows by induction hypothesis and by (5.1.2c) and (5.1.2e) that φ_m defines a call price function, i.e. φ_m is convex, $\varphi_m(0) = B_{m-1}$, $-\varphi_m'(0+) = 1$ and $\varphi_m(x) = 0$ for $x \geq B_m$. Hence, there exists a stopping time ϑ_m such that

$$\varphi_m(x) = \mathbb{E}\left[\left(X_{\vartheta_m}^{B_{m-1}} - x\right)^+\right] \qquad \forall x \in \mathbb{R}. \tag{5.1.7}$$

Therefore, by (5.1.6), (5.1.7) and induction hypothesis,

$$\begin{aligned} &\mathfrak{c}^{B_m}(x) \\ &= \mathbb{P}\left[\gamma_{m-1} = H_{B_{m-1}}\right] \cdot \mathbb{E}\left[\left(X_{\vartheta_m}^{B_{m-1}} - x\right)^+\right] + \mathbb{E}\left[\left(X_{\gamma_{m-1}} - x\right)^+ \mathbb{1}_{\left\{\gamma_{m-1} < H_{B_{m-1}}\right\}}\right] \\ &= \mathbb{E}\left[(X_{\gamma_m} - x)^+\right] \end{aligned} \tag{5.1.8}$$

where

$$\gamma_m := \begin{cases} \gamma_{m-1} & \text{if } \gamma_{m-1} < H_{B_{m-1}}, \\ \gamma_{m-1} + \vartheta_m & \text{if } \gamma_{m-1} = H_{B_{m-1}}. \end{cases} \tag{5.1.9}$$

Clearly, $\gamma_m \leq H_{B_m}$ and

$$\mathbb{P}\left[\gamma_m = H_{B_m}\right] = \mathbb{P}\left[X_{\gamma_m} = B_m\right] = \mathbb{P}\left[X_{\gamma_m} \geq B_m\right] = -\mathfrak{c}^{B_m}(B_m-) = b(B_m)$$

and

$$\mathbb{P}\left[\gamma_m \geq H_{B_j}\right] = \mathbb{P}\left[\gamma_{m-1} \geq H_{B_j}\right] = b(B_j) \qquad \forall j \leq m-1. \tag{5.1.10}$$

By the same argument as above, there exists $\gamma_{m+1} \geq \gamma_m$ such that

$$\begin{aligned} &\mathbb{P}\left[\gamma_{m+1} \geq H_{B_j}\right] = \mathbb{P}\left[\gamma_m \geq H_{B_j}\right] = b(B_j), \qquad j = 1, \ldots, m, \\ &\mathbb{E}\left[\left(X_{\gamma_{m+1}} - x\right)^+\right] = \mathfrak{c}^{B_{m+1}}(x) = c_\mu(x). \end{aligned}$$

Taking $S_t := X_{\frac{t}{T-t} \wedge \gamma_{m+1}}$ yields a market model.

As for the bounded support claim, we note that choosing a sufficiently large upper bound for the support of μ, will allow us to choose $c_\mu = \mathfrak{c}^{B_{m+1}}$ in a way to satisfy (5.1.2e) for $j = m$.

Finally, using the notation from the proof, we get by rearranging (5.1.8) and writing $\frac{\mathrm{d}^+}{\mathrm{d}x}$ for the right-derivative,

$$\begin{aligned}
&\mathbb{P}\left[S_T > x, M_T < B_j\right] \\
&= -\frac{\mathrm{d}^+}{\mathrm{d}x}\left(\mathbb{E}\left[\left(X_{\gamma_j} - x\right)^+ \mathbb{1}_{\{\gamma_j < H_{B_j}\}}\right]\right) \\
&= -\frac{\mathrm{d}^+}{\mathrm{d}x}\left(\mathfrak{c}^{B_{j+1}}(x) - b(B_j)\left[\frac{\mathfrak{c}^{B_{j+1}(x)} - \mathfrak{c}^{B_j}(x)}{b(B_j)} + (B_j - x)^+\right]\right) \\
&= -\frac{\mathrm{d}\mathfrak{c}^{B_j}}{\mathrm{d}x}(x+) - b(B_j)
\end{aligned}$$

for $0 \leq x < B_j$. □

5.1.3 Multiple Maturities

We now extend Theorem 5.1.2 to the setup of multiple maturities.

For simplicity we assume that the strikes and barriers at each maturity coincide. Denote the right endpoint of the support of the measure μ by $r_\mu := \inf\big\{x : \mu((x,\infty)) = 0\big\}$.

Take $0 < T_1 < \cdots < T_k$. Suppose $c_l(K_1), \ldots, c_l(K_n)$ are the prices for call options with strikes $0 < K_1 < \cdots < K_n$ and maturity T_l, $l = 1, \ldots, k$. Further, let $\boldsymbol{b}_l = \big(b_l(B_1), \ldots, b_l(B_m)\big)$ be the prices for simple barrier options with barrier levels $S_0 =: B_0 < B_1 \leq \cdots \leq B_m$ and (deterministic) maturity T_l, $l = 1, \ldots, k$. Set $b_l(B_0) := 1$ and $b_0 \equiv 0$.

Theorem 5.1.3 (Characterization Market Model – Multiple Maturities)**.** *Let Assumption* (i) *hold. Then there exists a market model if and only if*

1. *there exist call price functions* $c_{\mu_l}(\cdot) = \int (x-\cdot)^+ \mu_l(\mathrm{d}x)$ *which interpolate the given call prices for (deterministic) maturity* T_l, $l = 1, \dots, k$, *and satisfy* $c_{\mu_1} \leq \dots \leq c_{\mu_k}$, $-c'_{\mu_l}(0+) = 1$. *For* $B_{m+1} := r_{\mu_k}$ *we set* $b_l(B_{m+1}) := -\frac{\partial c_{\mu_l}}{\partial x}(B_{m+1}-)$.

2. *there exist* $\left\{ c_l^{B_j} \right\}$ *such that for all* $j = 1, \dots, m+1$ *and* $l = 1, \dots, k$,

$$c_l^{B_j} : \mathbb{R}_{\geq 0} \to [0, S_0] \qquad \text{is convex,} \tag{5.1.11a}$$

$$-\frac{\mathrm{d}c_l^{B_j}}{\mathrm{d}x}(0+) = b_l(B_{j-1}), \quad c_l^{B_j}(0) = b_l(B_{j-1}) \cdot B_{j-1}, \tag{5.1.11b}$$

$$-\frac{\mathrm{d}c_l^{B_j}}{\mathrm{d}x}(B_j-) = b_l(B_j), \quad c_l^{B_j}(x) = 0, \qquad x \geq B_j, \tag{5.1.11c}$$

$$c_{l-1}^{B_j}(x) + (b_l - b_{l-1})(B_{j-1}) \cdot (B_{j-1} - x)^+ \leq c_l^{B_j}(x) \qquad \forall x, \tag{5.1.11d}$$

$$\sum_{j=1}^{m} \left(c_l^{B_j}(x) - b_l(B_j)(B_j - x)^+ \right) + c_l^{B_{m+1}}(x) = c_{\mu_l}(x) \quad \forall x. \tag{5.1.11e}$$

This market model can be chosen with bounded support (but does not have to).

Furthermore, in the market models characterized by (5.1.11a)–(5.1.11e) *we have for* $0 \leq x < B_j$,

$$\mathbb{P}\left[S_{T_l} > x, M_{T_l} < B_j\right] = -\frac{\mathrm{d}\mathfrak{c}_l^{B_j}}{\mathrm{d}x}(x+) - b_l(B_j) \tag{5.1.12}$$

where

$$\mathfrak{c}_l^{B_j}(x) := \sum_{i=1}^{j-1} \left(c_l^{B_i}(x) - b_l(B_i) \cdot (B_i - x)^+ \right) + c_l^{B_j}(x). \tag{5.1.13}$$

Remark 5.1.4 (Connection to Theorem 5.1.2). Identifying for $j = 1, \dots, m+1$,

$$\mathfrak{c}^{B_j}(x) = \mathfrak{c}_1^{B_j}(x) \tag{5.1.14}$$

easily shows that Theorem 5.1.3 in the case $l = 1$ implies Theorem 5.1.2.

Conversely, identifying (using the notation of Theorems 5.1.2 and 5.1.3)

$$c_1^{B_j}(x) = \mathfrak{c}^{B_j}(x) - \mathfrak{c}^{B_{j-1}}(x) + b(B_{j-1})(B_{j-1} - x)^+ \tag{5.1.15}$$

for $j = 1, \ldots, m+1$, with the convention $\mathfrak{c}^{B_0}(x) := (B_0 - x)^+$, easily shows that Theorem 5.1.2 implies Theorem 5.1.3 in the case $l = 1$.

Therefore, Theorem 5.1.3 provides a slightly different alternative characterization than Theorem 5.1.2.

Proof of Theorem 5.1.3. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a market model, i.e. for $l = 1, \ldots, k$,

$$\begin{aligned} \mathbb{E}\left[(S_{T_l} - K_i)^+\right] &= c_{\mu_l}(K_i) \qquad \forall i = 1, \ldots, n, \\ \mathbb{P}\left[M_{T_l} \geq B_j\right] &= b_l(B_j) \qquad \forall j = 1, \ldots, m. \end{aligned}$$

By continuity of S the Dambis-Dubins-Schwarz time change yields

$$(S_t)_{t \leq T_k} = (X_{\rho_t})_{t \leq T_k}$$

where X is Geometric Brownian motion started at S_0 and $\rho_t = \langle \log(S) \rangle_t$.

Define for $j = 1, \ldots, m+1,\ l = 1, \ldots, k$ and $x \in \mathbb{R}$,

$$c_l^{B_j}(x) := \mathbb{E}\left[\mathbb{1}_{\left\{\rho_{T_l} \geq H_{B_{j-1}}\right\}} \left(X_{\rho_{T_l} \wedge H_{B_j}} - x\right)^+\right]. \tag{5.1.16}$$

Similarly as before, properties (5.1.11a)–(5.1.11c) are easily verified. We compute

$$c_l^{B_j}(x) - c_{l-1}^{B_j}(x) = \mathbb{E}\left[\mathbb{1}_{\left\{\rho_{T_{l-1}} < H_{B_{j-1}}, \rho_{T_l} \geq H_{B_{j-1}}\right\}} \left(X_{\rho_{T_l} \wedge H_{B_j}} - x\right)^+\right]$$

which defines a measure with mass $(b_l - b_{l-1})(B_{j-1})$ and mean $(b_l - b_{l-1})(B_{j-1}) \cdot B_{j-1}$ and hence (5.1.11d) follows. As for (5.1.11e) we note that

$$
\begin{aligned}
& \sum_{j=1}^{m} \left(c_l^{B_j}(x) - b_l(B_j)(B_j - x)^+ \right) + c_l^{B_{m+1}}(x) \\
= & \sum_{j=1}^{m} \mathbb{E}\left[\mathbb{1}_{\left\{\rho_{T_l} \geq H_{B_{j-1}}\right\}} \left(X_{\rho_{T_l} \wedge H_{B_j}} - x \right)^+ - \mathbb{1}_{\left\{\rho_{T_l} \geq H_{B_j}\right\}} (B_j - x)^+ \right] \\
& + \mathbb{E}\left[\mathbb{1}_{\left\{\rho_{T_l} \geq H_{B_m}\right\}} \left(X_{\rho_{T_l}} - x \right)^+ \right] \\
= & \sum_{j=1}^{m} \mathbb{E}\left[\mathbb{1}_{\left\{H_{B_{j-1}} \leq \rho_{T_l} < H_{B_j}\right\}} \left(X_{\rho_{T_l}} - x \right)^+ \right] + \mathbb{E}\left[\mathbb{1}_{\left\{\rho_{T_l} \geq H_{B_m}\right\}} \left(X_{\rho_{T_l}} - x \right)^+ \right] \\
= & c_{\mu_l}(x).
\end{aligned}
$$

Now assume that conditions (5.1.11a)–(5.1.11e) hold. By Remark 5.1.4 the claim is true for $k = 1$. Inductively, let us assume that the claim is true until $k - 1$, in particular there exist $\eta_l, \ l = 1 \dots, k - 1$, such that

$$
c_l^{B_j}(x) = \mathbb{E}\left[\mathbb{1}_{\left\{\eta_l \geq H_{B_{j-1}}\right\}} \left(X_{\eta_l \wedge H_{B_j}} - x \right)^+ \right], \qquad x \in \mathbb{R}, \tag{5.1.17a}
$$

$$
b_l(B_j) = \mathbb{P}\left[M_{\eta_l} \geq B_j\right] = \mathbb{P}\left[\eta_l \geq H_{B_j}\right]. \tag{5.1.17b}
$$

By (5.1.11a) the two functions

$$
\begin{aligned}
& x \mapsto s_k^{B_j}(x) := c_{k-1}^{B_j}(x) + (b_k - b_{k-1})(B_{j-1}) \cdot (B_{j-1} - x)^+, \\
& x \mapsto c_k^{B_j}(x),
\end{aligned}
$$

are convex and by (5.1.11d) we have

$$
0 \leq s_k^{B_j} \leq c_k^{B_j}.
$$

By (5.1.11b) the means of the measures corresponding to $s_k^{B_j}$ and $c_k^{B_j}$ coincide,

$$s_k^{B_j}(0) = b_{k-1}(B_{j-1}) \cdot B_{j-1} + (b_k - b_{k-1})(B_{j-1}) \cdot B_{j-1} = b_k(B_{j-1}) \cdot B_{j-1} = c_k^{B_j}(0),$$

and the same is true for the masses,

$$\frac{\partial s_k^{B_j}}{\partial x}(0+) = b_{k-1}(B_{j-1}) + (b_k - b_{k-1})(B_{j-1}) = b_k(B_{j-1}) = \frac{\partial c_k^{B_j}}{\partial x}(0+).$$

Hence, by Strassen [92] there exists a stopping time $\varrho_k^{B_j}$ which embeds the measure corresponding to $s_k^{B_j}$ into the measure corresponding to $c_k^{B_j}$.

We define recursively

$$\eta_k \wedge H_{B_1} = \begin{cases} \varrho_k^{B_1} & \text{if } \eta_{k-1} < H_{B_1}, \\ H_{B_1} & \text{else,} \end{cases} \tag{5.1.18a}$$

$$\eta_k \wedge H_{B_2} = \begin{cases} \eta_k \wedge H_{B_1} & \text{if } \eta_k \wedge H_{B_1} < H_{B_1}, \\ \eta_k \wedge H_{B_1} + \varrho_k^{B_2} \circ \theta_{\eta_k \wedge H_{B_1}} & \text{if } \eta_{k-1} < H_{B_2},\ \eta_k \wedge H_{B_1} = H_{B_1}, \\ H_{B_2} & \text{else,} \end{cases} \tag{5.1.18b}$$

$$\vdots$$

$$\eta_k \wedge H_{B_m} = \begin{cases} \eta_k \wedge H_{B_{m-1}} & \text{if } \eta_k \wedge H_{B_{m-1}} < H_{B_{m-1}}, \\ \eta_k \wedge H_{B_{m-1}} + \varrho_k^{B_m} \circ \theta_{\eta_k \wedge H_{B_{m-1}}} & \text{if } \eta_{k-1} < H_{B_m},\ \eta_k \wedge H_{B_{m-1}} = H_{B_{m-1}}, \\ H_{B_m} & \text{else,} \end{cases} \tag{5.1.18c}$$

$$\eta_k = \begin{cases} \eta_k \wedge H_{B_m} & \text{if } \eta_k \wedge H_{B_m} < H_{B_m}, \\ \eta_k \wedge H_{B_m} + \varrho_k^{B_{m+1}} \circ \theta_{\eta_k \wedge H_{B_m}} & \text{if } \eta_k \wedge H_{B_m} = H_{B_m}. \end{cases} \tag{5.1.18d}$$

Next we show that this construction recovers the correct quantities. It is already visible from (5.1.18a)–(5.1.18d) that the condition for $\varrho_k^{B_j}$ is triggered only for paths which hit the level B_{j-1} and hence $\varrho_k^{B_j}$ does not change the probability that $B_i, i \leq j-1$, is hit. However, it can change the probability that B_j is hit.

Denote $\mathbb{E}^{s_k^{B_j}}\left[(X_\gamma - x)^+\right]$ the expectation of X_γ where $X_0 \sim \nu_k^{B_j}$ and $\nu_k^{B_j}$ is the measure corresponding to $s_k^{B_j}$. With this definition it follows inductively that for $j = 1, \ldots, m+1$,

$$
\begin{aligned}
& \mathbb{E}\left[\mathbb{1}_{\{\eta_k \geq H_{B_{j-1}}\}}\left(X_{\eta_k \wedge H_{B_j}} - x\right)^+\right] \\
= \; & \mathbb{E}\left[\left(\mathbb{1}_{\{\eta_{k-1} \geq H_{B_{j-1}}\}} + \mathbb{1}_{\{\eta_{k-1} < H_{B_{j-1}}, \eta_k \wedge H_{B_{j-1}} = H_{B_{j-1}}\}}\right)\left(X_{\eta_k \wedge H_{B_j}} - x\right)^+\right] \\
= \; & \mathbb{E}\left[\mathbb{1}_{\{\eta_{k-1} \geq H_{B_{j-1}}\}}\left(X_{\left(\eta_{k-1} + \varrho_k^{B_j} \circ \theta_{\eta_{k-1}}\right) \wedge H_{B_j}} - x\right)^+\right] \\
& + \mathbb{E}\left[\mathbb{1}_{\{\eta_{k-1} < H_{B_{j-1}}, \eta_k \wedge H_{B_{j-1}} = H_{B_{j-1}}\}}\left(X_{\left(H_{B_{j-1}} + \varrho_k^{B_j} \circ \theta_{H_{B_{j-1}}}\right) \wedge H_{B_j}} - x\right)^+\right] \\
= \; & \mathbb{E}^{s_k^{B_j}}\left[\left(X_{\varrho_k^{B_j}} - x\right)^+\right] = c_k^{B_j}(x)
\end{aligned}
$$

and therefore

$$
\begin{aligned}
& \mathbb{P}\left[\eta_k \geq H_{B_j}\right] = -\frac{\partial c_k^{B_j}}{\partial x}(B_j-) \overset{(5.1.11c)}{=} b_k(B_j), \\
& \mathbb{P}\left[\eta_k \wedge H_{B_j} \geq H_{B_i}\right] = \mathbb{P}\left[\eta_k \wedge H_{B_i} \geq H_{B_i}\right] = b_k(B_i), \qquad \forall i < j.
\end{aligned}
$$

Finally, the embedding property follows as

$$\begin{aligned}
&\mathbb{E}\left[(X_{\eta_k}-x)^+\right] \\
= \quad & \sum_{j=1}^{m}\mathbb{E}\left[\mathbb{1}_{\left\{\eta_k\in\left[H_{B_{j-1}},H_{B_j}\right)\right\}}(X_{\eta_k}-x)^+\right]+\mathbb{E}\left[\mathbb{1}_{\{\eta_k\geq H_{B_m}\}}(X_{\eta_k}-x)^+\right] \\
= \quad & \sum_{j=1}^{m}\mathbb{E}\left[\mathbb{1}_{\{\eta_k\geq H_{B_{j-1}}\}}\left(X_{\eta_k\wedge H_{B_j}}-x\right)^+\right]-\mathbb{P}\left[\eta_k\geq H_{B_j}\right](B_j-x)^+ \\
+ \quad & \mathbb{E}\left[\mathbb{1}_{\{\eta_k\geq H_{B_m}\}}(X_{\eta_k}-x)^+\right] \\
= \quad & \sum_{j=1}^{m}\left(c_k^{B_j}(x)-b_k(B_j)(B_j-x)^+\right)+c_k^{B_{m+1}}(x)\overset{(5.1.11e)}{=}c_{\mu_k}(x).
\end{aligned}$$

The claim regarding the existence of a market model with a bounded support and equation (5.1.12) follow in the same way as in the proof of Theorem 5.1.2. □

5.2 Interpolation

Theorems 5.1.2 and 5.1.3 require to "decompose" some call price functions into a sequence of intermediate convex functions. Given this decomposition, equations (5.1.3) and (5.1.12) partially specify the joint marginal distributions of (S, M) implied by this decomposition. Next we discuss how to consistently interpolate these joint probabilities from (5.1.3) and (5.1.12) in barriers and time.

5.2.1 Computation of Decomposition

The strength of our main result hinges on the computation of the quantities $\left\{c_l^{B_j}\right\}$ as described in Theorem 5.1.3. In practice, a simple and efficient method to compute them would be to discretize in space and solve a linear programming (LP) problem. If N is the number of discretization points in space, the number of variables is $\mathcal{O}(N)$ in the one maturity case. Of course, a naive linear program which optimizes over the joint distribution of maximum and terminal value is also possible, but it has more variables: in the one maturity case it is of order $\mathcal{O}(N^2)$. In addition, one would also need to make sure that one obtains a valid joint distribution for the stock *and* its maximum, see conditions (5.2.3a)–(5.2.3b) below. In the case of k maturities, the number of variables is $\mathcal{O}(kN)$ to compute $\left\{c_l^{B_j}\right\}$ as described in Theorem 5.1.3. When trying to use a naive linear program one would need to ensure several condition regarding the *ordering* of the joint distributions, see Section 5.2.3 below. This might not be straightforward to implement.

By changing the objective function of the LP one is able to achieve two things. Firstly, one can regularize the solution by "penalizing" e.g. gaps in support or atoms. Secondly, one can find solutions with additional features such as maximizing the expectation of a given payoff, which would yield a upper robust price bound for this payoff. In particular, one could calculate the maximal price of a simple barrier option with barrier $B \neq B_j$.

5.2.2 Interpolation of Barrier Prices

In the models characterized by the decomposition $\left\{c_l^{B_j}\right\}$, equation (5.1.3) of Theorem 5.1.2 partially specifies the joint (tail) distribution for (S_T, M_T) as

$$\bar{F}_T(x, B_j) := \mathbb{P}\left[S_T > x, M_T \geq B_j\right] = -\frac{\mathrm{d}c_\mu}{\mathrm{d}x}(x+) + \left(\frac{\mathrm{d}\mathfrak{c}^{B_j}}{\mathrm{d}x}(x+) + b(B_j)\right)\mathbb{1}_{\{x<B_j\}} \tag{5.2.1}$$

for $x \geq 0$ and $j = 0, \ldots, m$.

Recall from Theorem 5.1.2 that in order to incorporate a bounded support for the joint distribution we can impose that $B_m = K_n$ and $b(B_m) = 0 = c_\mu(K_n)$ for B_m sufficiently large.

In order to obtain an unbounded support, we have to extrapolate. In this context, note that Theorem 5.1.2 is readily extended to countably many call and barrier options with increasing strikes and barriers.

Rogers [84, Theorem 2.2] characterizes the set of all these possible distributions by some integrability condition and two properties of the function

$$d(m) = \mathbb{E}\left[S_T \middle| M_T > m\right], \tag{5.2.2}$$

namely

$$d(m) \geq m, \tag{5.2.3a}$$

$$m \mapsto d(m) \qquad \text{is non-decreasing.} \tag{5.2.3b}$$

By construction we have for $j = 1, \ldots, m$,

$$d(B_j) \geq B_j, \tag{5.2.4a}$$

$$d(B_{j-1}) \leq d(B_j). \tag{5.2.4b}$$

The simplest interpolation is to use linear interpolation in barrier option prices. Let $a \in (0,1)$ and $B = aB_{j-1} + (1-a)B_j$. Setting

$$\bar{F}_T(x, B) := a\bar{F}_T(x, B_{j-1}) + (1-a)\bar{F}_T(x, B_j) \tag{5.2.5}$$

yields a joint (tail) distribution $\bar{F}_T$ which satisfies (5.2.3a)–(5.2.3b).

We will refer by $\mathcal{M}_T$ to the measure corresponding to $\bar{F}_T$.

5.2.3 Interpolation in Time

An interpolation in time is not as easily obtained because the interpolations for fixed maturities cannot be done independently of each other.

Rost [87, Theorem 4] characterizes when there exists a martingale S such that

$$(S_{T_1}, M_{T_1}) \sim \mathcal{M}_{T_1}, \qquad (S_{T_2}, M_{T_2}) \sim \mathcal{M}_{T_2}.$$

However, his characterization is not very explicit in our setup.

To see in a simple example that things can indeed go wrong, consider

$$\delta_{\{S_0\}} \preccurlyeq_{\mathrm{c}} \mu_1 \preccurlyeq_{\mathrm{c}} \mu_2, \quad b_{\mu_1} \not\leq b_{\mu_2} \quad \mathcal{M}_{T_1} := \mathcal{L}\left(X_{\tau^{\mathrm{AY}}_{\mu_1}}, M_{\tau^{\mathrm{AY}}_{\mu_1}}\right), \quad \mathcal{M}_{T_2} := \mathcal{L}\left(X_{\tau^{\mathrm{AY}}_{\mu_2}}, M_{\tau^{\mathrm{AY}}_{\mu_2}}\right)$$

where b_μ denotes the barycenter function of μ and τ^{AY}_μ denoted the Azéma-Yor embedding of μ. It is known, see e.g. Rogers [84], that both $\mathcal{M}_{T_1}$ and $\mathcal{M}_{T_2}$ can be embedded starting from the Dirac measure $\delta_{\{S_0\}}$ and that $M_{\tau^{\mathrm{AY}}_{\mu_1}} \sim \mu_1^{\mathrm{HL}}$ and $M_{\tau^{\mathrm{AY}}_{\mu_2}} \sim \mu_2^{\mathrm{HL}}$, respectively, where μ^{HL} denotes the Hardy-Littlewood transform of μ. However, it follows from Brown et al. [18] that it is not possible to embed $\mathcal{M}_{T_2}$ after $\mathcal{M}_{T_1}$ because of $b_{\mu_1} \not\leq b_{\mu_2}$.

Therefore, if the interpolation method from Section 5.2.2 yield $\mathcal{M}_{T_1}$ and $\mathcal{M}_{T_2}$ as above, then it follows that this interpolation is inconsistent with a model.

5.2.4 Interpolation via Skorokhod Embedding

One theoretical way to specify the joint laws $(\mathcal{M}_t)_{t\leq T}$ is to use one's favourite Skorokhod embedding $(\tau_t)_{t\leq T}$ as described in the proofs of Theorems 5.1.2 and 5.1.3. The functions $c_l^{B_j}$ can be interpreted as the footprint of the marginal law evolution of the process. This yields marginal laws as

$$\mathcal{M}_t := \mathcal{L}\left(X_{\tau_t}, M_{\tau_t}\right) \qquad \forall t \leq T. \tag{5.2.6}$$

Part II

Realized Variance Constraints

Chapter 6

Modelling Setup

Joint Publication Large parts of Chapters 6–8 will be submitted as part of a joint publication with Jan Obłój, see Obłój and Spoida [76].

As motivated in the introduction of the thesis, see Sections 1.2.3 and 1.3.2, we follow Mykland [67, 68, 69, 70] and incorporate constraints on the realized variance of price trajectories into the robust framework. That is, we formulate and solve robust pricing and hedging problems with a finite number of put options and constraints on realized variance, see the robust framework specification in Section 1.3.2. Tackling this problem via a time-change and Skorokhod embedding approach, we will see that this setup is different to the classical SEP in two ways: firstly, we are not given a unique measure μ to embed and secondly, the embedding has to be completed in a finite amount of time.

Notation The set of continuous functions $D \to R$, where $D, R \subseteq \mathbb{R}$, will be denoted by $\mathbb{C}(D; R)$. Similarly, we will use $\mathbb{C}^1$, $\mathbb{C}^{1,2}$, etc. for differentiable functions.

Denote the first hitting time of the level N by

$$H_N := \inf\{t > 0 \ : \ X_t = N\} \tag{6.0.1}$$

and write $H_{a,b}$ for the exit time of $[a, b]$. For an open set $\mathcal{O}$ we will denote

$$\tau_{\mathcal{O}} := \inf\{t > 0 \ : \ X_t \notin \mathcal{O}\}. \tag{6.0.2}$$

Recalling l_μ and r_μ from (2.1.8) we define the support of a measure μ as

$$\text{supp}(\mu) := [l_\mu, r_\mu]. \tag{6.0.3}$$

Organisation of Second Part of the Thesis In the remainder of this chapter we carefully introduce the formal modelling framework in its full generality and point out its flexibility to incorporate both market *information* and diverse types of *beliefs*.

Chapters 7 and 8 are devoted to convex Vanilla and Lookback options, respectively. More precisely, in the case of convex Vanilla options we recover the result of Mykland [67] by using the methodology of Skorokhod embeddings. We show that the robust pricing problem is essentially solved by Root's solution of the SEP, see Section 7.1. The robust pricing problem of Lookback and Barrier options requires new optimal solutions to the SEP. The solutions turn out to be Root stopping times for a certain two dimensional process, see Section 8.1.

Looking at the dual problem, we show how our robust price bounds can be enforced by pathwise hedging, i.e. we prove pathwise duality results in Sections 7.2 and 8.2. In our context, the superhedging strategy has to be valid if the realized variance of the stock does not exceed a given threshold.

We characterize existence of a market model by a suitable notion of arbitrage, see Sections 7.3 and 8.4.

Following the ideas of Dupire [38], Cox and Wang [24] and Oberhauser and dos Reis [72] we explain in Section 8.5 how to compute numerically the robust prices and hedging strategies.

6.1 Market Data and Skorokhod Embeddings

We impose the following standing assumption on the market. Similar assumptions are used by Davis et al. [32].

Assumption (ii) (Market)**.** *We assume that*

- *the market is frictionless, i.e. there are no transaction costs, no trading restrictions, no illiquidity, trading is executed instantaneously, arbitrary amounts of the asset S can be traded.*
- *the asset S has zero cost of carry. For example this is the case when S is a forward price, S pays dividends continuously at a rate equal to the prevailing interest rate, or S is the exchange rate between two currencies with the same interest rates.*
- *all counterparties are default-free and hence we ignore cash-flows due to mark-to-market and collateralization procedures.*

We suppose that n put options written on the underlying S with strikes $0 \leq K_1 \leq \cdots \leq K_n$ and maturity T are traded today at prices $p^\star(K_1), \dots, p^\star(K_n)$. In addition, call-put-parity is assumed to hold,

$$c^\star(K_i) = p^\star(K_i) + S_0 - K_i, \qquad i = 1, \dots, n, \tag{6.1.1}$$

where $c^\star(K_i)$ denotes today's price of a call option on S with strike K_i and maturity T.

Remark 6.1.1 (Bubbles). The assumption that call-put-parity (6.1.1) holds excludes the case that the asset S is a strict local martingale, which may be of interest in modelling financial bubbles, see e.g. Cox and Hobson [27], Jarrow et al. [55] and Cox et al. [26].

Our market data is insufficient to imply a unique risk-neutral distribution of S_T via the Breeden-Litzenberger formula. We think of this situation in the formalism of equivalence classes of probability measures on $\mathbb{R}_+$,

$$[\mu^\star] := \left\{ \mu \in \mathcal{M}^1(\mathbb{R}_+) : \int_{\mathbb{R}} (K_i - s)^+ \, \mu(\mathrm{d}s) = p^\star(K_i), \; i = 1, \dots, n \right\} \qquad (6.1.2)$$

where $\mathcal{M}^1(\mathbb{R}_+)$ denotes the set of probability measures on $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. We also note the one-to-one relation to

$$[p^\star] := \left\{ p \text{ is a put price function } : \; p(K_i) = p^\star(K_i) \; \forall i \text{ and } p(0) = 0 \right\} \qquad (6.1.3)$$

and similarly for $[c^\star]$ and $[U_{\mu^\star}]$, where we recall the definition of these quantities in (1.1.2)–(1.1.3). In the following we will use $p^\star$, $c^\star$ and $U_{\mu^\star}$ interchangeably, see (1.1.4).

Later, we will be interested in stopping times τ such that $\mathcal{L}\left(X_{\tau \wedge \Xi}\right) \in [\mu^\star]$ where X is a Geometric Brownian motion and $\Xi > 0$. By Tanaka's formula

$$u^\tau(t, x) := -\mathbb{E}\left[|X_{\tau \wedge t} - x|\right] = u^\tau(0, x) - \mathbb{E}\left[L^x_{\tau \wedge t}\right] \qquad (6.1.4)$$

and then the "embedding requirement" reads

$$\mathbb{E}\left[L^{K_i}_{\tau \wedge \Xi}\right] = -S_0 + K_i + 2c^\star(K_i) + u(0, K_i), \qquad \forall i = 1, \dots, n, \qquad (6.1.5)$$

and we note that the right hand side of (6.1.5) is known from market data. From (6.1.5) we get the following intuition: the local time $L^{K_i}_{\tau \wedge \cdot}$ increases when the process X "crosses" K_i. The idea in Sections 7.1 and 8.1 is to design τ in such a way to both reprice the puts and to maximize some criterion. The latter will be achieved by constructing a decision rule based on time and some other metric of the path up to this time, to control whether a path can cross the level K_i or not.

6.2 Path Spaces

We want to formalize the belief that price trajectories are continuous and have bounded realized variance. To this end, we start with the path space $\mathbb{C}\left([0,T];\mathbb{R}_+\right)$ of positive continuous functions[1] on $[0,T]$ and restrict to a subset $\mathfrak{P}_\Xi$ of paths which respects a bound on total realized variance. To achieve this, we need to define quadratic variation pathwise.

Let $\omega \in \mathbb{C}\left([0,T];\mathbb{R}_+\right)$. For $m \in \mathbb{N}$ define $\mathbb{D}^m := \{k2^{-m} : k \in \mathbb{Z}\}$, $\pi_0^m := 0$, and for $k \geq 1$,

$$\pi_k^m(\omega) := \inf\left\{t \in [\pi_{k-1}^m, T] \;:\; \omega(t) \in \mathbb{D}^m \backslash \omega(\pi_{k-1}^m)\right\} \wedge T. \tag{6.2.1}$$

Write $\boldsymbol{\pi}^m = \{\pi_k^m, k \in \mathbb{N}_0\}$ and $\boldsymbol{\pi} = \{\boldsymbol{\pi}^m, m \in \mathbb{N}\}$.

The following definition is based on Vovk [96, Section 2].

Definition 6.2.1 (Quadratic Variation)**.** *For a continuous function $\omega : [0,T] \to \mathbb{R}$ define*

$$A_t^m(\omega) := \sum_{k=1}^{\infty} \left(\omega(\pi_k^m(\omega) \wedge t) - \omega(\pi_{k-1}^m(\omega) \wedge t)\right)^2, \quad t \in [0,T], \quad m \in \mathbb{N}. \tag{6.2.2}$$

The function ω has quadratic variation if $(A^m(\omega))_{m\in\mathbb{N}}$ converges in the supremum norm on $\mathbb{C}\left([0,T];\mathbb{R}\right)$. The limit, if it exists, is called the quadratic variation of ω and will be denoted by $\langle\omega\rangle$.

[1] We are restricting here directly to continuous functions instead of càdlàg functions in order to avoid unnecessary technicalities, but the same formalism can be setup for càdlàg paths.

Our definition of quadratic variation above differs from the definition of Föllmer [40]. However, as shown by Vovk [96, Proposition 6.1], if ω has quadratic variation in the sense above, the quadratic variation of ω along $\boldsymbol{\pi}(\omega)$ exists in the sense of Föllmer [40] and the two definitions coincide. This will enable us to apply the Itô-Föllmer formula. In addition, as shown by Vovk [96, Section 7], if ω has quadratic variation, then also $f(\omega)$ has quadratic variation, where f is a continuously differentiable function.

Now we can incorporate the belief on total realized variance.

Definition 6.2.2 (Path Space)**.** *A path space $\mathfrak{P}$ is a Borel measurable subset of $\mathbb{C}([0,T];\mathbb{R}_+)$. First we exclude paths ω which do not have quadratic variation,*

$$\mathfrak{P}^{\mathrm{qv}} := \Big\{\omega \in \mathbb{C}([0,T];\mathbb{R}_+) \ : \ \omega(0) = S_0, \ \ \omega \textit{ has quadratic variation } \Big\}. \quad (6.2.3)$$

Fix $\Xi > 0$. The restricted path space with upper bound on realized variance is defined as the set

$$\mathfrak{P}_\Xi := \Big\{\omega \in \mathfrak{P}^{\mathrm{qv}} \ : \ \langle \log(\omega)\rangle_T \leq \Xi, \ t \mapsto \langle \log(\omega)\rangle_t \textit{ is strictly increasing}\Big\}. \quad (6.2.4)$$

6.3 Market Instruments, Pricing Operator and Market Models

We want to allow for dynamic trading in the stock and static positions in other market instruments, e.g. put options.

Definition 6.3.1 (Market Instruments)**.** *The set of market instruments is a set $\mathcal{Y}$ of functionals defined on a path space $\mathfrak{P}$.*

The set of market instruments we will consider consists only of constants and a finite number of put options, i.e.

$$\mathcal{Y} = \big\{F,\ (K - S_T)^+ \ :\ F \in \mathbb{R},\quad K \in \mathbb{K}\big\}, \tag{6.3.1}$$

where

$$\mathbb{K} = \{K_1, \dots, K_n\}. \tag{6.3.2}$$

Next we formalize which dynamic trading strategies are permitted. Let $\mathfrak{P}$ be a path space (later we will be mainly interested in $\mathfrak{P} = \mathfrak{P}_\Xi$). A mapping ψ : $[0,T] \times \mathfrak{P} \to \mathbb{R}$ is called progressively measurable if for any $\omega, \tilde{\omega} \in \mathfrak{P}$

$$\omega_s = \tilde{\omega}_s, \quad \forall s \leq [0,t] \qquad \Longrightarrow \qquad \psi(t,\omega) = \psi(t,\tilde{\omega}), \tag{6.3.3}$$

where we assume that $\psi(\cdot \wedge t, \cdot)$ is $B([0,t]) \otimes \sigma(X_s, s \leq t)$-measurable for every $t \leq T$. Take an increasing sequence of stopping times $0 =: \gamma_0 \leq \gamma_1 \leq \gamma_2 \leq \dots$ such that $\lim_{i\to\infty} \gamma_i(\omega) = \infty$ for all $\omega \in \mathfrak{P}$. Then we allow the following strategies,

$$\phi_t(\omega) = \sum_{i=0}^{\infty} \phi_i(t,\omega) \mathbb{1}_{\left[\gamma_i(\omega), \gamma_{i+1}(\omega) \wedge T\right)}(t) \tag{6.3.4}$$

where $\phi_i : [0,T] \times \mathfrak{P} \to \mathbb{R}$ is progressively measurable for all $i \in \mathbb{N}$ and of either form,

$$\phi_i(t,\omega) = f(t,\omega(t)) \qquad \text{for } f \in \mathbb{C}^{0,1}([0,T];\mathbb{R}), \tag{6.3.5a}$$

$$\phi_i(t,\omega) = \sum_{l=1}^{\infty} g_l \mathbb{1}_{\{B_l\}}(t,\omega) \qquad g_l \in \mathbb{R},\ B_l \subseteq [0,T] \times \mathfrak{P}. \tag{6.3.5b}$$

For these strategies we can define pathwise integrals w.r.t. every $\omega \in \mathfrak{P}^{\mathrm{qv}}$ which coincide a.s. with the classical definition of the stochastic integral.

Definition 6.3.2 (Admissible Pathwise Strategies)**.** *Let $\mathfrak{P} \subseteq \mathfrak{P}^{\text{qv}}$ be a path space. We define the set of admissible pathwise dynamic trading strategies as*

$$\Phi := \Big\{\phi : [0,T] \times \mathfrak{P} \to \mathbb{R} \ : \phi \textit{ is of the form } (6.3.4) \textit{ and } (\phi \cdot S)_t \geq -C_1 - C_2 S_t \\ \forall t \leq \Xi \textit{ and some } C_1, C_2 > 0\Big\}. \tag{6.3.6}$$

The set of admissible pathwise trading strategies is then defined by

$$\mathcal{A} := \left\{ \sum_{i=1}^{k} \lambda_i Y_i + \int_0^T \phi_u(S) \mathrm{d}S_u \ : \ \lambda_i \in \mathbb{R}, \ Y_i \in \mathcal{Y}, \ \phi \in \Phi, \ k \in \mathbb{N} \right\}. \tag{6.3.7}$$

Remark 6.3.3 (Interpretation of $\boldsymbol{\lambda}$). The constants $\boldsymbol{\lambda} = (\lambda_i)$ in the definition of $\mathcal{A}$ represent the static position in the market instruments, in our case in the put options.

Definition 6.3.4 (Pricing Operator)**.** *Let* Lin($\mathcal{Y}$) *denote the set of finite linear combinations of elements from the set of market instruments $\mathcal{Y}$ defined in* (6.3.1)*. A pricing operator on $\mathcal{Y}$ is a linear functional $\mathcal{P} : \text{Lin}(\mathcal{Y}) \to \mathbb{R}$ satisfying*

$$\mathcal{P}[1] = 1, \tag{6.3.8a}$$

$$\mathcal{P}\left[(K - S_T)^+\right] = p^\star(K) \qquad \forall K \in \mathbb{K}. \tag{6.3.8b}$$

By the assumption of frictionless markets we also set

$$\mathcal{P}\left[\int_0^T \phi_u \mathrm{d}S_u\right] := 0 \qquad \forall \phi \in \Phi \tag{6.3.9}$$

so that $\mathcal{P}$ is defined on Lin($\mathcal{A}$)*.*

Next we want to consistently price a non-traded instrument by superreplication

Definition 6.3.5 (Pathwise Superhedging Price)**.** *Take $\mathfrak{P} \subseteq \mathfrak{P}^{\text{qv}}$ and let $G : \mathfrak{P} \to \mathbb{R}$ be a measurable payoff functional. The pathwise superhedging price on* $(\mathfrak{P}, \mathcal{A}, \mathcal{P})$

for the payoff G is defined by

$$\mathbb{A}\left(G;\mathfrak{P},\mathcal{A},\mathcal{P}\right) \ := \ \inf\Big\{\mathcal{P}\left[V\right] \ : \ V\left(\omega\right) \geq G\left(\omega\right) \ \forall \omega \in \mathfrak{P}, \ V \in \mathcal{A}\Big\}. \tag{6.3.10}$$

The following definition is based on Cox and Obłój [28, Definition 1.1].

Definition 6.3.6 (Model and Market Model)**.** *Let $\mathfrak{P}$ be a path space. A $\mathfrak{P}$-model is a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual hypotheses, and an $\mathbb{F}$-adapted stochastic process S defined on this space with paths in $\mathfrak{P}$ $\mathbb{P}$-a.s.*

A $\mathfrak{P}$-model is called a $(\mathfrak{P}, \mathcal{Y}, \mathcal{P})$-market model if S is a $\mathbb{P}$-martingale and

$$\mathbb{E}\left[Y(S_t : t \leq T)\right] = \mathcal{P}\left[Y\right] \qquad \forall Y \in \mathcal{Y} \tag{6.3.11}$$

where we implicitly assume that the left hand side is well defined.

6.4 Time-Change

Next we introduce an auxiliary process X which will be the key mathematical tool for pricing and hedging. X is defined as the coordinate process on $\Omega^X := \mathbb{C}\left([0,\Xi];\mathbb{R}_+\right)$ and we fix $\mathcal{F} := \mathcal{B}(\Omega^X)$ as the Borel sigma algebra of Ω^X, $\mathbb{F}$ as the natural filtration of X and $\mathbb{P}$ as a probability measure such that X is a $\mathbb{P}$-Geometric Brownian motion on $(\Omega^X, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Defining

$$\Omega^{\text{GBM}} := \big\{\omega \in \mathbb{C}([0,\Xi];\mathbb{R}_+) \ : \ \omega(0) = S_0, \ \langle \log(\omega)\rangle_t = t \ \ \forall t \leq \Xi\big\}. \tag{6.4.1}$$

it is clear that such a $\mathbb{P}$ satisfies $\mathbb{P}\left[\Omega^{\text{GBM}}\right] = 1$ and hence we could alternatively replace Ω^X with Ω^{GBM}.

Next we explain the link of $\omega \in \Omega^{\mathrm{GBM}}$ to $\omega^S \in \mathfrak{P}_\Xi$. To this end let S be the coordinate process on $\mathfrak{P}_\Xi$, i.e. $S_t(\omega) = \omega(t)$ for $t \leq T$ and $\omega \in \mathfrak{P}_\Xi$. The process S has the interpretation of the asset price process. Denote the realized variance of S and its right-continuous inverse by ρ and $\rho^{\leftarrow}$, respectively,

$$\rho : [0,T] \times \mathfrak{P}_\Xi \to [0,\Xi], \quad (t,\omega) \mapsto \rho_t(\omega) \;:= \langle \log \omega \rangle_t, \tag{6.4.2a}$$

$$\rho^{\leftarrow} : [0,\Xi] \times \mathfrak{P}_\Xi \to [0,T], \quad (\xi,\omega) \mapsto \rho^{\leftarrow}_\xi(\omega) := \inf \{t > 0 : \rho_t(\omega) \geq \xi\}. \tag{6.4.2b}$$

Consider the time-change $\tilde{X}_\xi := S_{\rho^{\leftarrow}_\xi}$, $\xi \in [0,\Xi]$. By continuity of S we have $S_t = \tilde{X}_{\rho_t}$ for all $t \in [0,T]$. Now let $\left(\Omega^S, \mathcal{F}^S, \mathbb{F}^S, \mathbb{P}^S\right)$ be a $(\mathfrak{P}_\Xi, \mathcal{Y}, \mathcal{P})$-market model. Then we observe that by the Dambis-Dubins-Schwarz time change, cf. Oberhauser and dos Reis [72, Lemma 6 and 7] for more details, $\tilde{X}$ is a $\mathbb{P}^S$-Geometric Brownian motion with unit volatility w.r.t. the filtration $\tilde{\mathbb{F}} := \left(\tilde{\mathcal{F}}_\xi := \mathcal{F}^S_{\rho^{\leftarrow}_\xi}\right)_{\xi \in [0,\Xi]}$.

It is important to observe that, via the time-change ρ, every $\omega^S \in \mathfrak{P}_\Xi$ can be associated with $(\omega(u))_{u \leq s}$ for some $\omega \in \Omega^{\mathrm{GBM}}$ and some $s \leq \Xi$.

Chapter 7

Vanilla Options with Beliefs about Realized Variance

The goal of this chapter is to answer questions about robust pricing, hedging and arbitrage for convex Vanilla options $G(S_T)$, where G satisfies the following Assumption.

Assumption (iii) (Vanilla Payoff). *The payoff function $G : [0, \infty) \to \mathbb{R}$ is convex and satisfies the following polynomial growth conditions,*

$$\limsup_{s\to\infty} \frac{G'(s)}{s^p} < \infty, \qquad \limsup_{s\to 0} G'(s) \cdot s^p < \infty, \tag{7.0.1}$$

for some $p \geq 0$ and where G' denotes the left-derivative of the convex function G.

In the following, if not made clear otherwise, X is a $\mathbb{P}$-Geometric Brownian motion on the filtered probability space $(\Omega^X, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and expectations with respect to $\mathbb{P}$ are denoted by the operator $\mathbb{E}$, see Section 6.4.

7.1 Robust Pricing

The goal of this section is to compute the robust price of the convex Vanilla payoff $G(S_T)$, that is, we assume there are n traded put options maturing at time T with strikes $K_1, \dots, K_n$ and given prices $p^\star(K_1), \dots, p^\star(K_n)$, respectively, and our goal is to find the upper price bound for the payoff $G(S_T)$, consistent with these market traded puts and the realized variance constraint $\langle \log(S) \rangle_T \leq \Xi$.

The Root stopping time $\tau(\boldsymbol{b})$ based on $\boldsymbol{b} \in \mathbb{R}^n$,

$$\tau(\boldsymbol{b}) := \inf \Big\{ u > 0 \ : \ X_u = K_i, \ u \geq b_i \ \text{ for some } i = 1, \dots, n \Big\}, \tag{7.1.1}$$

will play a central role, see Figure 7.1.1 for an illustration. Recall from Section 1.1.3 the Root embedding $\tau_{\mathrm{R}}(\mu)$ for a measure μ. Intuitively, the Root stopping time yields an embedding which "stops as late as possible". Our time-change approach corresponds to change in volatility. Therefore, since convex options are non-decreasing with volatility, the Root stopping time appears as a natural candidate solution for the robust pricing problem. We now formalize this intuition.

7.1.1 The Case of Bounded Support

In this section we consider the case of a risk-neutral distribution of S_T with bounded support, i.e. we assume there is a traded put with strike $K_{n+1} = N$ and price $p^\star(K_{n+1}) = K_{n+1} - S_0$. In addition, we assume there is a put with strike $K_0 = \epsilon \geq 0$ and price $p^\star(K_0) = 0$.

Definition 7.1.1 (Maximal Law in Convex Order)**.** *Assume there exists a stopping time τ such that*

$$\mathbb{E}\left[(K_i - X_{\tau \wedge \Xi})^+ \right] = p^\star(K_i), \qquad \forall i = 0, \dots, n+1. \tag{7.1.2}$$

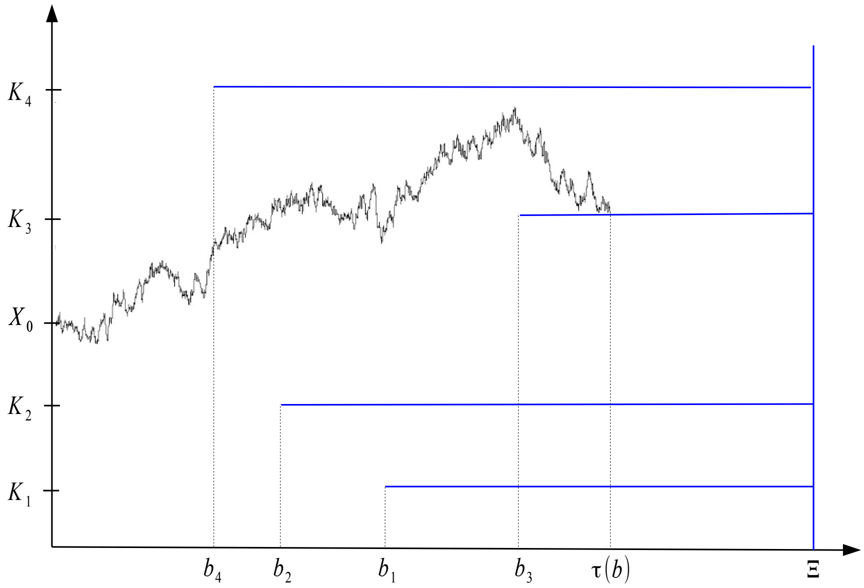

FIGURE 7.1.1: The stopping time $\tau(\boldsymbol{b}) \wedge \Xi$ is defined as the first time the process X hits the blue boundary.

Then we denote by $\mu^\star_{\epsilon,N}$ *the unique measure which is supported on* $\{K_0, K_1, \ldots, K_{n+1}\}$ *and satisfies*

$$\int (K_i - x)^+ \mu^\star_{\epsilon,N}(\mathrm{d}x) = p^\star(K_i), \qquad i = 0, \ldots, n+1. \tag{7.1.3}$$

Remark 7.1.2 (Uniform Integrability). It is clear, but important to note that for any stopping time τ which satisfies (7.1.2) the stopped process $(X_{t\wedge\tau})_{t\le\Xi}$ is a uniformly integrable martingale. More generally, this uniform integrability assertion remains true for any stopping time (assuming that $\Xi < \infty$ is a deterministic constant).

The following result links the measure $\mu^\star_{\epsilon,N}$ from Definition 7.1.1 to the robust superhedging price of convex Vanilla options.

Proposition 7.1.3 (Robust Superhedging Price for Convex Vanilla Options)**.** *Set* $\mathbb{K}$ *in* (6.3.2) *as* $\mathbb{K} := \{K_0, K_1, \ldots, K_{n+1}\}$ *and suppose* $\mathcal{Y}$, $\mathcal{A}$ *and* $\mathcal{P}$ *are as described in Section 6.3. Assume* (7.1.2) *holds for some stopping time* τ*. Let*

$G : [K_0, K_{n+1}] \to \mathbb{R}$ be a convex function. Then,

$$\begin{aligned}\mathbb{A}\left(G; \mathbb{C}\left([0,T]; [K_0, K_{n+1}]\right), \mathcal{A}, \mathcal{P}\right) &= \mathbb{A}\left(G; \mathbb{C}\left([0,T]; [K_0, K_{n+1}]\right), \mathcal{Y}, \mathcal{P}\right) \\ &= \int G \mathrm{d}\mu^\star_{\epsilon,N} \geq \int G \mathrm{d}\mu_{\epsilon,N}\end{aligned} \tag{7.1.4}$$

where $\mu_{\epsilon,N}$ is any probability measure which satisfies $\int (K_i - x)^+ \mu_{\epsilon,N}(\mathrm{d}x) = p^\star(K_i)$ for $i = 0, \ldots, n+1$.

Proof. By convexity of G there are unique $\lambda_0, \ldots, \lambda_{n+1}$ such that

$$G(S) \leq \sum_{i=0}^{n+1} \lambda_i \left(K_i - S\right)^+ \qquad \forall S \in [K_0, K_{n+1}], \tag{7.1.5}$$

with equality at $S = K_i, i = 0, \ldots, n+1$. Following Davis and Hobson [33] we obtain that the cost of this superhedging position is $\int_0^\infty G \mathrm{d}\mu^\star_{\epsilon,N}$.

Let τ be a Skorokhod embedding for $\mu^\star_{\epsilon,N}$. Then

$$\left(S^\star_t := X_{\frac{t}{T-t} \wedge \tau}\right)_{t \leq T} \tag{7.1.6}$$

yields a $(\mathbb{C}([0,T]; [K_0, K_{n+1}]), \mathcal{A}, \mathcal{P})$-market model which achieves equality in (7.1.5). The claims in (7.1.4) follow. □

Now we are in the position to prove a first version of our robust pricing result for convex Vanilla options.

Proposition 7.1.4 (Robust Pricing of Convex Vanilla Options – Bounded Support)**.** *Let G satisfy Assumption* (iii) *on page 137. Fix $\Xi > 0$ and assume there exists a stopping time τ such that* (7.1.2) *holds. Recall $\mu^\star_{\epsilon,N}$ from Definition 7.1.1 and write $\mu^\star_N = \mu^\star_{0,N}$. Then,*

$$\sup_\tau \mathbb{E}\left[G\left(X_{\tau \wedge \Xi}\right)\right] = \mathbb{E}\left[G\left(X_{\tau_{\mathrm{R}}\left(\mu^\star_N\right) \wedge \Xi}\right)\right] \tag{7.1.7}$$

where the supremum is over all stopping times τ satisfying (7.1.2).

Proof. We first consider the case when $G \in \mathbb{C}^2(\mathbb{R}_{\geq 0}; \mathbb{R})$ is strictly convex. Applying Itô's formula we get for any stopping time γ,

$$\begin{aligned} G(X_\gamma) &= G(X_0) + \int_0^\gamma G'(X_u)\mathrm{d}X_u \\ &\quad + \frac{1}{2}\int_0^{\gamma\wedge\Xi} G''(X_u)X_u^2\mathrm{d}u + \frac{1}{2}\int_{\gamma\wedge\Xi}^{\gamma} G''(X_u)X_u^2\mathrm{d}u. \end{aligned} \tag{7.1.8}$$

The stochastic integral $\int_0^\cdot G'(X_u)\mathrm{d}X_u$ is a martingale because by our assumption on G', see (7.0.1), we have $G'(x) \leq c_1 + c_2(x^p + x^{-p})$ for some constants c_1, c_2, and then $\int_0^\Xi \mathbb{E}\left[(G'(X_u)X_u^2)^2\right] \mathrm{d}u < \infty$.

We will denote by θ_t the time shift operator,

$$\theta_t : \mathbb{C}(\mathbb{R}_+; \mathbb{R}) \to \mathbb{C}(\mathbb{R}_+; \mathbb{R}), \qquad \omega \mapsto \theta_t(\omega) := (\omega_u)_{u\geq t}. \tag{7.1.9}$$

Let τ satisfy (7.1.2). By Proposition 7.1.3 we have $\mathcal{L}(X_{\tau\wedge\Xi}) \preccurlyeq_{\mathrm{c}} \mu_N^\star$. Hence, by e.g. Obłój [73, Proposition 8.1] there exists a stopping time η w.r.t. the filtration $(\mathcal{F}_{\tau+t})_{t\geq 0}$ such that

$$\varrho := \tau \wedge \Xi + \eta \circ \theta_{\tau\wedge\Xi}$$

embeds the law $\mu_N^\star$ and $(X_{\varrho\wedge t})_{t\geq 0}$ is a uniformly integrable martingale. The stopping time $\tau_{\mathrm{R}}(\mu_N^\star)$ embeds law $\mu_N^\star$ by definition. Since by assumption $G'' > 0$, the results of Cox and Wang [29, Sec. 4.3 and 5], see also Rost [86, Theorem 2], applied to the Geometric Brownian motion X and the stopping times $\varrho \leq H_N$ and $\tau_{\mathrm{R}}(\mu_N^\star) \leq H_N$ yield

$$\mathbb{E}\left[\int_{\Xi\wedge\varrho}^{\varrho} G''(X_u)X_u^2\mathrm{d}u\right] \geq \mathbb{E}\left[\int_{\Xi\wedge\tau_{\mathrm{R}}(\mu_N^\star)}^{\tau_{\mathrm{R}}(\mu_N^\star)} G''(X_u)X_u^2\mathrm{d}u\right].$$

Using the fact that

$$\mathbb{E}\left[G\left(X_{\varrho}\right)\right] = \mathbb{E}\left[G\left(X_{\tau_{\mathrm{R}}\left(\mu_N^\star\right)}\right)\right]$$

we obtain together with (7.1.8) that

$$\mathbb{E}\left[G\left(X_{\tau\wedge\Xi}\right)\right] \leq \mathbb{E}\left[G\left(X_{\varrho\wedge\Xi}\right)\right] \leq \mathbb{E}\left[G\left(X_{\tau_{\mathrm{R}}\left(\mu_N^\star\right)\wedge\Xi}\right)\right]. \tag{7.1.10}$$

Finally, in the case when G is not necessarily smooth, we can find a sequence G_k of smooth, strictly convex functions such that $G_k \downarrow G$ to obtain by monotone convergence and (7.1.10),

$$\begin{aligned}\mathbb{E}\left[G\left(X_{\tau\wedge\Xi}\right)\right] &= \lim_{k\to\infty} \mathbb{E}\left[G_k\left(X_{\tau\wedge\Xi}\right)\right] \\ &\leq \lim_{k\to\infty} \mathbb{E}\left[G_k\left(X_{\tau_{\mathrm{R}}\left(\mu_N^\star\right)\wedge\Xi}\right)\right] = \mathbb{E}\left[G\left(X_{\tau_{\mathrm{R}}\left(\mu_N^\star\right)\wedge\Xi}\right)\right]\end{aligned}$$

which proves (7.1.7). □

Remark 7.1.5 (Bound of Support). In the setting of Proposition 7.1.4, if $b_i = 0$ for some i then any stopping time τ which satisfies (7.1.2) also satisfies $\tau \leq H_{K_i}$, which corresponds to a bound on the support. These cases can be dealt with by specifying this property in the path space (i.e. support of the model) $\mathfrak{P}$. With this argument we will often restrict attention to

$$b_i^N > 0 \qquad \forall i = 1, \ldots, n. \tag{7.1.11}$$

Corollary 7.1.6 (Bound for Rays of Root Stopping Time)**.** *In the setting of Proposition 7.1.4 and Remark 7.1.5, write*[1] $\tau_{\mathrm{R}}(\mu_N^\star) = \tau(\boldsymbol{b}^N)$ *for a unique* $\boldsymbol{b}^N \in$

[1]Since $\mu_N^\star$ is atomic and of bounded support this is possible.

$(0,\infty)^{n+1}$. *Then,*

$$b_i^N \leq \Xi \qquad \forall\ i = 1, \dots, n+1. \tag{7.1.12}$$

Proof. By Proposition 7.1.4 we know that $\tau(\boldsymbol{b}^N) \wedge \Xi$ satisfies (7.1.2). Further, $\boldsymbol{b}^N$ is unique by (7.1.11) and (6.1.5). Moreover, by (6.1.5) we also deduce that

$$t \mapsto \mathbb{E}\left[L^{K_i}_{\tau(\boldsymbol{b}^N)\wedge t}\right] \qquad \text{is constant for } t \geq \Xi$$

which then implies (7.1.12). □

Corollary 7.1.7 (Maximization of Atoms at End of Support)**.** *In the setting of Propositions 7.1.3 and 7.1.4, writing* $\tau_{\mathrm{R}}(\mu^\star_{\epsilon,N}) = \tau(\boldsymbol{b}^{\epsilon,N})$ *for some* $\boldsymbol{b}^{\epsilon,N} \in \mathbb{R}^{n+2}$*, we have for all* $\xi \leq \Xi$*,*

$$\sup_\tau \mathbb{P}\left[X_{\tau\wedge\xi} = N\right] = \mathbb{P}\left[X_{\tau(\boldsymbol{b}^{\epsilon,N})\wedge\xi} = N\right], \tag{7.1.13a}$$

$$\sup_\tau \mathbb{P}\left[X_{\tau\wedge\xi} = \epsilon\right] = \mathbb{P}\left[X_{\tau(\boldsymbol{b}^{\epsilon,N})\wedge\xi} = \epsilon\right] \tag{7.1.13b}$$

where the supremum is over all stopping times τ *satisfying* (7.1.2)*.*

Proof. By the optimality property of the Root stopping time, cf. proof of Proposition 7.1.4, the call price function corresponding to $X_{\tau(\boldsymbol{b}^{\epsilon,N})\wedge\xi}$, denoted by $c^\star$, is at least as large as the call price function corresponding to any $X_{\tau\wedge\xi}$, denoted by c, where τ is a stopping times satisfying (7.1.2). Also note that $c(N) = 0 = c^\star(N)$ and $c(\epsilon) = X_0 - \epsilon = c^\star(\epsilon)$. Then, by the Breeden-Litzenberger formula, we have,

$$\begin{aligned}
\mathbb{P}\left[X_{\tau\wedge\xi} = N\right] = \mathbb{P}\left[X_{\tau\wedge\xi} \geq N\right] = -c'(N-) = \lim_{\epsilon\downarrow 0} \frac{c(N-\epsilon) - c(N)}{\epsilon} \\
\leq \lim_{\epsilon\downarrow 0} \frac{c^\star(N-\epsilon) - c^\star(N)}{\epsilon} = \mathbb{P}\left[X_{\tau(\boldsymbol{b}^{\epsilon,N})\wedge\xi} \geq N\right] = \mathbb{P}\left[X_{\tau(\boldsymbol{b}^{\epsilon,N})\wedge\xi} = N\right].
\end{aligned}$$

Analogously for the atom at ϵ. □

7.1.2 The Case of General Support

In this section we extend the result of the previous section to the case of a risk-neutral distribution of S_T with unbounded support.

We introduce the set of stopping times $\tau \leq \Xi$ which are consistent with the n put prices,

$$\mathcal{T}_\Xi^\star := \Big\{ \tau \in \mathcal{T}_\infty \ : \ \tau \leq \Xi, \ \ \mathbb{E}\left[(K_i - X_\tau)^+\right] = p^\star(K_i) \ \ \forall i = 1, \ldots, n \Big\}. \tag{7.1.14}$$

Theorem 7.1.8 (Robust Pricing of Convex Vanilla Options – General Support)**.** *Let $\mathcal{T}_\Xi^\star \neq \emptyset$ and suppose G satisfies Assumption* (iii) *on page 137. Then,*

$$\sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}\left[G\left(X_\tau\right)\right] = \lim_{N \to \infty} \mathbb{E}\left[G\left(X_{\tau_R\left(\mu_N^\star\right) \wedge \Xi}\right)\right] = \mathbb{E}\left[G\left(X_{\tau(\boldsymbol{b}^\star) \wedge \Xi}\right)\right] \tag{7.1.15}$$

for some $\boldsymbol{b}^\star \in [0, \Xi]^n$ which does not depend on Ξ and yields $\tau(\boldsymbol{b}^\star) \wedge \Xi \in \mathcal{T}_\Xi^\star$. Further, $b_i^\star$ is unique for every i such that $K_i \in \operatorname{supp}\left(\mathcal{L}\left(X_{\tau(\boldsymbol{b}^\star) \wedge \Xi}\right)\right)$ where we recall the definition of supp *in* (6.0.3)[2].

Proof. As in the proof of Proposition 7.1.4, Assumption (iii) ensures that all expectations we consider are finite and that the stochastic integral $\int_0^\cdot G'(X_u) \mathrm{d}X_u$ is a martingale. Moreover, it will be enough to prove the claim for $G \in \mathrm{C}^2(\mathbb{R}_{\geq 0}; \mathbb{R})$.

For any $\tau \in \mathcal{T}_\Xi^\star$ Propositions 7.1.3 and 7.1.4 yield a measure $\mu_{N,\tau}^\star$ with support contained in $[0, N]$ such that

$$\begin{aligned} &\mathcal{L}\left(X_{\tau \wedge H_N \wedge \Xi}\right) \in \left[\mathcal{L}\left(X_{\tau_{\mathrm{R}}\left(\mu_{N,\tau}^\star\right) \wedge \Xi}\right)\right], \\ &\mathcal{L}\left(X_{\tau \wedge H_N \wedge \Xi}\right) \preccurlyeq_{\mathrm{c}} \mathcal{L}\left(X_{\tau_{\mathrm{R}}\left(\mu_{N,\tau}^\star\right) \wedge \Xi}\right) \qquad \forall N \text{ sufficiently large,} \end{aligned} \tag{7.1.16}$$

[2] In the following we always take a "regular" version of $\boldsymbol{b}^\star$ in the sense that $b_i^\star = 0$ for $K_i \leq X_0$ ($K_i \geq X_0$) implies $b_j^\star = 0$ for all j such that $K_j \leq K_i$ ($K_j \geq K_i$).

where in (6.1.2) we use $K_1, \ldots, K_{n+1} = N$. Note that by monotone convergence,

$$\mathbb{E}\left[(K_i - X_{\tau\wedge H_N\wedge\Xi})^+\right] \uparrow p^\star(K_i) \qquad \text{as } N \to \infty \tag{7.1.17}$$

for all $i = 1, \ldots, n$ and $\tau \in \mathcal{T}_\Xi^\star$.

By convexity of G we have for any stopping time τ

$$\int_0^{\tau\wedge\Xi} G''(X_u)X_u^2 \mathrm{d}u \le \int_0^{\Xi} G''(X_u)X_u^2 \mathrm{d}u, \qquad \mathbb{E}\left[\int_0^{\Xi} G''(X_u)X_u^2 \mathrm{d}u\right] < \infty.$$

Using Itô's formula we obtain further

$$G(X_{\tau\wedge\Xi}) = G(X_0) + \int_0^{\tau\wedge\Xi} G'(X_u)\mathrm{d}X_u + \frac{1}{2}\int_0^{\tau\wedge\Xi} G''(X_u)X_u^2 \mathrm{d}u$$

and by taking expectation

$$\mathbb{E}\left[G(X_{\tau\wedge\Xi})\right] = G(X_0) + \frac{1}{2}\mathbb{E}\left[\int_0^{\tau\wedge\Xi} G''(X_u)X_u^2 \mathrm{d}u\right] < \infty. \tag{7.1.18}$$

Hence by applying dominated convergence at the second equality sign,

$$\begin{aligned}
\sup_{\tau\in\mathcal{T}_\Xi^\star} \mathbb{E}\left[G\left(X_{\tau\wedge\Xi}\right)\right] &\overset{(7.1.18)}{=} \sup_{\tau\in\mathcal{T}_\Xi^\star} \mathbb{E}\left[G(X_0) + \frac{1}{2}\int_0^{\tau\wedge\Xi} G''(X_u)X_u^2 \mathrm{d}u\right] \\
&= \sup_{\tau\in\mathcal{T}_\Xi^\star} \lim_{N\to\infty} \mathbb{E}\left[G(X_0) + \frac{1}{2}\int_0^{\tau\wedge\Xi\wedge H_N} G''(X_u)X_u^2 \mathrm{d}u\right] \\
&\overset{(7.1.18)}{=} \sup_{\tau\in\mathcal{T}_\Xi^\star} \lim_{N\to\infty} \mathbb{E}\left[G\left(X_{\tau\wedge\Xi\wedge H_N}\right)\right] \\
&\overset{(7.1.16)}{\le} \sup_{\tau\in\mathcal{T}_\Xi^\star} \lim_{N\to\infty} \mathbb{E}\left[G\left(X_{\tau_R\left(\mu^\star_{N,\tau}\right)\wedge\Xi}\right)\right]
\end{aligned} \tag{7.1.19}$$

We identify $\tau_{\mathrm{R}}(\mu^\star_{N,\tau}) = \tau(\boldsymbol{b}^{N,\tau})$ for some $\boldsymbol{b}^{N,\tau} \in \mathbb{R}^{n+1}$. From Corollary 7.1.6, applied to $\mu^\star_{\bar{N},\tau}$, $\bar{N} \ge N$, we know that

$$\sup_{\tau\in\mathcal{T}_\Xi^\star} \max_{i=1,\ldots,n} b_i^{\bar{N},\tau} \le \Xi \quad \forall \bar{N} \ge N,$$

and consequently

$$\left\{(b_1^{\bar N,\tau},\dots,b_n^{\bar N,\tau})\ :\ \tau\in\mathcal{T}_\Xi^\star,\ \bar N\geq N\right\}\subseteq[0,\Xi]^n.$$

Let $N_k\to\infty$ as $k\to\infty$. Choosing a maximizing sequence $(\tau_k)_{k\in\mathbb{N}}$ for (7.1.19) and considering a convergent subsequence if necessary, we can define

$$b_i^\star:=\lim_{k\to\infty}b_i^{N_k,\tau_k}\leq\Xi,\quad i=1,\dots,n. \tag{7.1.20}$$

Then by weak convergence and (7.1.17) we obtain $\tau(\boldsymbol{b}^\star)\wedge\Xi\in\mathcal{T}_\Xi^\star$.

Let us also show uniqueness of the $b_i^\star$'s and its independence of Ξ. Assume there exists another $\bar{\boldsymbol{b}}\in[0,\Xi^n]$ such that $\tau(\bar{\boldsymbol{b}})\wedge\Xi\in\mathcal{T}_\Xi^\star$. Let ℓ and $\bar\ell$ be such that $\min_{i=1,\dots,n}b_i^\star=b_\ell^\star$ and $\min_{i=1,\dots,n}\bar b_i=\bar b_{\bar\ell}$. For simplicity we present the argument only in the case where ℓ and $\bar\ell$ are unique and where every K_i is in the support of $\mathcal{L}(X_{\tau(\boldsymbol{b}^\star)\wedge\Xi})$. Then, if $\ell\neq\bar\ell$ it follows that either $\tau(\boldsymbol{b}^\star)\wedge\Xi$ "misprices" the put with strike $K_{\bar\ell}$ or $\tau(\bar{\boldsymbol{b}})\wedge\Xi$ "misprices" the put with strike K_ℓ. If $\ell=\bar\ell$ but $b_\ell^\star\neq\bar b_{\bar\ell}$ then it follows that either $\tau(\boldsymbol{b}^\star)\wedge\Xi$ or $\tau(\bar{\boldsymbol{b}})\wedge\Xi$ "misprices" the put with strike K_ℓ. Together this yields $\ell=\bar\ell$ and $b_\ell^\star=\bar b_{\bar\ell}$. With the same argument one shows iteratively that $b_i^\star=\bar b_i$ for all $i=1,\dots,n$.

Hence, continuing the estimation in (7.1.19),

$$\sup_{\tau\in\mathcal{T}_\Xi^\star}\lim_{N\to\infty}\mathbb{E}\left[G\left(X_{\tau_R\left(\mu_{N,\tau}^\star\right)\wedge\Xi}\right)\right]=\mathbb{E}\left[G\left(X_{\tau(\boldsymbol{b}^\star)\wedge\Xi}\right)\right]\leq\sup_{\tau\in\mathcal{T}_\Xi^\star}\mathbb{E}\left[G\left(X_{\tau\wedge\Xi}\right)\right].$$

The claim follows. □

Remark 7.1.9 (Characterization Bounded Embeddings). Theorem 7.1.8 in particular implies that

$$\mathcal{T}_\Xi^\star\neq\emptyset\qquad\Longleftrightarrow\qquad\exists\boldsymbol{b}^\star\in\mathbb{R}^n\ :\ \tau(\boldsymbol{b}^\star)\wedge\Xi\in\mathcal{T}_\Xi^\star. \tag{7.1.21}$$

Remark 7.1.10 (Interpretation of $\boldsymbol{b}^\star$). The vector $\boldsymbol{b}^\star$ from Theorem 7.1.8 is related to the annualized Black-Scholes implied volatilities $\sigma_{\text{impl}}(K_i)$ of the option prices $p^\star(K_i),\ i = 1, \ldots, n$. With T being the maturity in years, we have by our time-change approach, see Section 6.4 that

$$\min_{i=1,\ldots,n} T\sigma^2_{\text{impl}}(K_i) = \min_{i=1,\ldots,n} b_i^\star, \qquad\qquad T\sigma^2_{\text{impl}}(K_i) \leq b_i^\star. \tag{7.1.22}$$

Remark 7.1.11 (Case $b_i = \Xi$). If $b_i = \Xi$ then we can discard the option with strike K_i from the robust pricing problem. Indeed, according to Theorem 7.1.8, the robust pricing problem with puts $p^\star(K_1), \ldots, p^\star(K_{i-1}),\ p^\star(K_{i+1}), \ldots, p^\star(K_n)$ is solved by the vector $\tilde{\boldsymbol{b}} = (b_1, \ldots, b_{i-1},\ b_{i+1}, \ldots, b_n)$ and we have $\tau(\tilde{\boldsymbol{b}}) \wedge \Xi = \tau(\boldsymbol{b}) \wedge \Xi$.

7.1.3 Direct Extensions to Non-Vanilla Payoffs

Option on Stock and its Realized Variance Assume G satisfies

$$\frac{1}{2}x^2\frac{\partial^2 G}{\partial x^2}(t,x) + \frac{\partial G}{\partial t}(t,x) = g(x) \tag{7.1.23}$$

for some non-negative function g. Writing

$$\mathbb{E}\left[G(\tau \wedge \Xi, X_{\tau\wedge\Xi})\right] = G(0, X_0) + \mathbb{E}\left[\int_0^{\tau\wedge\Xi} g(X_u)\mathrm{d}u\right], \tag{7.1.24}$$

analogous arguments as in the proof of Proposition 7.1.4 and Theorem 7.1.8 yield

$$\sup_{\tau\in\mathcal{T}_\Xi^\star} \mathbb{E}\left[G(\tau, X_\tau)\right] = \mathbb{E}\left[G\left(\tau(\boldsymbol{b}^\star) \wedge \Xi, X_{\tau(\boldsymbol{b}^\star)\wedge\Xi}\right)\right]. \tag{7.1.25}$$

Options on Realized Variance If $g \equiv 1$, the r.h.s. in (7.1.24) becomes $\mathbb{E}\left[\tau \wedge \Xi\right]$ which corresponds to the price of a capped variance swap, i.e. of an option paying $\langle \log(S) \rangle_T \wedge \Xi$ at maturity T. More generally, we have by optimality of the Root stopping time and by disintegrating the concave function H,

$$\sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}\left[H(\tau \wedge \Xi)\right] = \mathbb{E}\left[H(\tau(\boldsymbol{b}^\star) \wedge \Xi)\right]. \tag{7.1.26}$$

Similar results have already been obtained by Cox and Wang [24] but in a different context with a slightly different Mathematical Finance interpretation.

Option on the Squared Difference of Running Maximum and Spot Let $M_t := \sup_{u \le t} X_u$. Consider the payoff $G(X_\tau, M_\tau)$ for a bivariate function G of the form $G(x, m) = (m - x)^2 g(m)$ for some non-negative g. Robust pricing and hedging without constraints on realized variance of payoffs G has been investigated by Hobson and Klimmek [54].

Note $G(X_0, M_0) = 0$ by $X_0 = M_0$. We have

$$G(X_{\tau\wedge\Xi}, M_{\tau\wedge\Xi}) = \int_0^{\tau\wedge\Xi} g(M_u)\mathrm{d}u - \int_0^{\tau\wedge\Xi} 2(M_u - X_u) g(M_u)\mathrm{d}X_u,$$

and hence

$$\mathbb{E}\left[G(X_{\tau\wedge\Xi}, M_{\tau\wedge\Xi})\right] = \mathbb{E}\left[\int_0^{\tau\wedge\Xi} g(M_u)\mathrm{d}u\right] = \mathbb{E}\left[\int_0^{\tau\wedge\Xi} \tilde{g}(X_u, M_u)\mathrm{d}u\right] \tag{7.1.27}$$

for some $\tilde{g}$. Therefore, by Rost [86] the optimal $\tau \in \mathcal{T}_\Xi^\star$ for (7.1.27) can be chosen as a Root embedding for a certain bivariate measure.

If we let $g \equiv 1$ then (7.1.27) reads

$$\mathbb{E}\left[(M_{\tau\wedge\Xi} - X_{\tau\wedge\Xi})^2\right] = \mathbb{E}\left[\int_0^{\tau\wedge\Xi} 1\mathrm{d}u\right] = \mathbb{E}\left[\tau \wedge \Xi\right] \tag{7.1.28}$$

and the Root stopping time $\tau(\boldsymbol{b}^\star) \wedge \Xi$ maximizes (7.1.28) over the set $\mathcal{T}_\Xi^\star$.

7.2 Robust Hedging and Duality

In this section we state and prove our duality result for convex Vanilla options. We also characterize the optimal static position $\boldsymbol{\lambda}^\star \in \mathbb{R}^n$. The key part of the proof is the construction of the cheapest pathwise superhedging strategy for payoff $G(S_T)$ which works on $\mathfrak{P}_\Xi$. It turns out that this strategy has a dynamic component. Dynamic trading using a pathwise approach has been considered by Bick and Willinger [12] and Davis et al. [32].

For $\boldsymbol{\lambda} \in \mathbb{R}^n$ define the *penalized payoff*

$$G^{\boldsymbol{\lambda}}(x) := G(x) - \sum_{i=1}^{n} \lambda_i \left(K_i - x\right)^+. \tag{7.2.1}$$

A version of the following result has been proven by Mykland [67].

Theorem 7.2.1 (Duality and Optimal Static Position)**.** *Let G satisfy Assumption* (iii) *on page 137. Recall the formalism of Section 6.3. Then,*

$$\sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}\left[G\left(X_\tau\right)\right] = \inf_{\boldsymbol{\lambda} \in \mathbb{R}^n} \sup_{\tau \in \mathcal{T}_\Xi} \left\{ \mathbb{E}\left[G^{\boldsymbol{\lambda}}\left(X_\tau\right)\right] + \sum_{i=1}^{n} \lambda_i p^\star(K_i) \right\} \tag{7.2.2}$$

and there is no duality gap, i.e.

$$\mathbb{A}(G(S_T); \mathfrak{P}_\Xi, \mathcal{A}, \mathcal{P}) = \sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}\left[G\left(X_\tau\right)\right] = \sup_{\mathbb{P}^S} \mathbb{E}_{\mathbb{P}^S}\left[G(S_T)\right] \tag{7.2.3}$$

where the last supremum is over all $(\mathfrak{P}_\Xi, \mathcal{Y}, \mathcal{P})$-market models $\left(\Omega^S, \mathcal{F}^S, \mathbb{F}^S, \mathbb{P}^S\right)$.

Assume $\mathcal{T}_\Xi^\star \neq \emptyset$ and recall $\boldsymbol{b}^\star$ from Theorem 7.1.8. Then there is existence in both problems in (7.2.2) *and the optimal static position*[3] $\boldsymbol{\lambda}^\star \in \mathbb{R}$ *is uniquely*

[3] The financial interpretation of $\boldsymbol{\lambda}$ was given in Remark 6.3.3.

determined by the linear system

$$\frac{\partial \mathbb{E}\left[G\left(X_{\tau(\boldsymbol{b}^\star)\wedge\Xi}\right)\right]}{\partial b_i} = \sum_{j=1}^{n} \lambda_j^\star \frac{\partial \mathbb{E}\left[\left(K_j - X_{\tau(\boldsymbol{b}^\star)\wedge\Xi}\right)^+\right]}{\partial b_i}, \quad i = 1, \ldots, n. \tag{7.2.4}$$

The proof of this result will be developed in the next sections.

7.2.1 An Optimal Stopping Problem

Let $\mathcal{T}_{t,\Xi}$ denote the set of stopping times τ such that $t \leq \tau \leq \Xi$. We write

$$V^{\boldsymbol{\lambda}}(t,x) := \sup_{\tau\in\mathcal{T}_{t,\Xi}} \mathbb{E}\left[G^{\boldsymbol{\lambda}}(X_\tau) \;\middle|\; X_t = x\right] \tag{7.2.5}$$

for the Snell envelope of $\left(G^{\boldsymbol{\lambda}}(X_t)\right)_{t\leq\Xi}$. Equation (7.2.5) is a standard optimal stopping problem. The stopping time

$$\tau_t^\star = \inf\left\{u > t : V^{\boldsymbol{\lambda}}(u, X_u) = G^{\boldsymbol{\lambda}}(X_u)\right\} \wedge \Xi. \tag{7.2.6}$$

is optimal for (7.2.5), see e.g. Peskir and Shiryaev [81].

Because $V^{\boldsymbol{\lambda}}$ is non-increasing in time, $\tau^\star := \tau_0^\star$ must be a Root stopping time. In addition, we conclude with Theorem 7.1.8 (or directly by the piecewise convexity of $G^{\boldsymbol{\lambda}}$) that

$$\tau_t^\star = \inf\left\{u > t \;:\; X_u = K_i,\ u \geq b_i \text{ for some } i = 1, \ldots, n\right\} \wedge \Xi \tag{7.2.7}$$

for some $\boldsymbol{b} \in \mathbb{R}^n$. Therefore, the continuation region for the optimal stopping problem (7.2.5) can be chosen as[4]

$$D^{\boldsymbol{\lambda}} = \Big\{(t,x) \in [0,\Xi) \times \mathbb{R}_+ \ : \ (t,x) \notin \bigcup_{i=1}^{n} \big([b_i,\infty) \times \{K_i\}\big)\Big\}. \tag{7.2.8}$$

Lemma 7.2.2 (Properties of $V^{\boldsymbol{\lambda}}$)**.** *The value function $V^{\boldsymbol{\lambda}}$ is continuous on $\bar{D}^{\boldsymbol{\lambda}} = [0,\Xi] \times \mathbb{R}_+$ and $V^{\boldsymbol{\lambda}} \in \mathbb{C}^{1,2}(D^{\boldsymbol{\lambda}})$.*

Further, for $(t,x) \in D^{\boldsymbol{\lambda}}$,

$$\frac{\partial V^{\boldsymbol{\lambda}}}{\partial t}\big(t,x\big) \leq 0, \qquad \frac{\partial^2 V^{\boldsymbol{\lambda}}}{\partial x^2}\big(t,x\big) \geq 0, \tag{7.2.9a}$$

$$\frac{\partial V^{\boldsymbol{\lambda}}}{\partial t}\big(t,x\big) + \frac{1}{2}x^2\frac{\partial^2 V^{\boldsymbol{\lambda}}}{\partial x^2}\big(t,x\big) = 0. \tag{7.2.9b}$$

Moreover,

$$\left|\frac{\partial V^{\boldsymbol{\lambda}}}{\partial x}\big(t,K_i\pm\big)\right| < \infty \qquad \text{for } t \geq b_i, \quad i = 1,\ldots,n. \tag{7.2.10}$$

Proof. The proof follows closely the arguments outlined by Peskir and Shiryaev [81, Chapter III, Section 7]. Since $(t,X_t)_{t\geq 0}$ and $G^{\boldsymbol{\lambda}}$ are continuous, continuity of $V^{\boldsymbol{\lambda}}$ is clear.

Take $(t,x) \in D^{\boldsymbol{\lambda}}$ and let $\epsilon > 0$ small enough such that $B_\epsilon(t,x) := \{(s,y) : |s-t|^2 + |y-x|^2 < \epsilon^2\} \subseteq D^{\boldsymbol{\lambda}}$. Classical PDE results, see e.g. Friedman [43, p. 134, Theorem 2.4], give existence of a smooth solution f to the Dirichlet problem

$$\frac{\partial f}{\partial t} + \frac{1}{2}x^2\frac{\partial^2 f}{\partial x^2} = 0 \qquad \text{in } B_\epsilon(t,x), \tag{7.2.11a}$$

$$f\big|_{\partial B_\epsilon(t,x)} = V^{\boldsymbol{\lambda}}. \tag{7.2.11b}$$

[4] This is not exactly true in the bounded support case. However, by continuity of paths, this does not make a difference because the other parts of the time-space domain are never visited.

Then by Itô's formula, the optional sampling theorem and properties of the Snell envelope $V^{\boldsymbol{\lambda}}$,

$$\begin{aligned}
f(t,x) &= \mathbb{E}\left[V^{\boldsymbol{\lambda}}\left(\tau_{B_\epsilon(t,x)}, X_{\tau_{B_\epsilon(t,x)}}\right)\Big|X_t = x\right] = \mathbb{E}\left[V^{\boldsymbol{\lambda}}\left(\tau_{D^{\boldsymbol{\lambda}}}, X_{\tau_{D^{\boldsymbol{\lambda}}}}\right)\Big|X_t = x\right] \\
&= \mathbb{E}\left[G^{\boldsymbol{\lambda}}\left(X_{\tau_{D^{\boldsymbol{\lambda}}}}\right)\Big|X_t = x\right] = V^{\boldsymbol{\lambda}}(t,x)
\end{aligned}$$

proving that $V^{\boldsymbol{\lambda}}$ is smooth at (t,x).

Since $V^{\boldsymbol{\lambda}}$ is non-increasing in time and $\left(V^{\boldsymbol{\lambda}}(t,X_t)\right)_{t\geq 0}$ is a martingale for $(t,X_t)\in D^{\boldsymbol{\lambda}}$ we deduce (7.2.9a)–(7.2.9b).

Now we argue (7.2.10). If $t \geq b_i$ then $V^{\boldsymbol{\lambda}}(t,K_i) = G^{\boldsymbol{\lambda}}(K_i)$, and it is generally true by definition that $V^{\boldsymbol{\lambda}} \geq G^{\boldsymbol{\lambda}}$. Therefore, by (7.2.9a) and Assumption (iii) we get

$$\begin{aligned}
-\infty < \frac{\partial V^{\boldsymbol{\lambda}}}{\partial x}(t,K_i-) \leq \frac{\partial G^{\boldsymbol{\lambda}}}{\partial x}(K_i-) < \infty, \\
-\infty < \frac{\partial G^{\boldsymbol{\lambda}}}{\partial x}(K_i+) \leq \frac{\partial V^{\boldsymbol{\lambda}}}{\partial x}(t,K_i+) < \infty.
\end{aligned}$$

□

7.2.2 Pathwise Superhedging Strategy

In this section we show how to construct a pathwise superhedging strategy which dominates $G^{\boldsymbol{\lambda}}(X_t)$ for all $t \leq \Xi$. It turns out that we can always do this with cost $V^{\boldsymbol{\lambda}}(0,X_0)+\delta$ for arbitrarily small $\delta > 0$. Suitable time-changing will yield a pathwise superhedging strategy for $G(S_T)$ on $\mathfrak{P}_\Xi$.

In the following, all integrals w.r.t. X or S are understood pathwise.

Proposition 7.2.3 (Pathwise Superhedging Strategy)**.** *Let* $\boldsymbol{\lambda}\in\mathbb{R}^n$ *and recall the definition of* $V^{\boldsymbol{\lambda}}$ *in* (7.2.5).

Then, for all $\delta > 0$ *there exists a progressively measurable* $\tilde{\phi}$ *such that for all* $t \leq \Xi$ *and* $X \in \Omega^{\mathrm{GBM}}$*:*

$$\delta + V^{\boldsymbol{\lambda}}(0, X_0) + \int_0^t \tilde{\phi}_u(X) \mathrm{d}X_u \geq V^{\boldsymbol{\lambda}}(t, X_t) \geq G^{\boldsymbol{\lambda}}(X_t) \tag{7.2.12}$$

is well defined pathwise. Setting

$$\phi_u(S) := \tilde{\phi}_{\langle \log(S) \rangle_u}(S) \tag{7.2.13}$$

yields for all $S \in \mathfrak{P}_\Xi$*:*

$$\delta + V^{\boldsymbol{\lambda}}(0, X_0) + \int_0^T \phi_u(S) \mathrm{d}S_u \geq G^{\boldsymbol{\lambda}}(S_T) \tag{7.2.14}$$

and $\phi \in \Phi$*, defined in* (6.3.6)*.*

Moreover, if $t = \tau_{D^{\boldsymbol{\lambda}}}$ *for* $D^{\boldsymbol{\lambda}}$ *from* (7.2.8) *then there is equality in* (7.2.12) *with* $\delta = 0$*.*

Proof. By the Itô-Föllmer formula we have by (7.2.9b),

$$V^{\boldsymbol{\lambda}}(0, X_0) + \int_0^{t \wedge \tau_{D^{\boldsymbol{\lambda}}}} \frac{\partial V^{\boldsymbol{\lambda}}}{\partial x}\big(u, X_u\big) \mathrm{d}X_u = V^{\boldsymbol{\lambda}}(t \wedge \tau_{D^{\boldsymbol{\lambda}}}, X_{t \wedge \tau_{D^{\boldsymbol{\lambda}}}}) \tag{7.2.15}$$

and we can take $\tilde{\phi}_u(X) := \frac{\partial V^{\boldsymbol{\lambda}}}{\partial x}(u, X_u) \mathbb{1}_{\{u < \tau_{D^{\boldsymbol{\lambda}}}\}}$. This is the pathwise superhedging property until time $\tau_{D^{\boldsymbol{\lambda}}}$. It also proves the claim regarding equality in (7.2.12) for $t = \tau_{D^{\boldsymbol{\lambda}}}$.

Next we find a pathwise superhedging strategy after time $\tau_{D^{\boldsymbol{\lambda}}}$. Let $u < b_i$. By continuity of $V^{\boldsymbol{\lambda}}$ we have $V^{\boldsymbol{\lambda}}(b_i, \cdot) = \lim_{u \uparrow b_i} V^{\boldsymbol{\lambda}}(u, \cdot)$. Therefore, for $b_i > 0$ the value function $V^{\boldsymbol{\lambda}}(b_i, \cdot)$ is convex in a neighbourhood of K_i by convexity of $V^{\boldsymbol{\lambda}}(u, \cdot),\ u < b_i$, in a neighbourhood of K_i.

For $a, b \in \mathbb{R}$, denote $[\![a, b]\!] = [a \wedge b, a \vee b]$. Take $t \geq b_i$ such that $X_t = K_i$. Then clearly $V^{\boldsymbol{\lambda}}(t, K_i) = G^{\boldsymbol{\lambda}}(K_i)$. We claim that we can choose

$$\tilde{\phi}_t^{\pm} \in \left[\!\!\left[\frac{\partial V^{\boldsymbol{\lambda}}}{\partial x}(t, K_i-), \frac{\partial V^{\boldsymbol{\lambda}}}{\partial x}(t, K_i+)\right]\!\!\right] \tag{7.2.16}$$

such that for every $\delta > 0$ there exists $\epsilon > 0$ such that

$$\delta + V^{\boldsymbol{\lambda}}(t, X_t) + \tilde{\phi}_t^{\pm} \cdot (X_s - X_t) \geq V^{\boldsymbol{\lambda}}(s, X_s) \tag{7.2.17}$$

for all $(s, X_s) \in [t, \Xi] \times [K_i - \epsilon, K_i + \epsilon]$.

In fact, for $t > b_i$ we now argue that we can take $\delta = 0$. Observe that by the supermartingale property of $V^{\boldsymbol{\lambda}}$ we have

$$G^{\boldsymbol{\lambda}}(K_i) = V^{\boldsymbol{\lambda}}(b_i, K_i) \geq \mathbb{E}\left[V^{\boldsymbol{\lambda}}(\rho, X_\rho) \,\middle|\, X_{b_i} = K_i\right]$$

where $\rho := H_{K_i - \epsilon, K_i + \epsilon} \wedge t$. Now if $V^{\boldsymbol{\lambda}}(t, \cdot)$ was convex in a neighbourhood of K_i, then by (7.2.9a) also $V^{\boldsymbol{\lambda}}(u, \cdot)$, $b_i \leq u \leq t$, would be convex in a neighbourhood of K_i and hence

$$\mathbb{E}\left[V^{\boldsymbol{\lambda}}(\rho, X_\rho) \,\middle|\, X_{b_i} = K_i\right] \geq G^{\boldsymbol{\lambda}}(K_i)$$

and there is equality only if $V^{\boldsymbol{\lambda}}(u, \cdot)$ is linear on $[K_i - \epsilon, K_i + \epsilon]$ for all $u \in [b_i, t]$. It follows that $V^{\boldsymbol{\lambda}}(t, \cdot)$ is either linear on $[K_i - \epsilon, K_i + \epsilon]$ or not convex at K_i and the same holds for $V^{\boldsymbol{\lambda}}(t', \cdot)$ for $t' \geq t$. It is now clear that there exists $\tilde{\phi}_t^{\pm}$ which satisfies (7.2.17) with $\delta = 0$.

When $t = b_i$ it is clear that given $\delta > 0$ we can find $\epsilon > 0$ to satisfy (7.2.17).

In summary, the above argument shows that we can define a pathwise superheding strategy from time $t = \tau_{D^{\boldsymbol{\lambda}}}$ until the time first time s after t such that $X_s \notin (K_i - \epsilon, K + \epsilon)$. It is a buy and hold strategy given by $\tilde{\phi}_u \equiv \tilde{\phi}_t^{\pm}$ for $t \leq u < s$.

However, at time s, by (7.2.17), we have enough capital to switch back to the trading strategy $\tilde{\phi}_u(X) := \frac{\partial V^{\boldsymbol{\lambda}}}{\partial x}(u, X_u)$. By the same arguments as above, this strategy will superhedge until the path leaves $D^{\boldsymbol{\lambda}}$ again. It is clear that we can iterate these arguments to obtain a pathwise superhedging strategy $\tilde{\phi}$ as claimed in (7.2.12).

Recalling (7.2.13) we now argue that (7.2.14) makes sense pathwise for $S \in \mathfrak{P}_\Xi$. By definition of $\boldsymbol{\pi}$ in (6.2.1) we observe the fact $\sup_k |\pi^m_{k+1} - \pi^m_k| \to 0$ as $m \to \infty$. Then by $\langle \log(S) \rangle$ being non-decreasing we infer that $\langle \log(S) \rangle$ has zero quadratic variation in the sense of Definition 6.2.1, i.e. in particular the quadratic variation of $\langle \log(S) \rangle$ is well-defined. In that case the multi-dimensional Itô-Föllmer formula given in Föllmer [40, eq. (12)] holds for $Y = (S, \langle \log(S) \rangle)$ and yields for $\langle \log(S) \rangle_t < \tau_{D^{\boldsymbol{\lambda}}}$, (we use the same short-hand notation as in classical stochastic calculus)

$$
\begin{aligned}
& V^{\boldsymbol{\lambda}}\big(\langle \log(S) \rangle_t, S_t\big) - V^{\boldsymbol{\lambda}}(0, S_0) \\
= \quad & \int_0^t \frac{\partial V^{\boldsymbol{\lambda}}}{\partial x}\Big(\langle \log(S) \rangle_u, S_u\Big) \mathrm{d}S_u + \int_0^t \frac{\partial V^{\boldsymbol{\lambda}}}{\partial t}\Big(\langle \log(S) \rangle_u, S_u\Big) \mathrm{d}\langle \log(S) \rangle_u \\
& + \frac{1}{2} \int_0^t \frac{\partial^2 V^{\boldsymbol{\lambda}}}{\partial x^2}\Big(\langle \log(S) \rangle_u, S_u\Big) \mathrm{d}\langle S \rangle_u.
\end{aligned}
$$

Note that for $S \in \mathfrak{P}_\Xi$ Föllmer [40, eq. (14)] yields $S_t^2 \mathrm{d}\langle \log(S) \rangle_t = \mathrm{d}\langle S \rangle_t$ and then we obtain for $\langle \log(S) \rangle_t < \tau_{D^{\boldsymbol{\lambda}}}$ by (7.2.9b)

$$
\frac{\partial V^{\boldsymbol{\lambda}}}{\partial t}\Big(\langle \log(S) \rangle_t, S_t\Big) \mathrm{d}\langle \log(S) \rangle_t + \frac{1}{2} \frac{\partial^2 V^{\boldsymbol{\lambda}}}{\partial x^2}\Big(\langle \log(S) \rangle_t, S_t\Big) \mathrm{d}\langle S \rangle_t = 0. \tag{7.2.18}
$$

Together, we compute for $\langle \log(S) \rangle_t \leq \tau_{D^{\boldsymbol{\lambda}}}$,

$$
V^{\boldsymbol{\lambda}}\big(\langle \log(S) \rangle_t, S_t\big) - V^{\boldsymbol{\lambda}}(0, S_0) \overset{(7.2.18)}{=} \int_0^t \underbrace{\frac{\partial V^{\boldsymbol{\lambda}}}{\partial x}\Big(\langle \log(S) \rangle_u, S_u\Big)}_{=\phi_u(S)} \mathrm{d}S_u \tag{7.2.19}
$$

and in particular the stochastic integral on the r.h.s. of (7.2.19) is well defined. It is clear that when $\tilde{\phi}$ is constant in time and space on some rectangle then the integral in (7.2.14) is well defined. Together, this proves (7.2.14). The property $\phi \in \Phi$ can be deduced by using convexity of G. □

7.2.3 Proof of Theorem 7.2.1

We require one last regularity result in order to prove Theorem 7.2.1.

Lemma 7.2.4 (Smoothness in $\boldsymbol{b}$)**.** *Let G satisfy Assumption* (iii) *on page 137. Then the mapping*

$$[0,\Xi]^n \to \mathbb{R}, \quad \boldsymbol{b} \mapsto \mathbb{E}\left[G^{\boldsymbol{\lambda}}(X_{\tau(\boldsymbol{b})\wedge\Xi})\right] \tag{7.2.20}$$

is continuous and is further differentiable at any point $\boldsymbol{b}$ in the interior of $[0,\Xi]^n$.

Proof. Cox and Wang [24, Lemma 3.3] yields that the measure corresponding to $\mathcal{L}(X_t; t < \tau(\boldsymbol{b}) \wedge \Xi)$ has density $p_{D^{\boldsymbol{\lambda}}}$ with respect to the Lebesgue measure on $D^{\boldsymbol{\lambda}}$ and that the density is smooth. From this it follows that the transition density for Geometric Brownian motion absorbed at two points is smooth in time and space on the interior of the interval between the two points.

Denote $K_0 := 0$, $b_0 := 0$, $K_{n+1} := +\infty$, $b_{n+1} := 0$. Fix $i \in \{1,\dots,n\}$. Define $j_- := \max\{j \in \{0,\dots,n+1\} : b_j \le b_i,\ K_j < K_i\}$ and $j_+ := \min\{j \in \{0,\dots,n+1\} : b_j \le b_i,\ K_j > K_i\}$. Now set $K_- := K_{j_-}$, $K_+ := K_{j_+}$ and $\bar{b} := b_{j_-} \vee b_{j_+} \le b_i$. Note that $\bar{b} = \bar{b}(b_i)$ is piecewise constant and right-continuous in b_i.

Take $K \in \mathbb{R}_+$. Our first claim is that $b_i \mapsto \mathbb{E}\left[\left(K - X_{\tau(\boldsymbol{b})\wedge\Xi}\right)^+\right]$ is continuously differentiable if and only if the same holds for $b_i \mapsto \mathbb{E}\left[\left(K - X_{\tau(\boldsymbol{b})\wedge\Xi\wedge H_{K_-,K_+}}\right)^+\right]$ and in that case the derivatives coincide. If $\bar{b} < b_i$ this is trivial. If $\bar{b} = b_i$ then also trivially the right-derivatives coincide. Observe for $x < K$,

$$\begin{aligned}
\mathbb{E}\left[L_\epsilon^K \mid X_0 = x\right] &\leq \mathbb{P}\left[\sup_{u\leq\epsilon} X_u \geq K \,\middle|\, X_0 = x\right] \mathbb{E}\left[L_\epsilon^K \mid X_0 = K\right], \\
\lim_{\epsilon\downarrow 0} \frac{\mathbb{P}\left[\sup_{u\leq\epsilon} X_u \geq K \mid X_0 = x\right]}{\epsilon} &= 0
\end{aligned}$$

and similarly for $x > K$. From this we deduce together with (6.1.4) that also the left-derivatives coincide.

The previous argument allows to restrict to the case $\bar{b} < b_i$ because the other case follows in the same way. Then,

$$\begin{aligned}
\mathbb{E}\left[\left(K - X_{\tau(\boldsymbol{b})\wedge\Xi}\right)^+\right] = {} & \mathbb{E}\left[\left(K - X_{\tau(\boldsymbol{b})\wedge\Xi}\right)^+ \mathbb{1}_{\left\{X_{\tau(\boldsymbol{b})\wedge\bar{b}}\notin(K_-,K_+)\right\}}\right] \\
& + \mathbb{E}\left[\left(K - X_{\tau(\boldsymbol{b})\wedge\Xi}\right)^+ \mathbb{1}_{\left\{X_{\tau(\boldsymbol{b})\wedge\bar{b}}\in(K_-,K_+)\right\}}\right].
\end{aligned}$$

Observe that by choice of K_-, K_+ and $\bar{b}$ it follows that the first term on the r.h.s. is constant in b_i in a neighbourhood of b_i. Hence it remains to show that the second term is continuously differentiable in b_i. To this end, note that

$$\begin{aligned}
& \mathbb{E}\left[\left(K - X_{\tau(\boldsymbol{b})\wedge\Xi}\right)^+ \mathbb{1}_{\left\{X_{\tau(\boldsymbol{b})\wedge\bar{b}}\in(K_-,K_+)\right\}}\right] \\
= {} & \int_{(K_-,K_+)} \mathbb{P}\left[X_{\tau(\boldsymbol{b})\wedge\bar{b}} \in \mathrm{d}x\right] \Big\{ (K-K_-)^+ \mathbb{P}\left[X_{H_{K_-,K_+}\wedge b_i} = K_- \middle| X_{\bar{b}} = x\right] \\
& + (K-K_+)^+ \mathbb{P}\left[X_{H_{K_-,K_+}\wedge b_i} = K_+ \middle| X_{\bar{b}} = x\right] \\
& + \mathbb{E}\left[\left(K - X_{\tau(\boldsymbol{b})\wedge\Xi}\right)^+ \mathbb{1}_{\left\{X_{b_i}\in(K_-,K_+)\right\}} \middle| X_{\bar{b}} = x\right] \Big\}
\end{aligned}$$

By $b_i > \bar{b}$, the quantity $\mathbb{P}\left[X_{\tau(\boldsymbol{b})\wedge\bar{b}} \in \mathrm{d}x\right]$ does not depend on b_i. Further, clearly, the first two terms in the curly brackets are continuously differentiable in b_i. Hence, it remains to verify smoothness of the third term in the curly brackets. We write for $x \in (K_-, K_+)$,

$$\begin{aligned}&\mathbb{E}\left[\left(K - X_{\tau(\boldsymbol{b})\wedge\Xi}\right)^+ \mathbb{1}_{\{X_{b_i}\in(K_-,K_+)\}}\middle|X_{\bar{b}} = x\right]\\ &= \int_{(K_-,K_+)} \mathbb{P}\left[X_{b_i} \in \mathrm{d}y\middle|X_{\bar{b}} = x\right] \mathbb{E}\left[\left(K - X_{\tau(\boldsymbol{b})\wedge\Xi}\right)^+ \;\middle|\; X_{b_i} = y\right]\end{aligned}$$

The last expression is continuously differentiable in b_i by Lemma 7.2.2.

The claim for general G now follows from the integration by parts formula. □

Putting the results of the previous sections together we can now prove Theorem 7.2.1.

Proof of Theorem 7.2.1. For $\boldsymbol{\lambda} \in \mathbb{R}^n$ we get essentially by an application of Theorem 7.1.8,

$$\sup_{\tau\in\mathcal{T}_\Xi} \mathbb{E}\left[G^{\boldsymbol{\lambda}}(X_\tau)\right] = \sup_{\boldsymbol{b}\in[0,\Xi]^n} \mathbb{E}\left[G^{\boldsymbol{\lambda}}(X_{\tau(\boldsymbol{b})\wedge\Xi})\right]. \tag{7.2.21}$$

Note that existence of a maximizer is always guaranteed by the continuity of

$$[0,\Xi]^n \to \mathbb{R}, \quad \boldsymbol{b} \mapsto \mathbb{E}\left[G^{\boldsymbol{\lambda}}(X_{\tau(\boldsymbol{b})\wedge\Xi})\right],$$

cf. Lemma 7.2.4, and the compactness of $[0,\Xi]^n$.

In order to prove (7.2.2) we will apply Theorem B.1.2. To this end, we make the identification

$$\mathscr{X} = [0,\Xi]^n, \quad \mathscr{Y} = \mathbb{R}^n, \quad f(\boldsymbol{b},\boldsymbol{\lambda}) = \mathbb{E}\left[G^{\boldsymbol{\lambda}}\left(X_{\tau(\boldsymbol{b})\wedge\Xi}\right)\right] + \sum_{i=1}^n \lambda_i p^\star(K_i)$$

and verify the Theorem's assumptions. Clearly, $\mathscr{X}$ is compact and $f(\cdot,\boldsymbol{\lambda})$ is continuous on $\mathscr{X}$ for each $\boldsymbol{\lambda} \in \mathscr{Y}$. Now we check that f is concave-convex-like in the sense of Definition B.1.1. Property (i) holds essentially by Theorem 7.1.8:

Let $x_1 = \boldsymbol{b}^1, x_2 = \boldsymbol{b}^2$ and let α be an independent Bernoulli random variable with parameter a. The stopping time

$$\bar{\tau} := \left(\tau(\boldsymbol{b}^1) \wedge \Xi\right) \mathbb{1}_{\{\alpha=0\}} + \left(\tau(\boldsymbol{b}^2) \wedge \Xi\right) \mathbb{1}_{\{\alpha=1\}} \leq \Xi$$

has put option prices given by

$$a\mathbb{E}\left[\left(K_i - X_{\tau(\boldsymbol{b}^1)\wedge\Xi}\right)^+\right] + (1-a)\mathbb{E}\left[\left(K_i - X_{\tau(\boldsymbol{b}^2)\wedge\Xi}\right)^+\right].$$

Consequently, by Theorem 7.1.8 there exists $x_3 = \boldsymbol{b}^3$ such that $\tau(\boldsymbol{b}^3) \wedge \Xi$ embeds these puts and

$$\mathbb{E}\left[G\left(X_{\tau(\boldsymbol{b}^3)\wedge\Xi}\right)\right] \geq \mathbb{E}\left[G\left(X_{\bar{\tau}}\right)\right] = a\mathbb{E}\left[G\left(X_{\tau(\boldsymbol{b}^1)\wedge\Xi}\right)\right] + (1-a)\mathbb{E}\left[G\left(X_{\tau(\boldsymbol{b}^2)\wedge\Xi}\right)\right],$$

which implies property (i). Property (ii) holds by the linearity of $f(\boldsymbol{b}, \cdot)$.

Therefore, we can indeed interchange the inf and sup in (7.2.2) and then observe

$$\inf_{\boldsymbol{\lambda}\in\mathbb{R}^n}\left(\mathbb{E}\left[G^{\boldsymbol{\lambda}}\left(X_\tau\right)\right] + \sum_{i=1}^n \lambda_i p^\star(K_i)\right) = \begin{cases} \mathbb{E}\left[G(X_\tau)\right] & \text{if } \tau \in \mathcal{T}_\Xi^\star, \\ -\infty & \text{else.} \end{cases}$$

This establishes (7.2.2).

It will follow from our arguments that the superhedging price cannot be reduced by using put options $p^\star(K_i)$ where $b_i^\star = \Xi$ for hedging purposes, see also Remark 7.1.11. Similar considerations allow to exclude put options $p^\star(K_i)$ where $b_i^\star = 0$, see Remark 7.1.5. Therefore, we assume WLOG that $\boldsymbol{b}^\star \in (0, \Xi)^n$. Then, by

Lemma 7.2.4 the following system is well-defined,

$$\begin{aligned} & D_{\boldsymbol{b}}\mathbb{E}\left[G^{\boldsymbol{\lambda}}(X_{\tau(\boldsymbol{b}^\star)\wedge\Xi})\right] = \boldsymbol{0} \\ \iff \quad & \frac{\partial \mathbb{E}\left[G(X_{\tau(\boldsymbol{b}^\star)\wedge\Xi})\right]}{\partial b_i} = \sum_{j=1}^{n} \lambda_j \frac{\partial \mathbb{E}\left[\left(K_j - X_{\tau(\boldsymbol{b}^\star)\wedge\Xi}\right)^+\right]}{\partial b_i} \end{aligned} \tag{7.2.22}$$

where $i = 1, \ldots, n$. Further, system (7.2.22) has a unique solution $\boldsymbol{\lambda}^\star \in \mathbb{R}^n$ because

$$\frac{\partial \mathbb{E}\left[\left(K_j - X_{\tau(\boldsymbol{b}^\star)\wedge\Xi}\right)^+\right]}{\partial b_i} = 0 \qquad \text{for } i \neq j \text{ such that } b_i^\star \geq b_j^\star,$$

and

$$\frac{\partial \mathbb{E}\left[\left(K_i - X_{\tau(\boldsymbol{b}^\star)\wedge\Xi}\right)^+\right]}{\partial b_i} > 0 \qquad \text{for } i = 1, \ldots, n \text{ such that } 0 < b_i^\star < \Xi,$$

and hence (7.2.22) is a linear system of triangular form with non-zero diagonal. By choice of $\boldsymbol{\lambda}^\star$,

$$\mathbb{E}\left[G^{\boldsymbol{\lambda}^\star}(X_{\tau(\boldsymbol{b}^\star)\wedge\Xi})\right] = \sup_{\tau\in\mathcal{T}_\Xi} \mathbb{E}\left[G^{\boldsymbol{\lambda}^\star}(X_{\tau\wedge\Xi})\right] = V^{\boldsymbol{\lambda}^\star}(0, X_0). \tag{7.2.23}$$

Using Proposition 7.2.3, for all $\delta > 0$ we can find $\phi \in \Phi$ such that for all $S \in \mathfrak{P}_\Xi$:

$$\delta + V^{\boldsymbol{\lambda}^\star}(0, X_0) + \int_0^T \phi_u(S)\mathrm{d}S_u + \sum_{i=1}^{n} \lambda_i^\star \left(K_i - S_T\right)^+ \geq G(S_T). \tag{7.2.24}$$

We also conclude by Proposition 7.2.3 that the $(\mathfrak{P}_\Xi, \mathcal{Y}, \mathcal{P})$-market model obtained by $S^\star$, where

$$S^\star := \left(X_{\rho_t^\star}\right)_{t\in[0,T]}, \qquad \rho_t^\star = \langle X^\star\rangle_t, \qquad X^\star := \left(X_{\tau(\boldsymbol{b}^\star)\wedge t}\right)_{t\in[0,\Xi]}, \tag{7.2.25}$$

achieves equality in (7.2.24) with $\delta = 0$. This implies that there cannot exist a cheaper pathwise superhedging strategy, i.e.

$$\mathbb{A}\left(G(S_T); \mathfrak{P}_\Xi, \mathcal{A}, \mathcal{P}\right) = \sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}\left[G(X_\tau)\right].$$

Together with the time-change considerations of Section 6.4 this proves the duality statement (7.2.3) and establishes (7.2.4). □

7.3 Arbitrage

We investigate the forms of arbitrage which can arise in our robust framework by adopting the definition of arbitrage from Acciaio et al. [1], see also Davis and Hobson [33] and Cox and Obłój [28]. In particular, we ask the question whether this notion of arbitrage is suitable to characterize existence of market models.

7.3.1 Arbitrage on the Restricted Path Space

Definition 7.3.1 (Arbitrage Relative to $(\mathfrak{P}, \mathcal{A}, \mathcal{P})$)**.** *Let $\mathfrak{P}$ be a path space. Recall the definitions of $\mathbb{K}$, $\mathcal{A}$ and $\mathcal{P}$ in Section 6.3.*

We say the pricing operator $\mathcal{P}$ admits no arbitrage relative to $\mathfrak{P}$ and $\mathcal{A}$ if

$$\forall V \in \mathcal{A} \;:\; V > 0 \quad on\ \mathfrak{P} \qquad \Longrightarrow \qquad \mathcal{P}\left[V\right] > 0. \tag{7.3.1}$$

The following result is based on Mykland [67, Proposition 3] and shows that the notion of arbitrage relative to $(\mathfrak{P}, \mathcal{A}, \mathcal{P})$ is suitable to characterize existence of a market model.

Theorem 7.3.2 (Arbitrage Relative to $(\mathfrak{P}_\Xi, \mathcal{A}, \mathcal{P})$)**.** *Let $\mathbb{K} = \{K_1, \ldots, K_n\}$.*

Then, there is arbitrage relative to $(\mathfrak{P}_\Xi, \mathcal{A}, \mathcal{P})$ *if and only if* $\mathcal{T}^\star_\Xi = \emptyset$*.*

Now let $\mathcal{T}^\star_\Xi \neq \emptyset$*. Suppose* G *satisfies Assumption* (iii) *on page 137. We add the option* $G(S_T)$ *to* $\mathcal{A}$ *and extend* $\mathcal{P}\left[G(S_T)\right] := p_G$*.*

Then, there is arbitrage relative to $(\mathfrak{P}_\Xi, \mathcal{A}, \mathcal{P})$ *if* $p_G > \sup_{\tau \in \mathcal{T}^\star_\Xi} \mathbb{E}\left[G(X_\tau)\right]$*.*

Proof. Let us distinguish two cases: firstly when there exists a probability measure μ such that

$$\int (K_i - x)^+ \mu(\mathrm{d}x) = p^\star(K_i) \qquad \forall i = 1, \ldots, n \tag{7.3.2}$$

and secondly when there does not. Davis and Hobson [33] use the notion of weak arbitrage to distinguish between these two cases.

When (7.3.2) holds for some μ, we have seen in Section 7.1.1 that there exists $\tau(\boldsymbol{b}^\star) \in \mathcal{T}^\star_\infty$ for some $\boldsymbol{b}^\star \in \mathbb{R}^n$. Further, we recall from Remark 7.1.9,

$$\mathcal{T}^\star_\Xi \neq \emptyset \qquad \Longleftrightarrow \qquad \exists \boldsymbol{b}^\star \in [0, \Xi]^n \ : \ \tau(\boldsymbol{b}^\star) \wedge \Xi \in \mathcal{T}^\star_\Xi.$$

If $\max_{i=1,\ldots,n} b^\star_i \leq \Xi$ then the Root embedding $\tau(\boldsymbol{b}^\star) \wedge \Xi$ provides a $(\mathfrak{P}_\Xi, \mathcal{Y}, \mathcal{P})$-market model and consequently there is no arbitrage in the sense of Definition 7.3.1.

Now we construct an arbitrage in the case of $\max_{i=1,\ldots,n} b^\star_i > \Xi$. Let $\mathcal{S}$ be the indices for which the $b^\star_i$'s are less or equal to Ξ and let I be the index corresponding to the smallest $b^\star_i$ strictly greater than Ξ (if it is non-unique choose one). Considering only the puts $p^\star(K_i)$ where $i \in \mathcal{S}$, Theorem 7.1.4 gives the maximal price $\bar{p}$ for the put option with strike K_I. We conclude that the observed market price for the put with strike K_I is strictly larger than $\bar{p}$ and that there does not exist a $(\mathfrak{P}_\Xi, \mathcal{Y}, \mathcal{P})$-market model in which the price for the put with strike K_I is reproduced. Hence the arbitrage consists in selling the put with strike K_I for its market price and setting up a superhedge for $(K_I - S_T)^+$. From Proposition 7.2.3 it follows that we can superreplicate $(K_I - S_T)^+$ with cost $\bar{p}+\delta$ and by the previous argument there exists $\delta > 0$ small enough such that $p^\star(K_I) - (\bar{p} + \delta) > 0$. This arbitrage is guaranteed to realize whenever $S \in \mathfrak{P}_\Xi$.

Now assume that (7.3.2) does not hold for any probability measure μ. If there is arbitrage relative to $(\mathbb{C}([0,T];\mathbb{R}_+, S_0), \mathcal{Y}, \mathcal{P})$, i.e. when no dynamic trading is allowed[5], it is clear how to realize the arbitrage. Following Davis and Hobson [33] it remains to consider the case when

$$c^\star(K_{n-1}) = c^\star(K_n) > 0. \tag{7.3.3}$$

Remove the option with strike K_n. Now we argue analogously to the first case: if $\max_{i=1,\ldots,n-1} b_i^\star > \Xi$ we showed how to realize an arbitrage on $\mathfrak{P}_\Xi$. If $\max_{i=1,\ldots,n-1} b_i^\star \leq \Xi$ we sell the call with strike K_n for $c^\star(K_n) = c^\star(K_{n-1}) > 0$ and setup a superhedge using trading and static positions in the options with strikes $K_1, \ldots, K_{n-1}$ (cf. Proposition 7.2.3). It is clear that this superhedging cost can be chosen strictly smaller than $c^\star(K_n)$ and hence the arbitrage is constructed.

Finally, if $\mathcal{T}_\Xi^\star \neq \emptyset$ and $p_G > \sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}[G(X_\tau)]$ then the arbitrage consists in selling the option $G(S_T)$ for price p_G and superreplicating it with cost $\sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}[G(X_\tau)] + \delta < p_G$. □

7.3.2 Remarks

Dichotomy As seen in Theorem 7.3.2, on the restricted path space $\mathfrak{P}_\Xi$ and in the case where $\mathcal{A}$ includes only the n put options and the dynamic trading strategies in the stock, there is dichotomy in the sense that either there is a market model or we can construct an arbitrage relative to $(\mathfrak{P}, \mathcal{A}, \mathcal{P})$.

[5]This situation is commonly referred to in the literature as model-independent arbitrage, cf. Davis and Hobson [33]; Cox and Obłój [28].

Consistent Beliefs Model-independent and weak arbitrage in the sense of Davis and Hobson [33] or Cox and Obłój [28] aside, if one does not want to challenge the market prices, it makes sense to impose

$$\Xi \geq \max_{i=1,\dots,n} b_i^\star \tag{7.3.4}$$

to make the *belief* about realized variance consistent with actually observed market prices.

If market prices turn out to be inconsistent with the belief, an arbitrage strategy can be setup as explained in Section 7.3.

Lower Bound on Realized Variance At first sight it might seem interesting to redo the analysis of robust pricing, hedging and arbitrage with both lower and upper bounds on realized variance, i.e. $\Xi_- \leq \langle \log(S) \rangle_T \leq \Xi_+$.

However, it is straightforward to see that there exists an embedding $\tau \in \mathcal{T}_\infty^\star$ such that $\tau \geq \Xi_-$ if and only if there is $\boldsymbol{b}^\star \in [\Xi_-, \infty]^n$ such that $\tau(\boldsymbol{b}^\star) \in \mathcal{T}_\infty^\star$. Indeed, if $\mathbb{E}\left[\left(K_i - X_{\Xi_-}\right)^+\right] > p^\star(K_i)$ there cannot exist a $\tau \in \mathcal{T}_\infty^\star$ such that $\tau \geq \Xi_-$. In the other case we can clearly use the Root embedding $\tau(\boldsymbol{b}^\star)$. Then, with the previous Remark in mind, it is reasonable to assume that $\min_{i=1,\dots,n} b_i^\star \geq \Xi_-$ and therefore a lower bound on realized variance would not affect the analysis for convex Vanilla options.

Lower Bound Analysis An open problem is to repeat the analysis to obtain robust lower price bounds and subhedging strategies. In this context the Rost embedding, see Obłój [73, Section 7.3], first comes to mind. However, the embedding requirement under realized variance constraints appears to be harder in this case. We leave this question for future research.

Perfect Hedge vs. Asymptotically Perfect Hedge in Worst-Case Model

In Section 7.1 we showed that the upper price bound is attained by a Root stopping time, both in the case of a risk-neutral distribution of S_T with bounded and unbounded support. More precisely, we proved that

$$\mathcal{L}\left(X_{\tau_R\left(\mu_M^\star\right)\wedge\Xi}\right) \longrightarrow \mathcal{L}\left(X_{\tau(\boldsymbol{b}^\star)\wedge\Xi}\right) \qquad \text{weakly as } M\to\infty \tag{7.3.5}$$

for some $\boldsymbol{b}^\star \in \mathbb{R}^n$ and $\Xi < \infty$. Note that $\left(X_{\tau(\boldsymbol{b}^\star)\wedge\Xi\wedge t}\right)_{t\geq 0}$ is uniformly integrable. This is the reason why we have a perfect hedge in one model. The sequence $(\mu_M^\star)_M$ does not necessarily converge weakly and consequently there is only an "asymptotically perfect hedge" in the case of $\Xi = \infty$. This situation corresponds to weak arbitrage, cf. Davis and Hobson [33].

Chapter 8

Lookback Options with Beliefs about Realized Variance

The goal of this chapter is to extend the results of Chapter 7 to certain Lookback options $G\left(\sup_{t\leq T} S_t\right)$. We denote

$$M_t := \sup_{0\leq u\leq t} X_u, \qquad \bar{S}_t := \sup_{0\leq u\leq t} S_u. \tag{8.0.1}$$

If not specified further, we impose that the payoff G is right-continuous, non-decreasing and $\mathbb{E}\left[G(M_\Xi)\right] < \infty$. Since these payoffs are invariant under time-change our approach via time-change can be used to analyse the problem.

Recall the definition of $\mathcal{T}^\star_\Xi$ in (7.1.14) and the characterization that $\mathcal{T}^\star_\Xi \neq \emptyset$ if and only if there exists $\boldsymbol{b}^\star \in \mathbb{R}^n$ such that $\tau(\boldsymbol{b}^\star) \wedge \Xi \in \mathcal{T}^\star_\Xi$, cf. Remark 7.1.9.

8.1 Robust Pricing

We have seen in Section 7.1 that the optimal $\tau \in \mathcal{T}^\star_\Xi$ for the robust pricing problem of the convex Vanilla option features only t in its stopping rule. We argued that "stopping late is beneficial". Similarly, "stopping late is beneficial" also for Lookback options because the maximum is non-decreasing. However, in the example of the simple barrier option, it is clear that after the barrier B is hit, the payoff cannot become larger anymore along this path. Therefore, as we will show below, in the robust pricing problem of a pricewise constant Lookback option it is optimal to use initially a Root embedding based on some $\boldsymbol{\kappa}^0$ and after the option starts knocking in we successively switch to Root embeddings based on $\boldsymbol{\kappa}^1, \boldsymbol{\kappa}^2$, etc.

8.1.1 Simple Barrier Option

We consider the robust pricing problem of the simple barrier option,

$$\sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}\left[\mathbb{1}_{\{M_\tau \geq B\}}\right] \qquad \text{for } B > X_0 = M_0. \tag{8.1.1}$$

Theorem 8.1.1 (Optimal Embedding for Simple Barrier Option). *Let $\mathcal{T}_\Xi^\star \neq \emptyset$. Then there exist $\boldsymbol{\kappa}^0, \boldsymbol{\kappa}^1 \in [0, \Xi]^n$ such that $\tau(\boldsymbol{\kappa}) \wedge \Xi$ is optimal for problem* (8.1.1) *where*

$$\tau(\boldsymbol{\kappa}) = \inf\left\{u > 0 \ : \ X_u = K_i, \ u \geq \kappa_i^{\mathbb{1}\{H_B \leq u\}} \ \textit{ for some } i = 1, \ldots, n\right\}. \tag{8.1.2}$$

Before we can prove this result, we need an auxiliary statement.

Lemma 8.1.2 (Root Embedding for Process Started at Random Time). *Suppose $(\Omega, \mathcal{G})$ is a probability space with filtration $\mathbb{G} := (\mathcal{G}_t)_{t \geq 0}$. Assume that* $\mathrm{GBM}^{(\tilde{\vartheta}, B)}$ *is a Geometric Brownian motion relative to $\mathbb{G}$ started in B at time $\tilde{\vartheta}$, independently of $\mathcal{G}_{\tilde{\vartheta}}$ and that $\tilde{\vartheta}$ is $\mathcal{G}_0$ measurable. For $B \in (0, \infty)$ consider the process $\tilde{X}$ defined* *by*

$$\tilde{X}_t := \begin{cases} B & \textit{if } t < \tilde{\vartheta}, \\ \mathrm{GBM}_t^{(\tilde{\vartheta},B)} & \textit{if } t \geq \tilde{\vartheta}. \end{cases} \tag{8.1.3}$$

Let $\mu^{>B}$ *be a measure such that the SEP of* $\mu^{>B}$ *into* $\tilde{X}$ *has a solution* $\gamma \geq \tilde{\vartheta}$ *which satisfies* $\mathcal{L}(\tilde{X}_{\gamma\wedge\Xi}) \in [\mu^{>B}]$ *where we recall the formalism of* (6.1.2).

Then there exists $\tilde{\boldsymbol{\delta}} \in [0,\Xi]^n$ *such that*

$$\tilde{\tau}(\tilde{\boldsymbol{\delta}}) := \inf\left\{u > \tilde{\vartheta} \ : \ \tilde{X}_u = K_i, \ u \geq \tilde{\delta}_i \ \textit{ for some } i = 1,\dots,n\right\} \tag{8.1.4}$$

satisfies

$$\mathcal{L}(\tilde{X}_{\tilde{\tau}(\tilde{\boldsymbol{\delta}})\wedge\Xi}) \in [\mu^{>B}], \tag{8.1.5a}$$

$$\mathcal{L}\left(\tilde{X}_{\tilde{\tau}(\tilde{\boldsymbol{\delta}})\wedge\xi}\right) \succcurlyeq_{\mathrm{c}} \mathcal{L}\left(\tilde{X}_{\gamma\wedge\xi}\right) \qquad \forall \xi \leq \Xi. \tag{8.1.5b}$$

Furthermore, let $\hat{X}$ *be the process from* (8.1.3) *with* $\tilde{\vartheta}$ *replaced by* $\hat{\vartheta}$*. Assume*

$$\mathbb{P}\left[\hat{\vartheta} \leq \xi\right] \geq \mathbb{P}\left[\tilde{\vartheta} \leq \xi\right] \qquad \forall \xi \geq 0, \tag{8.1.6a}$$

$$\mathbb{P}\left[\hat{\vartheta} \leq \alpha\right] = \mathbb{P}\left[\tilde{\vartheta} \leq \Xi\right] \qquad \textit{for some } \alpha \leq \Xi. \tag{8.1.6b}$$

Then, there exists $\hat{\boldsymbol{\delta}} \in [0,\Xi]^n$ *such that*

$$\mathcal{L}(\hat{X}_{\hat{\tau}(\hat{\boldsymbol{\delta}})\wedge\Xi}) \in [\mu^{>B}], \tag{8.1.7a}$$

$$\mathcal{L}\left(\hat{X}_{\hat{\tau}(\hat{\boldsymbol{\delta}})\wedge\xi}\right) \succcurlyeq_{\mathrm{c}} \mathcal{L}\left(\tilde{X}_{\tilde{\tau}(\tilde{\boldsymbol{\delta}})\wedge\xi}\right) \succcurlyeq_{\mathrm{c}} \mathcal{L}\left(\tilde{X}_{\gamma\wedge\xi}\right) \qquad \forall \xi \leq \Xi. \tag{8.1.7b}$$

Proof of Lemma 8.1.2. Our first observation is that it is enough to prove the claim for (probability) measures $\mu^{>B}$ with support $\mathrm{supp}(\mu^{>B}) \subseteq \{K_1,\dots,K_n\} \cup \{B\}$, i.e.

for atomic measures with bounded support. In view of the arguments in Section 7.1 the general case then follows from compactness arguments.

Set[1]

$$\tilde{Y}_t := \begin{cases} B - (\tilde{\vartheta} - t) & \text{if } t < \tilde{\vartheta}, \\ \mathrm{GBM}_t^{(\tilde{\vartheta}, B)} & \text{if } t \geq \tilde{\vartheta}, \end{cases} \tag{8.1.8}$$

and introduce the continuous, time-homogeneous, transient Markov process $\tilde{Z}$ defined by

$$\tilde{Z}_t := \left(\tilde{Y}_t, \mathbb{1}_{\left\{\tilde{\vartheta} \leq t\right\}}\right) \in \left\{[B - r_{\tilde{\vartheta}}, B) \times \{0\}\right\} \cup \left\{\mathbb{R}_+ \times \{1\}\right\} \tag{8.1.9}$$

where $r_{\tilde{\vartheta}} := \inf\left\{x : \mathbb{P}\left[\tilde{\vartheta} > x\right] = 0\right\}$ is the right end of support of $\tilde{\vartheta}$.

Consider the SEP of $\mu^{>B} \otimes \delta_{\{1\}}$ into $\tilde{Z}$ to which γ is a solution by assumption. Then Rost [86, Theorem 1] implies existence of a stopping time T of minimal residual expectation and Rost [86, Theorem 3] gives existence of a barrier $\mathcal{B}$ with sections $\mathcal{B}_t$ such that

$$\inf\left\{u \geq 0 \ : \ \tilde{Z}_t \in \mathcal{B}_{t+}\right\} \leq T \leq \inf\left\{u \geq 0 \ : \ \tilde{Z}_t \in \mathcal{B}_t\right\}. \tag{8.1.10}$$

Now, since the measure $\mu^{>B} \otimes \delta_{\{1\}}$ is concentrated on $\left\{K_1, \ldots, K_n, B\right\} \times \{1\}$ we must have

$$\mathcal{B}_t \subseteq \left\{K_1, \ldots, K_n, B\right\} \times \{1\} \qquad \forall t \geq 0 \tag{8.1.11}$$

[1] We are thankful to Alexander Cox for suggesting this definition.

and then because the one-point sets $\{x\} \times \{1\}$ are regular for the process $\tilde{Z}$ for all $x \in \mathbb{R}_+$ we conclude

$$T = \tilde{\tau}(\tilde{\boldsymbol{\delta}}) \tag{8.1.12}$$

for some $\tilde{\boldsymbol{\delta}} \in \mathbb{R}^n$.

For any non-negative $g(x) = G''(x)x^2$ we set

$$f(\tilde{Z}_u) := g(B) \cdot \mathbb{1}_{\{\tilde{\vartheta} > u\}} + g(\tilde{Y}_u) \cdot \mathbb{1}_{\{\tilde{\vartheta} \leq u\}} = G''(\tilde{X}_u)\tilde{X}_u^2.$$

By assumption, $\tilde{Z}$ is a semi-martingale. Then, by applying Itô's formula to $F(\tilde{Z})$, where $F'' = f$, and using the embedding property, we obtain

$$\mathbb{E}\left[F\left(\tilde{Z}_{\tilde{\tau}(\tilde{\boldsymbol{\delta}})}\right)\right] = \mathbb{E}\left[F\left(\tilde{Z}_{\gamma}\right)\right].$$

Therefore, as in the proof of Proposition 7.1.4, the minimal residual expectation property of $\tilde{\tau}(\tilde{\boldsymbol{\delta}})$ implies that

$$\begin{aligned}
\mathbb{E}\left[\int_0^{\tilde{\tau}(\tilde{\boldsymbol{\delta}})\wedge\xi} G''(\tilde{X}_u)\tilde{X}_u^2 \mathrm{d}u\right] &= \mathbb{E}\left[\int_0^{\tilde{\tau}(\tilde{\boldsymbol{\delta}})\wedge\xi} f(\tilde{Z}_u)\mathrm{d}u\right] \\
&\geq \mathbb{E}\left[\int_0^{\gamma\wedge\xi} f(\tilde{Z}_u)\mathrm{d}u\right] = \mathbb{E}\left[\int_0^{\gamma\wedge\xi} G''(\tilde{X}_u)\tilde{X}_u^2 \mathrm{d}u\right]
\end{aligned}$$

and hence

$$\mathbb{E}\left[G\left(\tilde{X}_{\tilde{\tau}(\tilde{\boldsymbol{\delta}})\wedge\xi}\right)\right] \geq \mathbb{E}\left[G\left(\tilde{X}_{\gamma\wedge\xi}\right)\right] \qquad \forall \xi \geq 0.$$

Therefore, we observe further,

$$\mu^{>B} = \mathcal{L}\left(\tilde{Y}_{\tilde{\tau}(\tilde{\boldsymbol{\delta}})}\right) = \mathcal{L}\left(\tilde{X}_{\tilde{\tau}(\tilde{\boldsymbol{\delta}})}\right) \succcurlyeq_{\mathrm{c}} \mathcal{L}\left(\tilde{X}_{\tilde{\tau}(\tilde{\boldsymbol{\delta}})\wedge\Xi}\right) \succcurlyeq_{\mathrm{c}} \mathcal{L}\left(\tilde{X}_{\gamma\wedge\Xi}\right) \in \left[\mu^{>B}\right]$$

and hence

$$\mathcal{L}\left(\tilde{X}_{\tilde{\tau}(\tilde{\boldsymbol{\delta}})\wedge\Xi}\right) \in \left[\mu^{>B}\right].$$

The previous equation implies $\tilde{\boldsymbol{\delta}} \in [0,\Xi]^n$. Together, this proves (8.1.5a) and (8.1.5b).

By (8.1.6a)–(8.1.6b) and Shaked and Shanthikumar [88, Theorem 1.A.1] there exists a random variable $\chi \geq 0$ such that $\hat{\vartheta} + \chi \sim \tilde{\vartheta}$. Therefore, the stopping time

$$\hat{\gamma} := \inf\left\{u > \hat{\vartheta} \ : \ \hat{X}_u = K_i,\ u \geq (\tilde{\delta}_i - \chi)\mathbb{1}_{\left\{\tilde{\delta}_i - \tilde{\vartheta} > 0\right\}} \ \text{ for some } i = 1, \ldots, n\right\} \geq \hat{\vartheta}$$

satisfies

$$\mathcal{L}(\hat{X}_{\hat{\gamma}\wedge\Xi}) \in [\mathcal{L}(\tilde{X}_{\gamma\wedge\Xi})] = [\mu^{>B}] \tag{8.1.13}$$

and

$$\mathcal{L}(\hat{X}_{\hat{\gamma}\wedge\Xi}) \succcurlyeq_{\mathrm{c}} \mathcal{L}(\tilde{X}_{\gamma\wedge\Xi}). \tag{8.1.14}$$

By similar arguments as before we obtain $\hat{\tau}(\hat{\boldsymbol{\delta}})$ which satisfies

$$\mathcal{L}(\hat{X}_{\hat{\tau}(\hat{\boldsymbol{\delta}})\wedge\Xi}) \in [\mu^{>B}]$$

and

$$\begin{aligned}\mathbb{E}\left[\int_0^{\hat{\tau}(\hat{\boldsymbol{\delta}})\wedge\xi} G''(\hat{X}_u)\hat{X}_u^2 \mathrm{d}u\right] &= \mathbb{E}\left[\int_0^{\hat{\tau}(\hat{\boldsymbol{\delta}})\wedge\xi} f(\hat{Z}_u)\mathrm{d}u\right] \\ &\geq \mathbb{E}\left[\int_0^{\hat{\gamma}\wedge\xi} f(\hat{Z}_u)\mathrm{d}u\right] = \mathbb{E}\left[\int_0^{\hat{\gamma}\wedge\xi} G''(\hat{X}_u)\hat{X}_u^2 \mathrm{d}u\right]\end{aligned}$$

and hence

$$\mathbb{E}\left[G\left(\hat{X}_{\hat{\tau}(\hat{\boldsymbol{\delta}})\wedge\xi}\right)\right] \geq \mathbb{E}\left[G\left(\hat{X}_{\hat{\gamma}\wedge\xi}\right)\right] \qquad \forall \xi \geq 0.$$

Equations (8.1.7a)–(8.1.7b) follow. □

Proof of Theorem 8.1.1. Let $\tau \in \mathcal{T}_{\Xi}^{\star}$. Observe

$$\mathbb{P}\left[M_\tau \geq B\right] = \mathbb{P}\left[M_{\tau\wedge H_B} \geq B\right] = \mathbb{P}\left[X_{\tau\wedge H_B} = B\right].$$

Define

$$\mu^{\leq B} := \mathcal{L}\left(X_{\tau\wedge H_B}\right).$$

By Proposition 7.1.4 and Corollary 7.1.7 there exists $\boldsymbol{\kappa}^0 \in [0, \Xi]$ (here we set $\kappa_j^0 := 0$ for j such that $K_j \geq B$) such that

$$\mathcal{L}\left(X_{\tau(\boldsymbol{\kappa}^0)\wedge H_B\wedge\Xi}\right) \in \left[\mathcal{L}\left(X_{\tau\wedge H_B}\right)\right]$$

and for all $\xi \leq \Xi$

$$\begin{aligned} \mathbb{P}\left[M_{\tau(\boldsymbol{\kappa}^0)\wedge\xi} \geq B\right] &= \mathbb{P}\left[X_{\tau(\boldsymbol{\kappa}^0)\wedge H_B\wedge\xi} = B\right] \\ &\geq \mathbb{P}\left[X_{\tau\wedge H_B\wedge\xi} = B\right] = \mathbb{P}\left[M_{\tau\wedge\xi} \geq B\right]. \end{aligned} \tag{8.1.15}$$

Using the notation of Lemma 8.1.2 we choose

$$\begin{aligned} \tilde{\vartheta} &:= \inf\left\{u > 0 : X_{\tau\wedge u} = B\right\} \wedge \Xi, \\ \hat{\vartheta} &:= \inf\left\{u > 0 : X_{\tau(\boldsymbol{\kappa}^0)\wedge u} = B\right\} \wedge \Xi, \\ \gamma &:= \tau \cdot \mathbb{1}_{\{\tau > H_B\}} + \Xi \cdot \mathbb{1}_{\{\tau \leq H_B\}}, \\ \mu^{>B} &:= \mathcal{L}\left(\tilde{X}_\gamma\right). \end{aligned}$$

It follows from (8.1.15) that $\tilde{\vartheta}$ and $\hat{\vartheta}$ satisfy the conditions (8.1.6a) and (8.1.6b) of Lemma 8.1.2. The other assumptions of Lemma 8.1.2 are easily verified and hence we obtain $\hat{\boldsymbol{\delta}} = \boldsymbol{\kappa}^1$ such that

$$\mathcal{L}\left(\hat{X}_{\hat{\tau}(\boldsymbol{\kappa}^1)\wedge\Xi}\right) \in \left[\mu^{>B}\right].$$

Denote by $L_t^{K_i}(U)$ the local time at K_i of the process U until time t. Finally, we verify the embedding property of $\tau(\boldsymbol{\kappa}) \wedge \Xi$ by verifying (6.1.5). We compute, recalling the process $\mathrm{GBM}^{(H_B,B)}$ from Lemma 8.1.2 and (8.1.7a),

$$\begin{aligned}
\mathbb{E}\left[L_\tau^{K_i}(X)\right] &= \mathbb{E}\left[L_{\tau\wedge H_B}^{K_i}(X)\right] + \mathbb{E}\left[\left(L_\tau^{K_i}(X) - L_{H_B}^{K_i}(X)\right)\mathbb{1}_{\{\tau > H_B\}}\right] \\
&= \mathbb{E}\left[L_{\tau(\boldsymbol{\kappa}^0)\wedge H_B\wedge\Xi}^{K_i}(X)\right] + \mathbb{E}\left[L_\tau^{K_i}\left(\mathrm{GBM}^{(H_B,B)}\right)\mathbb{1}_{\{\tau > H_B\}}\right] \\
&= \mathbb{E}\left[L_{\tau(\boldsymbol{\kappa}^0)\wedge H_B\wedge\Xi}^{K_i}(X)\right] + \mathbb{E}\left[L_{\hat{\tau}(\boldsymbol{\kappa}^1)\wedge\Xi}^{K_i}(\hat{X})\right] \\
&= \mathbb{E}\left[L_{\tau(\boldsymbol{\kappa})}^{K_i}(X)\right].
\end{aligned} \tag{8.1.16}$$

Together, we have constructed $\tau(\boldsymbol{\kappa}) \wedge \Xi \in \mathcal{T}_\Xi^\star$ which yields a larger probability of hitting the barrier B than τ. □

Remark 8.1.3 (Lower Bound on Realized Variance). Similar to the remark about lower bounds on realized variance in Section 7.3.2, it follows from the construction in the proof of Theorem 8.1.1 that a lower bound on Realized Variance is not effective because it would either be inconsistent with the option prices or satisfied automatically.

8.1.2 Double-Touch Barrier Option

Now we consider the example of the Double-Touch Barrier option. We are able to find the optimal embedding with similar arguments as in the proof of Theorem

8.1.1. Mathematically, the problem reads,

$$\sup_{\tau\in\mathcal{T}^\star_\Xi} \mathbb{E}\left[\mathbb{1}_{\{I_\tau\le B_1, M_\tau\ge B_2\}}\right] \qquad \text{for } B_1 < X_0 = M_0 < B_2 \tag{8.1.17}$$

where $I_t := \inf_{u\le t} X_u$.

Theorem 8.1.4 (Optimal Embedding for Double-Touch Barrier Option)**.** *Let $\mathcal{T}^\star_\Xi \neq \emptyset$. Then, there exist $\boldsymbol{\kappa}^{0,0,0}, \boldsymbol{\kappa}^{0,1,0}, \boldsymbol{\kappa}^{1,0,0}, \boldsymbol{\kappa}^{1,1,0}, \boldsymbol{\kappa}^{1,1,1} \in [0,\Xi]^n$ such that $\tau(\boldsymbol{\kappa})\wedge\Xi$ is optimal for problem* (8.1.17) *where*

$$\begin{aligned}\tau(\boldsymbol{\kappa}) = \inf\Bigg\{u>0\ :\ X_u = K_i,\ u \ge \kappa_i^{\mathbb{1}\{H_{B_1}\le u\},\mathbb{1}\{H_{B_2}\le u\},\mathbb{1}\{H_{B_1}\le H_{B_2}\le u\}}\\ \textit{for some } i=1,\dots,n\Bigg\}.\end{aligned} \tag{8.1.18}$$

Proof. Take $\tau\in\mathcal{T}^\star_\Xi$. The idea of the proof is to show that we can improve both terms in the decomposition

$$\begin{aligned}&\mathbb{P}\left[I_\tau\le B_1, M_\tau\ge B_2\right]\\ =&\mathbb{P}\left[X_{\tau\wedge H_{B_1}} = B_1, H_{B_2}\le H_{B_1}\wedge\tau\right] + \mathbb{P}\left[X_{\tau\wedge H_{B_2}} = B_2, H_{B_1}\le H_{B_2}\wedge\tau\right].\end{aligned} \tag{8.1.19}$$

Define

$$\mu^{0,0,0} := \mathcal{L}\left(X_{\tau\wedge H_{B_1}\wedge H_{B_2}}\right).$$

Similarly as in the proof of Theorem 8.1.1 we use $\tau(\boldsymbol{\kappa}^{0,0,0})\wedge H_{B_1}\wedge H_{B_2}\wedge\Xi$ instead of $\tau\wedge H_{B_1}\wedge H_{B_2}$ to embed $\mu^{0,0,0}$ which yields for $\xi\le\Xi$,

$$\mathbb{P}\left[X_{\tau(\boldsymbol{\kappa}^{0,0,0})\wedge H_{B_1}\wedge H_{B_2}\wedge\xi} = B_1\right] \ge \mathbb{P}\left[X_{\tau\wedge H_{B_1}\wedge H_{B_2}} = B_1\right], \tag{8.1.20a}$$

$$\mathbb{P}\left[X_{\tau(\boldsymbol{\kappa}^{0,0,0})\wedge H_{B_1}\wedge H_{B_2}\wedge\xi} = B_2\right] \ge \mathbb{P}\left[X_{\tau\wedge H_{B_1}\wedge H_{B_2}\wedge\xi} = B_2\right]. \tag{8.1.20b}$$

Using the notation of Lemma 8.1.2 we define

$$
\begin{aligned}
B^{0,1,0} &:= B_2, \\
\tilde{\vartheta}^{0,1,0} &:= \inf\left\{u > 0 : X_{\tau \wedge H_{B_1} \wedge H_{B_2} \wedge u} = B_2\right\} \wedge \Xi, \\
\hat{\vartheta}^{0,1,0} &:= \inf\left\{u > 0 : X_{\tau(\boldsymbol{\kappa}^{\mathbf{0,0,0}}) \wedge H_{B_1} \wedge H_{B_2} \wedge u} = B_2\right\} \wedge \Xi, \\
\gamma^{0,1,0} &:= \tau \wedge H_{B_1} \cdot \mathbb{1}_{\left\{\tau \wedge H_{B_1} > H_{B_2}\right\}} + \Xi \cdot \mathbb{1}_{\left\{\tau \wedge H_{B_1} \leq H_{B_2}\right\}}, \\
\mu^{0,1,0} &:= \mathcal{L}\left(\tilde{X}^{0,1,0}_{\gamma^{0,1,0}}\right),
\end{aligned}
$$

where $\tilde{X}^{0,1,0}$ is the process defined in (8.1.3) with $B = B^{0,1,0}$ and $\tilde{\vartheta} = \tilde{\vartheta}^{0,1,0}$. Similarly for $\hat{X}^{0,1,0}$. Then Lemma 8.1.2 yields

$$
\hat{\tau}^{0,1,0}(\boldsymbol{\kappa}^{0,1,0}) = \inf\left\{u > \hat{\vartheta}^{0,1,0} \;:\; X_u = K_i,\; u \geq \kappa_i^{0,1,0} \;\text{ for some } i = 1, \ldots, n\right\}
$$

such that $\mathcal{L}\left(\hat{X}^{0,1,0}_{\hat{\tau}^{0,1,0}(\boldsymbol{\kappa}^{0,1,0}) \wedge H_{B_1} \wedge \Xi}\right) \in [\mu^{0,1,0}]$. Further, for convex, non-negative G and $\xi \leq \Xi$ equation (8.1.7b) reads

$$
\begin{aligned}
& \mathbb{E}\left[G\left(X_{\hat{\tau}^{0,1,0}(\boldsymbol{\kappa}^{0,1,0}) \wedge H_{B_1} \wedge \xi}\right) \mathbb{1}_{\left\{H_{B_2} < \tau(\boldsymbol{\kappa}^{0,0,0}) \wedge H_{B_1} \wedge \xi\right\}}\right] \\
& \quad + G(B_2)\mathbb{P}\left[H_{B_2} \geq \tau(\boldsymbol{\kappa}^{0,0,0}) \wedge H_{B_1} \wedge \xi\right] \\
\geq \quad & \mathbb{E}\left[G\left(X_{\tau \wedge H_{B_1} \wedge \xi}\right) \mathbb{1}_{\left\{H_{B_2} < \tau \wedge H_{B_1} \wedge \xi\right\}}\right] \\
& \quad + G(B_2)\mathbb{P}\left[H_{B_2} \geq \tau \wedge H_{B_1} \wedge \xi\right].
\end{aligned}
$$

Consequently, by $G(B_2) \geq 0$, equation (8.1.20b) and the same argument as in the proof of Corollary 7.1.7, we deduce

$$
\begin{aligned}
& \mathbb{P}\left[X_{\hat{\tau}^{0,1,0}(\boldsymbol{\kappa}^{0,1,0}) \wedge H_{B_1} \wedge \xi} = B_1, H_{B_2} < \tau(\boldsymbol{\kappa}^{0,0,0}) \wedge H_{B_1} \wedge \xi\right] \\
\geq \; & \mathbb{P}\left[X_{\tau \wedge H_{B_1} \wedge \xi} = B_1, H_{B_2} < \tau \wedge H_{B_1} \wedge \xi\right]
\end{aligned}
\tag{8.1.21}
$$

and hence an improvement of the first term on the right hand side of (8.1.19).

Completely analogously, we obtain $\hat{\tau}^{1,0,0}(\boldsymbol{\kappa}^{1,0,0})$ which improves the second term on the right hand side of (8.1.19).

Again, using the notation of Lemma 8.1.2 we define

$$\begin{aligned}
B^{1,1,0} &:= B_1, \\
\tilde{\vartheta}^{1,1,0} &:= \inf\left\{u > \tilde{\vartheta}^{0,1,0} : X_{\tau\wedge H_{B_1}\wedge u} = B_1\right\} \wedge \Xi, \\
\hat{\vartheta}^{1,1,0} &:= \inf\left\{u > \hat{\vartheta}^{0,1,0} : X_{\hat{\tau}^{0,1,0}(\boldsymbol{\kappa}^{\mathbf{0,1,0}})\wedge H_{B_1}\wedge u} = B_1\right\} \wedge \Xi, \\
\gamma^{1,1,0} &:= \tau \cdot \mathbb{1}_{\left\{H_{B_2}<H_{B_1}\leq\tau\right\}} + \Xi \cdot \mathbb{1}_{\left\{H_{B_2}<H_{B_1}\leq\tau\right\}^{\mathrm{C}}}, \\
\mu^{1,1,0} &:= \mathcal{L}\left(\tilde{X}^{1,1,0}_{\gamma^{1,1,0}}\right),
\end{aligned}$$

Recalling (8.1.21), Lemma 8.1.2 yields $\hat{\tau}^{1,1,0}(\boldsymbol{\kappa}^{1,1,0})$ such that $\mathcal{L}\left(\hat{X}^{1,1,0}_{\hat{\tau}^{1,1,0}(\boldsymbol{\kappa}^{1,1,0})\wedge\Xi}\right) \in [\mu^{1,1,0}]$.

Completely analogously (after marking the corresponding definitions as above), we obtain $\hat{\tau}^{1,1,1}(\boldsymbol{\kappa}^{1,1,1})$ such that $\mathcal{L}\left(\hat{X}^{1,1,0}_{\hat{\tau}^{1,1,1}(\boldsymbol{\kappa}^{1,1,1})\wedge\Xi}\right) \in [\mu^{1,1,1}]$.

The embedding property follows similarly as in the proof of Theorem 8.1.1. □

8.1.3 Piecewise Constant Lookback Option

Now we consider the robust pricing problem of a piecewise constant Lookback option,

$$\sup_{\tau\in\mathcal{T}^{\star}_{\Xi}} \mathbb{E}\left[G\left(M_\tau\right)\right] \tag{8.1.22}$$

for piecewise constant, right-continuous, non-decreasing $G : [S_0, \infty) \to \mathbb{R}$ with discontinuities $B_1, \ldots, B_k$ such that $X_0 =: B_0 < B_1 < \cdots < B_k < \infty$.

Theorem 8.1.5 (Optimal Embedding for Piecewise Constant Lookback Option)**.** *Let $\mathcal{T}^{\star}_{\Xi} \neq \emptyset$. Then there exist $\boldsymbol{\kappa}^0, \boldsymbol{\kappa}^1, \ldots, \boldsymbol{\kappa}^k \in [0, \Xi]^n$ such that $\tau(\boldsymbol{\kappa}) \wedge \Xi$ is optimal* *for problem* (8.1.22) *where*

$$\tau(\boldsymbol{\kappa}) = \inf\left\{u > 0 \ : \ X_u = K_i, \ u \geq \kappa_i^{\varpi_u} \ \textit{ for some } i = 1, \dots, n\right\}, \tag{8.1.23a}$$

$$\varpi_u = \max\left\{j \leq k : M_u \geq B_j\right\}. \tag{8.1.23b}$$

Proof. The proof goes along the lines of the proofs of Theorems 8.1.1 and 8.1.4, essentially making repeatedly use of Lemma 8.1.2. □

8.1.4 Different Optimal Embedding for Different Barrier Options

We now demonstrate that in general the optimal embedding depends on the payoff function. More precisely, there is no "universally" optimal solution to the robust pricing problem of Lookback options in the sense that one solution simultaneously maximizes the price of each simple barrier option. This is different from the case $\Xi = \infty$ where it is known that the Azéma-Yor solution maximizes the law of the maximum in stochastic order amongst all embeddings. However, the analysis of Cox and Obłój [22, 28] for double touch/no-touch barrier options also arrived at different optimal embedding for different barrier options (also in the case $\Xi = \infty$).

We consider the simplest, non-trivial setup, namely the case of just one market-traded option. Our aim is to show that there exist $0 < \Xi < \infty$, a price of the put option $p^\star(K)$ and barriers $K < B_1 < B_2$ such that the optimal solution for the B_2-barrier option is suboptimal for the B_1-barrier option.

Define

$$p^\star(K) := \mathbb{E}\left[\left(K - X_{H_{B_1}\wedge\Xi}\right)^+\right]$$

and hence, by construction the stopping time $H_{B_1} \wedge \Xi$ solves the robust pricing problem for the simple barrier option problem with barrier B_1.

As for the simple barrier option with barrier B_2 we conclude from Theorem 8.1.1 that the solution is obtained by choosing $\xi < \Xi$ such that

$$p^\star(K) = \mathbb{E}\left[\left(K - X_{H_{B_2} \wedge \xi}\right)^+\right].$$

Clearly,

$$\mathbb{P}\left[M_{H_{B_1} \wedge \Xi} \geq B_1\right] > \mathbb{P}\left[M_{H_{B_2} \wedge \xi} \geq B_1\right]$$

and therefore the optimal embeddings for these two barrier options are different.

8.2 Robust Hedging

Subsequently, we consider payoffs satisfying the following assumption.

Assumption (iv) (Lookback Payoff)**.** *The function* $G : [X_0, \infty) \to [0, \infty)$ *is right-continuous, non-decreasing and pricewise constant with discontinuities at* $B_0 := S_0 < B_1 < \cdots < B_k$.

Following the same arguments as in Section 7.2, define the *penalized Lookback payoff*

$$G^{\boldsymbol{\lambda}}(x, m) := G(m) - \sum_{i=1}^{n} \lambda_i \left(K_i - x\right)^+. \tag{8.2.1}$$

Theorem 8.2.1 (Duality)**.** *Let G satisfy Assumption* (iv). *Recall the formalism of Section 6.3. Then,*

$$\sup_{\tau\in\mathcal{T}_\Xi^\star}\mathbb{E}\left[G\left(M_\tau\right)\right]=\inf_{\boldsymbol{\lambda}\in\mathbb{R}^n}\sup_{\tau\in\mathcal{T}_\Xi}\left\{\mathbb{E}\left[G^{\boldsymbol{\lambda}}\left(X_\tau,M_\tau\right)\right]+\sum_{i=1}^n\lambda_i p^\star(K_i)\right\}. \tag{8.2.2}$$

and there is no duality gap, i.e.

$$\mathbb{A}(G(\bar{S}_T);\mathfrak{P}_\Xi,\mathcal{A},\mathcal{P})=\sup_{\tau\in\mathcal{T}_\Xi^\star}\mathbb{E}\left[G(M_\tau)\right]=\sup_{\mathbb{P}^S}\mathbb{E}_{\mathbb{P}^S}\left[G(\bar{S}_T)\right] \tag{8.2.3}$$

where the supremum is over all $(\mathfrak{P}_\Xi,\mathcal{Y},\mathcal{P})$-market models $\left(\Omega^S,\mathcal{F}^S,\mathbb{F}^S,\mathbb{P}^S\right)$.

The proof of this result will be developed in the next sections.

8.2.1 An Optimal Stopping Problem

Let Assumption (iv) hold.

Define the *penalized Lookback value function* for $x\le m$,

$$V^{\boldsymbol{\lambda}}(t,x,m):=\sup_{\tau\in\mathcal{T}_{t,\Xi}}\mathbb{E}\left[G^{\boldsymbol{\lambda}}(X_\tau,M_\tau)\;\Big|\;X_t=x,M_t=m\right]. \tag{8.2.4}$$

Set $\tilde{V}^{\boldsymbol{\lambda}}(u,x,m)=V^{\boldsymbol{\lambda}}(u,x,m)-G^{\boldsymbol{\lambda}}(x,m)$ and

$$\tau_t^\star=\inf\left\{u>t\;:\;\tilde{V}^{\boldsymbol{\lambda}}(u,X_u,M_u)=0\right\}\wedge\Xi. \tag{8.2.5}$$

The stopping time $\tau_t^\star$ is optimal for (8.2.4). Further, $V^{\boldsymbol{\lambda}}$ is non-increasing in time and hence $\tau_t^\star$ defines a Root stopping time. It follows from Theorem 8.1.5 that the continuation region of the optimal stopping problem can be chosen as[2]

$$D^{\boldsymbol{\lambda}} = \Big\{(t,x,m) \in [0,\Xi) \times \mathbb{R}_+ \times [M_0,\infty) \;:\; x \leq m, \\ x \neq K_i \;\text{ or }\; t < \kappa_i^{\ell(m)} \;\text{ for all } i = 1,\dots,n\Big\}, \tag{8.2.6}$$

for some $\boldsymbol{\kappa}^0,\dots,\boldsymbol{\kappa}^k$ and $\ell(m) := \max\{j \leq k : m \geq B_j\}$. Observe further that

$$(t,x,m_1) \notin D^{\boldsymbol{\lambda}} \implies (t,x,m_2) \notin D^{\boldsymbol{\lambda}} \qquad \text{whenever } m_1 \leq m_2, \tag{8.2.7}$$

which is because for every $\tau \in \mathcal{T}_{t,\Xi}$ we have

$$G(M_\tau \vee m_1) - G(m_1) \geq G(M_\tau \vee m_2) - G(m_2)$$

and hence, since there is "no interaction between x and m",

$$\begin{aligned} 0 \;&\geq\; \mathbb{E}\Big[G^{\boldsymbol{\lambda}}(X_\tau, M_\tau) \;\Big|\; X_t = x, M_t = m_1\Big] - G^{\boldsymbol{\lambda}}(x,m_1) \\ &\geq\; \mathbb{E}\Big[G^{\boldsymbol{\lambda}}(X_\tau, M_\tau) \;\Big|\; X_t = x, M_t = m_2\Big] - G^{\boldsymbol{\lambda}}(x,m_2). \end{aligned}$$

For simplicity, the following result is stated for $G(m) = \mathbb{1}_{\{m \geq B\}}$. However, it is clear that it extends to payoffs satisfying Assumption (iv) on page 179. We will use this observation in the proofs of Proposition 8.2.3 and Theorem 8.2.1.

Lemma 8.2.2 (Properties of $V^{\boldsymbol{\lambda}}$)**.** *Suppose $G(m) = \mathbb{1}_{\{m \geq B\}}$ for some $B > X_0$. Set*

$$\Delta^{<B} := \{(t,x,m) \in (0,\Xi) \times \mathbb{R}_+ \times [X_0,B) : x \leq m\}, \tag{8.2.8a}$$

$$\Delta^{>B} := \{(t,x,m) \in (0,\Xi) \times \mathbb{R}_+ \times [B,\infty) : x \leq m\}. \tag{8.2.8b}$$

[2] A similar remark as in the footnote on page 151 applies.

Then, firstly, $(t, x) \mapsto V^{\boldsymbol{\lambda}}(t, x, m \vee x)$ is continuous on $[0, \Xi - \epsilon] \times \mathbb{R}_+$ for every $\epsilon > 0$ and $m \geq X_0$.

Secondly, for $(t, x, m),\ (t, x, \tilde{m}) \in \Delta^{<B} \cap D^{\boldsymbol{\lambda}}$ the value function $V^{\boldsymbol{\lambda}}$ satisfies

$$V^{\boldsymbol{\lambda}}(t, x, m) = V^{\boldsymbol{\lambda}}(t, x, \tilde{m}) = f(t, x), \tag{8.2.9a}$$

$$f(t, x) \geq 0 - \sum_{i=1}^{n} \lambda_i \left(K_i - x\right)^+, \tag{8.2.9b}$$

$$\frac{\partial f}{\partial t}(t, x) + \frac{1}{2} x^2 \frac{\partial^2 f}{\partial x^2}(t, x) = 0, \tag{8.2.9c}$$

Thirdly, for $(t, x, m),\ (t, x, \tilde{m}) \in \Delta^{>B} \cap D^{\boldsymbol{\lambda}}$ the value function $V^{\boldsymbol{\lambda}}$ satisfies[3]

$$V^{\boldsymbol{\lambda}}(t, x, m) = V^{\boldsymbol{\lambda}}(t, x, \tilde{m}) = f(t, x), \tag{8.2.10a}$$

$$f(t, x) \geq 1 - \sum_{i=1}^{n} \lambda_i \left(K_i - x\right)^+, \tag{8.2.10b}$$

$$\frac{\partial f}{\partial t}(t, x) + \frac{1}{2} x^2 \frac{\partial^2 f}{\partial x^2}(t, x) = 0. \tag{8.2.10c}$$

Proof. Properties (8.2.9a)–(8.2.9b) and (8.2.10a)–(8.2.10b) are clear.

For every $\epsilon > 0$ the function f is continuous on $[0, \Xi - \epsilon] \times \mathbb{R}_+$. This yields in turn the claim regarding continuity of $\tilde{V}^{\boldsymbol{\lambda}}$.

Now take $\epsilon > 0$ small enough such that $(t, x, x + \epsilon) \in \Delta^{<B} \cap D^{\boldsymbol{\lambda}}$ and $B_\epsilon(t, x) \times \{x + \epsilon\} \subseteq \Delta^{<B} \cap D^{\boldsymbol{\lambda}}$. Classical PDE results, see e.g. Friedman [43, p. 134, Theorem 2.4], give existence of a smooth solution g to the Dirichlet problem:

$$\frac{\partial g}{\partial t} + \frac{1}{2} x^2 \frac{\partial^2 g}{\partial x^2} = 0 \qquad \text{in } B_\epsilon(t, x), \tag{8.2.11a}$$

$$g\big|_{\partial B_\epsilon(t,x)} = f. \tag{8.2.11b}$$

[3] In the present case of $G(m) = \mathbb{1}_{\{m \geq B\}}$, since $G(m) = 1$ on $\Delta^{>B}$, the continuation region is determined solely by $\lambda_i, i = 1, \ldots, n$. Note that λ_i can have either sign.

Then by Itô's formula, the optional sampling theorem, (8.2.9a), the Markov property and properties of the Snell envelope $V^{\boldsymbol{\lambda}}$,

$$\begin{aligned} g(t,x) &= \mathbb{E}\left[V^{\boldsymbol{\lambda}}\left(\tau_{B_\epsilon(t,x)}, X_{\tau_{B_\epsilon(t,x)}}, X_0 \vee X_{\tau_{B_\epsilon(t,x)}}\right) \,\middle|\, X_t = x\right] \\ &= \mathbb{E}\left[V^{\boldsymbol{\lambda}}\left(\tau_{D^{\boldsymbol{\lambda}}}, X_{\tau_{D^{\boldsymbol{\lambda}}}}, M_{\tau_{D^{\boldsymbol{\lambda}}}}\right) \,\middle|\, X_t = x, M_t = x\right] \\ &= \mathbb{E}\left[G^{\boldsymbol{\lambda}}\left(X_{\tau_{D^{\boldsymbol{\lambda}}}}, M_{\tau_{D^{\boldsymbol{\lambda}}}}\right) \,\middle|\, X_t = x, M_t = x\right] = V^{\boldsymbol{\lambda}}(t,x,x) = f(t,x), \end{aligned}$$

proving that $V^{\boldsymbol{\lambda}}$ is smooth at (t,x), and then (8.2.9c) follows. Completely analogously, we prove (8.2.10c). □

8.2.2 Pathwise Superhedging Strategy

In the following, all integrals w.r.t. X or S are understood pathwise.

Proposition 8.2.3 (Pathwise Superhedging Strategy)**.** *Let Assumption* (iv) *on page 179 hold and let $\boldsymbol{\lambda} \in \mathbb{R}^n$. Recall the definition of $V^{\boldsymbol{\lambda}}$ in* (8.2.4)*.*

Then for all $\delta > 0$ there exists a progressively measurable $\tilde{\phi}$ such that for all $t \leq \Xi$ and $X \in \Omega^{\mathrm{GBM}}$:

$$\delta + V^{\boldsymbol{\lambda}}(0, X_0, M_0) + \int_0^t \tilde{\phi}_u(X)\mathrm{d}X_u \geq V^{\boldsymbol{\lambda}}(t, X_t, M_t) \geq G^{\boldsymbol{\lambda}}(X_t, M_t) \tag{8.2.12}$$

is well-defined pathwise. Setting

$$\phi_u(S) := \tilde{\phi}_{\langle \log(S)\rangle_u}(S) \tag{8.2.13}$$

yields for all $S \in \mathfrak{P}_\Xi$:

$$\delta + V^{\boldsymbol{\lambda}}(0, X_0, M_0) + \int_0^T \phi_u(S)\mathrm{d}S_u \geq G^{\boldsymbol{\lambda}}(S_T, M_T) \tag{8.2.14}$$

and $\phi \in \Phi$, defined in (6.3.6)*.*

Proof. Take $G(m) = \mathbb{1}_{\{m \geq B\}}$ for $B > M_0 = X_0$. Let $\epsilon > 0$. Because $V^{\boldsymbol{\lambda}}(\Xi, \cdot, \cdot)$ is not continuous at (B, B) we consider an extended superhedging horizon $\Xi + \epsilon$ by defining

$$V_\epsilon^{\boldsymbol{\lambda}}(t, x, m) = \sup_{\tau \in \mathcal{T}_{t,\Xi+\epsilon}} \mathbb{E}\left[G^{\boldsymbol{\lambda}}(X_\tau, M_\tau) \;\middle|\; X_t = x, M_t = m \right]. \tag{8.2.15}$$

We clearly have

$$\lim_{\epsilon \to 0} V_\epsilon^{\boldsymbol{\lambda}}(0, X_0, M_0) = V^{\boldsymbol{\lambda}}(0, X_0, M_0),$$
$$V_\epsilon^{\boldsymbol{\lambda}}(t, x, m) \geq V^{\boldsymbol{\lambda}}(t, x, m).$$

As in (8.2.6) we denote by $D_\epsilon^{\boldsymbol{\lambda}}$ the continuation region of the stopping problem (8.2.15) which is characterized by some $\boldsymbol{\kappa}^{0,\epsilon}, \boldsymbol{\kappa}^{1,\epsilon}$. It follows from Lemma 8.2.2 that as long as the path is in $D_\epsilon^{\boldsymbol{\lambda}} \cap \Delta^{<B}$, we obtain a valid pathwise hedge by delta-hedging with

$$\tilde{\phi}_t(X) := \frac{\partial V_\epsilon^{\boldsymbol{\lambda}}}{\partial x}(t, X_t, M_0) \qquad \forall t < \tau_{D_\epsilon^{\boldsymbol{\lambda}}} \wedge H_B \wedge \Xi \tag{8.2.16}$$

which works pathwise by the Itô-Föllmer formula. Furthermore, by (8.2.9c) and the continuity assertion of Lemma 8.2.2 we have

$$\begin{aligned} &V_\epsilon^{\boldsymbol{\lambda}}(0, X_0, M_0) + \int_0^{\tau_{D_\epsilon^{\boldsymbol{\lambda}}} \wedge H_B \wedge \Xi} \tilde{\phi}_t(X) \mathrm{d}X_t \\ &= V_\epsilon^{\boldsymbol{\lambda}}\left(\tau_{D_\epsilon^{\boldsymbol{\lambda}}} \wedge H_B \wedge \Xi, X_{\tau_{D_\epsilon^{\boldsymbol{\lambda}}} \wedge H_B \wedge \Xi}, M_{\tau_{D_\epsilon^{\boldsymbol{\lambda}}} \wedge H_B \wedge \Xi} \right). \end{aligned}$$

Hence, when $X_t = B$, $t < \Xi$, we switch to the strategy

$$\tilde{\phi}_t(X) := \frac{\partial V_\epsilon^{\boldsymbol{\lambda}}}{\partial x}(t, X_t, B) \qquad \forall t \in [H_B \wedge \Xi, \tau_{D_\epsilon^{\boldsymbol{\lambda}}} \wedge \Xi) \tag{8.2.17}$$

which also works pathwise by the Itô-Föllmer formula and (8.2.10a)–(8.2.10c)

Similar to the construction in the proof of Proposition 7.2.3 we now explain how to superhedge after the continuation region $D_\epsilon^{\boldsymbol{\lambda}}$ of the stopping problem (8.2.15) is left. Observe that X can hit the points $(\kappa_i^{l,\epsilon}, K_i),\ l = 0, 1,\ i = 1, \dots, n$, at most once because the time variable is strictly increasing. Therefore, given an additional initial wealth of δ, say, we can upon hitting of any of these points, switch locally to a simple superhedging strategy, see the argument in the proof of Proposition 7.2.3. If X hits a point (t, K_i), where $t > \kappa_i^{\mathbb{1}_{\{M_t<B\}},\epsilon}$, then the value function (here $m \in \{M_0, B\}$ depending on $\mathbb{1}_{\{M_t<B\}}$),

$$x \mapsto V^{\boldsymbol{\lambda}}(t, x, m \vee x) \qquad \text{is linear or not convex at } K_i.$$

Hence, as in the proof of Proposition 7.2.3 we can switch locally to a simple superhedging strategy.

In summary, we have shown, that for arbitrarily small positive additional wealth, we can superreplicate the barrier payoff pathwise. If G satisfies Assumption (iv), then by distinguishing between $k+1$ "regions" one obtains an extension of Lemma 8.2.2 which then allows to put together the pathwise superhedging strategies for each separate region. The same time-changing arguments as in the proof of Proposition 7.2.3 yield (8.2.14). Finally, $\phi \in \Phi$ can be deduced by using that G is non-decreasing. □

8.2.3 Proof of Theorem 8.2.1

We require one last regularity result in order to prove Theorem 8.2.1.

Lemma 8.2.4 (Smoothness in $\boldsymbol{\kappa}$)**.** *Let G satisfy Assumption* (iv) *on page 179.*

Then for every $\epsilon > 0$ the mapping

$$[0, \Xi]^{n\cdot(k+1)} \to \mathbb{R}, \quad \boldsymbol{\kappa} \mapsto \mathbb{E}\left[G^{\boldsymbol{\lambda}}(X_{\tau(\boldsymbol{\kappa})\wedge(\Xi+\epsilon)}, M_{\tau(\boldsymbol{\kappa})\wedge(\Xi+\epsilon)})\right] \tag{8.2.18}$$

is continuous and is further differentiable at any point $\boldsymbol{\kappa}$ *in the interior of* $[0,\Xi]^{n\cdot(k+1)}$.

Proof. It is enough to consider $G(m) = \mathbb{1}_{\{m\geq B\}}$ for some $B > X_0$ because G satisfying Assumption (iv) is just a sum of these indicators. Further, in view of Lemma 7.2.4 it remains to prove that $\kappa_i^j \mapsto \mathbb{P}\left[M_{\tau(\boldsymbol{\kappa})\wedge(\Xi+\epsilon)} \geq B\right]$ is continuously differentiable. It is clear that $t \mapsto \mathbb{P}\left[M_{t\wedge H_K} \geq B\right]$ is continuously differentiable and hence by writing out $\mathbb{P}\left[M_{\tau(\boldsymbol{\kappa})\wedge(\Xi+\epsilon)} \geq B\right]$ as in the proof of Lemma 7.2.4 the claim follows. □

Proof of Theorem 8.2.1. Let us prove equation (8.2.2). Suppose G satisfies Assumption (iv). Then by Theorem 8.1.5 it is enough to consider only $\tau = \tau(\boldsymbol{\kappa}) \wedge \Xi$ in the optimal stopping problem on the r.h.s. of (8.2.2). This allows us to apply Theorem B.1.2. Indeed, making the identification

$$\mathscr{X} = [0,\Xi]^{n\cdot(k+1)}, \quad \mathscr{Y} = \mathbb{R}^n,$$
$$f(\boldsymbol{\kappa},\boldsymbol{\lambda}) = \mathbb{E}\left[G^{\boldsymbol{\lambda}}\left(X_{\tau(\boldsymbol{\kappa})\wedge\Xi}, M_{\tau(\boldsymbol{\kappa})\wedge\Xi}\right)\right] + \sum_{i=1}^{n} \lambda_i p^\star(K_i),$$

we verify the Theorem's assumptions. Clearly, $\mathscr{X}$ is compact and $f(\cdot,\boldsymbol{\lambda})$ is continuous on $\mathscr{X}$ for each $\boldsymbol{\lambda} \in \mathscr{Y}$. Now we verify that f is concave-convex-like in the sense of Definition B.1.1. Property (i) holds essentially by Theorem 8.1.5: Let $x_1 = \boldsymbol{\kappa}^1, x_2 = \boldsymbol{\kappa}^2$ and let α be an independent Bernoulli random variable with parameter a. The stopping time

$$\bar{\tau} := \left(\tau(\boldsymbol{\kappa}^1) \wedge \Xi\right) \mathbb{1}_{\{\alpha=0\}} + \left(\tau(\boldsymbol{\kappa}^2) \wedge \Xi\right) \mathbb{1}_{\{\alpha=1\}} \leq \Xi$$

embeds the put prices

$$a\mathbb{E}\left[\left(K_i - X_{\tau(\boldsymbol{\kappa}^1)\wedge\Xi}\right)^+\right] + (1-a)\mathbb{E}\left[\left(K_i - X_{\tau(\boldsymbol{\kappa}^2)\wedge\Xi}\right)^+\right].$$

Consequently, by Theorem 8.1.5 there exists a $x_3 = \boldsymbol{\kappa}^3 \in \mathscr{X}$ such that $\tau(\boldsymbol{\kappa}^3) \wedge \Xi$ embeds these puts and

$$\mathbb{E}\left[G\left(M_{\tau(\boldsymbol{\kappa}^3)\wedge\Xi}\right)\right] \geq \mathbb{E}\left[G\left(M_{\bar{\tau}}\right)\right] = a\mathbb{E}\left[G\left(M_{\tau(\boldsymbol{\kappa}^1)\wedge\Xi}\right)\right] + (1-a)\mathbb{E}\left[G\left(X_{\tau(\boldsymbol{\kappa}^2)\wedge\Xi}\right)\right],$$

which implies property (i). Property (ii) holds by the linearity of $f(\boldsymbol{\kappa}, \cdot)$. This establishes (8.2.2).

Take $G(m) = \mathbb{1}_{\{m \geq B\}}$. Let $\epsilon > 0$ and recall $V_\epsilon^{\boldsymbol{\lambda}}$ from (8.2.15) which clearly satisfies

$$V_\epsilon^{\boldsymbol{\lambda}}(t, X_t, M_t) \geq G^{\boldsymbol{\lambda}}(X_t, M_t) \qquad \forall t \leq \Xi, \tag{8.2.19a}$$

$$0 \leq V_\epsilon^{\boldsymbol{\lambda}}(0, X_0, M_0) - V^{\boldsymbol{\lambda}}(0, X_0, M_0) < \delta(\epsilon) \longrightarrow 0 \qquad \text{as } \epsilon \to 0. \tag{8.2.19b}$$

Recalling (8.2.6), let $\boldsymbol{\kappa}^0, \boldsymbol{\kappa}^1 \in [0, \Xi]^n$ be some optimal constants for $V^{\boldsymbol{\lambda}}$ from (8.2.4). Next we want to differentiate w.r.t. κ_j^l, see Lemma 8.2.4. Since we are considering an extended superhedging horizon of $\Xi + \epsilon > \Xi$ the case $\kappa_j^l = \Xi$ is not causing problems, and the case $\kappa_j^l = 0$ can be naturally excluded, see Remark 7.1.5. Then, as in the proof of Theorem 7.2.1 it is clear that the linear system

$$\frac{\partial \mathbb{E}\left[G\left(M_{\tau(\boldsymbol{\kappa})\wedge(\Xi+\epsilon)}\right)\right]}{\partial \kappa_i^l} = \sum_{j=1}^{n} \sum_{m=0}^{1} \tilde{\lambda}_j^m \frac{\partial \mathbb{E}\left[\left(K_j - X_{\tau(\boldsymbol{\kappa})\wedge(\Xi+\epsilon)}\right)^+\right]}{\partial \kappa_i^l} \tag{8.2.20}$$

for $i = 1, \ldots, n$ and $l = 0, 1$, is of triangular form[4] and hence has a unique solution $\tilde{\boldsymbol{\lambda}}^\epsilon \in \mathbb{R}^{2n}$. Set $\lambda_i^\epsilon := \tilde{\lambda}_i^{0,\epsilon} + \tilde{\lambda}_i^{1,\epsilon}$ for $i = 1, \ldots, n$. It further follows analogously as in the proof of Theorem 7.2.1 that $\tau(\boldsymbol{\kappa}) \wedge (\Xi + \epsilon)$ is optimal for $V_\epsilon^{\boldsymbol{\lambda}^\epsilon}$.

The final step is to argue that the (cheapest) superhedging cost for payoff G is $V^{\boldsymbol{\lambda}}(0, X_0, M_0)$. We will achieve this by fixing the market model characterized by

[4] In fact, in the present case of $G(m) = \mathbb{1}_{\{m \geq B\}}$ the partial derivatives w.r.t. κ_i^1, $i = 1, \ldots, n$, and κ_j^0 where $K_j \geq B$ of the left-hand expression of (8.2.20) are zero. We are not simplifying this expression to demonstrate that analogous arguments hold for G satisfying Assumption (iv).

$\tau(\boldsymbol{\kappa}) \wedge \Xi$ and identifying two payoffs G^- and G^+ whose superhedging cost can be made arbitrarily close to $V^{\boldsymbol{\lambda}}(0, X_0, M_0)$ and which satisfy $G^- \leq G \leq G^+$.

From Proposition 8.2.3 we know that we can "superreplicate" $V_\epsilon^{\boldsymbol{\lambda}^\epsilon}$ pathwise with initial cost $V_\epsilon^{\boldsymbol{\lambda}^\epsilon}(0, X_0, X_0) + \delta(\epsilon)$ for arbitrariliy small $\delta(\epsilon) > 0$. By (8.2.19a)–(8.2.19b) we obtain that the strategy $V_\epsilon^{\boldsymbol{\lambda}^\epsilon}$ is dominating $G^{\boldsymbol{\lambda}^\epsilon}(X_t)$ for every $t \leq \Xi$ and its initial cost converges to $V^{\boldsymbol{\lambda}}(0, X_0, M_0)$. It follows by choice of $\boldsymbol{\lambda}^\epsilon$ and the definition of the trading strategy $\tilde{\phi}^\epsilon$ from Proposition 8.2.3 that

$$V_\epsilon^{\boldsymbol{\lambda}^\epsilon}\left(\tau(\boldsymbol{\kappa}) \wedge \Xi, X_{\tau(\boldsymbol{\kappa})\wedge\Xi}, M_{\tau(\boldsymbol{\kappa})\wedge\Xi}\right) = V_\epsilon^{\boldsymbol{\lambda}^\epsilon}(0, X_0, M_0) + \int_0^{\tau(\boldsymbol{\kappa})\wedge\Xi} \tilde{\phi}_t^\epsilon(X)\mathrm{d}X_t.$$

It follows again by choice of $\boldsymbol{\lambda}^\epsilon$ and the definition of $\tilde{\phi}^\epsilon$ that on $\{\tau(\boldsymbol{\kappa}) \wedge \Xi = \tau(\boldsymbol{\kappa}) \wedge (\Xi + \epsilon)\}$,

$$V_\epsilon^{\boldsymbol{\lambda}^\epsilon}\left(\tau(\boldsymbol{\kappa}) \wedge \Xi, X_{\tau(\boldsymbol{\kappa})\wedge\Xi}, M_{\tau(\boldsymbol{\kappa})\wedge\Xi}\right) = G^{\boldsymbol{\lambda}^\epsilon}\left(X_{\tau(\boldsymbol{\kappa})\wedge\Xi}, M_{\tau(\boldsymbol{\kappa})\wedge\Xi}\right).$$

On $\{\tau(\boldsymbol{\kappa}) \wedge \Xi < \tau(\boldsymbol{\kappa}) \wedge (\Xi + \epsilon)\}$ we have $\tau(\boldsymbol{\kappa}) > \Xi$. Denote $\underline{K} := \max_{K_i<B} K_i$. If further $X_\Xi \notin (\underline{K}, B)$ or $M_\Xi \geq B$ then by $\boldsymbol{\kappa}^0, \boldsymbol{\kappa}^1 \in [0, \Xi]^n$ we must have

$$V_\epsilon^{\boldsymbol{\lambda}^\epsilon}(\Xi, X_\Xi, M_\Xi) = G^{\boldsymbol{\lambda}^\epsilon}(X_\Xi, M_\Xi).$$

In the other case, $x = X_\Xi \in (\underline{K}, B)$ and $m = M_\Xi < B$. We compute

$$\begin{aligned} V_\epsilon^{\boldsymbol{\lambda}^\epsilon}(\Xi, x, m) - G^{\boldsymbol{\lambda}^\epsilon}(x, m) &= \mathbb{P}\left[M_{\Xi+\epsilon} \geq B \;\middle|\; X_\Xi = x, M_\Xi = m\right] \\ &=: \tilde{G}^+(x, m) \leq 1. \end{aligned}$$

Construction of G^+ Set

$$G^+(x, m) := G(m) + \tilde{G}^+(x, m)\mathbb{1}_{\{x\in(\underline{K},B),\ x\leq m<B\}} \geq G(m).$$

The preceding arguments show

$$\mathbb{A}\left(G^{+}(S_T,\bar{S}_T);\mathfrak{P}_{\Xi},\mathcal{A},\mathcal{P}\right)=\mathbb{E}\left[G(M_{\tau(\boldsymbol{\kappa})\wedge(\Xi+\epsilon)})\right]=\sup_{\tau\in\mathcal{T}^{\star}_{\Xi+\epsilon}}\mathbb{E}\left[G(M_{\tau})\right]. \tag{8.2.21}$$

Construction of G^- Let $\delta'>0$ and take $\tilde{B}:=B+\delta'$. Repeating the arguments for the barrier level $\tilde{B}$ yields the function G^+ from above with B replaced by $\tilde{B}$. We denote it by $\tilde{G}^-(x,m)$. We have $0\le\tilde{G}^-(x,m)\le G(m)$ for all $x,m\notin[\underline{K},B)$ but $\tilde{G}^-(x,m)\ge G(m)$ for $x,m\in[\underline{K},B)$. Note however that by choosing $\epsilon>0$ small enough we can achieve

$$0\le\sup_{x,m\in[\underline{K},B)}\tilde{G}^-(x,m)<\delta'.$$

This gives rise to the definition

$$G^-(x,m):=\tilde{G}^-(x,m)\mathbb{1}_{\{x,m\notin[\underline{K},B)\}}\le G(m).$$

It follows that for all $\epsilon>0$ small enough,

$$\sup_{\tau\in\mathcal{T}^{\star}_{\Xi+\epsilon}}\mathbb{E}\left[\mathbb{1}_{\left\{M_\tau\ge\tilde{B}\right\}}\right]-\delta'\le\mathbb{A}\left(G^{-}(S_T,\bar{S}_T);\mathfrak{P}_{\Xi},\mathcal{A},\mathcal{P}\right)\le\sup_{\tau\in\mathcal{T}^{\star}_{\Xi+\epsilon}}\mathbb{E}\left[\mathbb{1}_{\left\{M_\tau\ge\tilde{B}\right\}}\right].$$

The proof of (8.2.3) in the case $G(m)=\mathbb{1}_{\{m\ge B\}}$ is complete by noting that

$$\begin{aligned}
\sup_{\tau\in\mathcal{T}^{\star}_{\Xi+\epsilon}}\mathbb{E}\left[\mathbb{1}_{\left\{M_\tau\ge\tilde{B}\right\}}\right]&\longrightarrow\sup_{\tau\in\mathcal{T}^{\star}_{\Xi}}\mathbb{E}\left[\mathbb{1}_{\left\{M_\tau\ge\tilde{B}\right\}}\right]&&\text{as }\epsilon\downarrow0,\\
\sup_{\tau\in\mathcal{T}^{\star}_{\Xi}}\mathbb{E}\left[\mathbb{1}_{\left\{M_\tau\ge\tilde{B}\right\}}\right]&\longrightarrow\sup_{\tau\in\mathcal{T}^{\star}_{\Xi}}\mathbb{E}\left[\mathbb{1}_{\{M_\tau\ge B\}}\right]&&\text{as }\delta'\downarrow0,\ \text{ i.e. }\tilde{B}\downarrow B.
\end{aligned}$$

The argument for payoffs satisfying Assumption (iv) is analogous, but notationally more involved. □

8.2.4 Structure in Optimal Embedding

We conclude this section with a "monotonicity result" regarding the optimal $\boldsymbol{\kappa}^0, \boldsymbol{\kappa}^1, \ldots, \boldsymbol{\kappa}^k$ from Theorem 8.1.5.

Proposition 8.2.5 (Structure in Optimal Embedding)**.** *In the setting of Theorem 8.1.5* $\boldsymbol{\kappa} = (\boldsymbol{\kappa}^0, \boldsymbol{\kappa}^1, \ldots, \boldsymbol{\kappa}^k)$ *can be chosen such that*

$$\kappa_i^j \leq \kappa_i^{j-1} \tag{8.2.22}$$

for all $i = 1, \ldots, n$ *and* $j = 1, \ldots, k$.

Proof. Under our assumptions, we have argued in the proof of Theorem 8.2.1 that it is enough to consider stopping times $\tau = \tau(\boldsymbol{\kappa}) \wedge \Xi$ for some $\boldsymbol{\kappa} \in [0, \Xi]^{n\cdot(k+1)}$ in both stopping problems in (8.2.2). We also proved in Theorem 8.1.5 that $\tau(\boldsymbol{\kappa}) \wedge \Xi$ is optimal for the l.h.s. of (8.2.2) and we proved in the proof of Theorem 8.2.1 existence of $\boldsymbol{\lambda} \in \mathbb{R}^n$ such that $\tau(\boldsymbol{\kappa}) \wedge (\Xi + \epsilon)$ is optimal for $V_\epsilon^{\boldsymbol{\lambda}}$ defined in (8.2.15), see (8.2.20).

Noting that some κ_i^j's are ambiguous, we set $\kappa_i^j := 0$ for i such that $K_i \geq B_{j+1},\ j = 1, \ldots, k-1,\ i = 1, \ldots, n$. In addition we take "regular barriers" $\boldsymbol{\kappa}^j$ in the sense that $\kappa_i^j = 0$ implies $\kappa_l^j = 0$ for $l > i$ if $K_i \geq B$ and $\kappa_i^j = 0$ implies $\kappa_l^j = 0$ for $l < i$ if $K_i \leq B$. Then the claim follows from (8.2.7) applied for $V_\epsilon^{\boldsymbol{\lambda}}$. □

8.3 Lookback Options Revisited

Until now we have considered Lookback payoffs satisfying Assumption (iv) on page 179 and obtained robust pricing and hedging results. A natural extension is to consider general right-continuous, non-decreasing payoff functions G. We did not pursue this question until now because we would like to point out a formal link to an alternative approach based on a recent paper by Beiglböck and Huesmann [7]. The intuition we obtain by formally applying their results gives another explanation for optimality of certain embeddings. However, as will become clear, the arguments in this section are not yet powerful enough to prove e.g. the results of Section 8.1.

8.3.1 Link to the Beiglböck and Huesmann [7] Variational Principle

Based on an idea of Hobson [52, Figure 15], formalized by Beiglböck and Huesmann [7], we recall the notion of *bad-pairs*.

Bad-Pairs Consider the sample space

$$\Big\{(f,s)\ :\ s \geq 0,\ f:[0,s]\to\mathbb{R}_+ \text{ is continuous},\ \ f(0)=S_0\Big\}$$

and a payoff function $\gamma : S \to \mathbb{R}$. The following notion is based on Beiglböck and Huesmann [7, Section 1.3]. Suppose that $\hat{\tau}$ stops a path once it reaches (g,t) and some other path is still living in (f,s). Assume also that

$f(s) = g(t)$,

" stop in (f,s), do not stop in (g,t)" leads to a better payoff.

Then we will say that $\big((f,s),(g,t)\big)$ is a *bad pair* relative to γ.

Formal Application of the Variational Principle The Variational Principle of Beiglböck and Huesmann [7, Theorem 1.3] shows that a necessary feature of an optimal embedding is that locally at level $X = K_i$ there are no *bad pairs*.

Consider the payoff function $\gamma(f,s) = G\left(\max_{0\le u\le s\wedge\Xi} f(u)\right)$ where $G \ge 0$ is right-continuous and non-decreasing, and note

$$\sup_{\tau\in\mathcal{T}^\star_\Xi} \mathbb{E}\left[G\left(\max_{u\le\tau} X_u\right)\right] \le \sup_{\tau\in\mathcal{T}^\star_\infty} \mathbb{E}\left[G\left(\max_{u\le\tau\wedge\Xi} X_u\right)\right] = \sup_{\tau\in\mathcal{T}^\star_\infty} \mathbb{E}\left[\gamma(X_\tau,\tau)\right]. \quad (8.3.1)$$

Now, applying the variational principle of Beiglböck and Huesmann [7, Theorem 1.3] formally for X being a Geometric Brownian motion and $\mu = \mu^\star_N$ yields the following heuristics.

From the no *bad-pair* assertion and the atomic nature of the target measure $\mu^\star_N$ we deduce that it is not optimal (or does not make a difference) to stop paths at $x \ne K_i$. Therefore, it remains to look at the stopping rule at K_i. The following properties appear natural:

(a) at each K_i there is a non-increasing stopping boundary d_i. We illustrate this intuition in Figure 8.3.1.

(b) there is no benefit in continuing after Ξ.

Hence, if there exists an embedding $\tau \le \Xi$, i.e. $\mathcal{T}^\star_\Xi \ne \emptyset$, then for the optimal stopping time $\tau^\star$ there will be no crossing of K_i after Ξ — i.e. $\tau^\star \wedge \Xi \in \mathcal{T}^\star_\Xi$. In other words, if $\mathcal{T}^\star_\Xi \ne \emptyset$, then a solution to $\sup_{\tau\in\mathcal{T}^\star_\infty} \mathbb{E}\left[\gamma(X_\tau,\tau)\right]$ is also a solution to $\sup_{\tau\in\mathcal{T}^\star_\Xi} \mathbb{E}\left[G(\bar{X}_\tau)\right]$, i.e. there is equality in (8.3.1).

Remark 8.3.1 (Convex Vanilla Option). There is an analogous argument for the convex Vanilla option case, encoded by the choice $\gamma(f,s) = G\left(f(s\wedge\Xi)\right)$ which boils down to the *bad-pair* argument given by Hobson [52, Figure 15].

8.3.2 Robust Pricing and Hedging of Lookback Options

The heuristics in the previous section motivate the following class of stopping times.

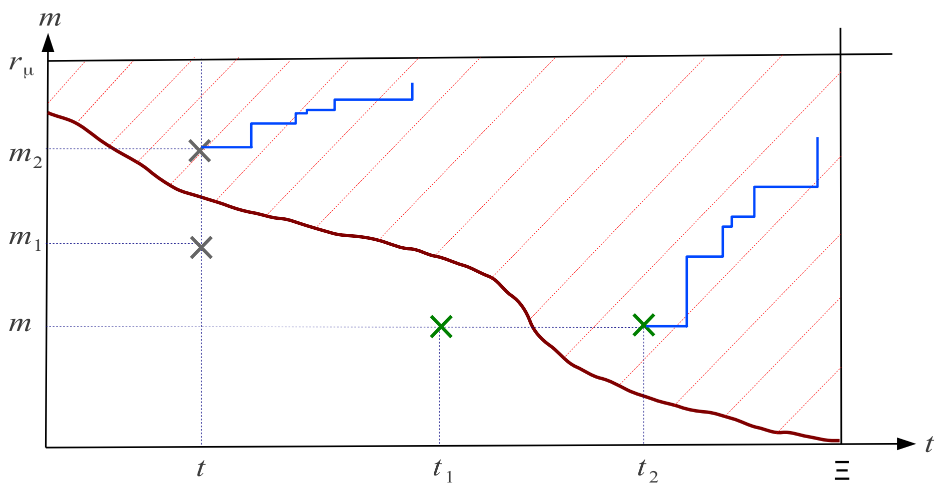

FIGURE 8.3.1: WLOG a path is only stopped at $X = K_i$. The above picture illustrates the optimal stopping rule in that case. Firstly, consider the two grey points $p_1 = (t, m_1), p_2 = (t, m_2)$. Let us assume that a path continues from p_2 but stops at p_1 (as indicated in the picture above). Then we can just "swap these paths", i.e. let the process continue from p_1 but stop it at p_2. The result of this is that regardless what the path does, we achieve at least the same payoff as before. Secondly, consider the two green points $q_1 = (t_1, m), q_2 = (t_2, m)$ and assume that a path continues from q_2 but stops at q_1 (as indicated in the picture above). Then we can "swap the paths at q_1 and q_2" as before without affecting the payoff. These arguments suggest that an optimal embedding is determined by a non-decreasing boundary $\boldsymbol{d}$.

Definition 8.3.2 (Optimal Stopping Boundaries)**.** *Define*

$$\mathscr{D} := \Big\{ \boldsymbol{d} = (d_1, \dots, d_n) \ : d_i : [0,\infty] \to [0,\infty], \quad d_i \ \textit{non-increasing and right-continuous} \Big\}. \tag{8.3.2}$$

For $\boldsymbol{d} \in \mathscr{D}$ we define the stopping time

$$\tau(\boldsymbol{d}) := \inf \big\{ u > 0 \ : \ X_u = K_i, \ u \geq d_i(M_u) \ \textit{ for some } i = 1, \dots, n \big\} \tag{8.3.3}$$

and denote

$$\mathcal{T}^{\star}_{\Xi}(\mathscr{D}) := \big\{ \tau(\boldsymbol{d}) \wedge \Xi \ : \ \boldsymbol{d} \in \mathscr{D} \big\} \cap \mathcal{T}^{\star}_{\Xi}. \tag{8.3.4}$$

We obtain the following result for Lookback options.

Theorem 8.3.3 (Robust Pricing and Hedging of Lookback Options)**.** *Let $G \geq 0$ be non-decreasing and right-continuous and $\mathbb{E}[G(M_\Xi)] < \infty$. Then,*

$$\sup_{\tau \in \mathcal{T}^\star_\Xi} \mathbb{E}[G(M_\tau)] = \sup_{\tau \in \mathcal{T}^\star_\Xi(\mathscr{D})} \mathbb{E}[G(M_\tau)] = \mathbb{E}\left[G\left(M_{\tau(\boldsymbol{d}^\star)\wedge\Xi}\right)\right] \tag{8.3.5}$$

for some $\boldsymbol{d}^\star \in \mathscr{D}$ and there is no duality gap, i.e.

$$\mathbb{A}(G(\bar{S}_T); \mathfrak{P}_\Xi, \mathcal{A}, \mathcal{P}) = \sup_{\tau \in \mathcal{T}^\star_\Xi} \mathbb{E}[G(M_\tau)] = \sup_{\mathbb{P}^S} \mathbb{E}_{\mathbb{P}^S}\left[G(\bar{S}_T)\right] \tag{8.3.6}$$

where the last supremum is over all $(\mathfrak{P}_\Xi, \mathcal{Y}, \mathcal{P})$-market models $\left(\Omega^S, \mathcal{F}^S, \mathbb{F}^S, \mathbb{P}^S\right)$.

The proof of Theorem 8.3.3 will be developed in the next sections.

If $\Xi = \infty$, our class of embeddings contains both the Root and the Azéma-Yor embeddings.

Example 8.3.4 (Root Embedding)**.** *If we let*

$$d_i(m) = d_i^{\mathrm{R}}(m) = b_i \tag{8.3.7}$$

then we recover the classical Root embedding τ_{R}. This, in fact, shows

$$\mathcal{T}^\star_\Xi(\mathscr{D}) = \emptyset \quad \Longleftrightarrow \quad \mathcal{T}^\star_\Xi = \emptyset. \tag{8.3.8}$$

Example 8.3.5 (Azéma-Yor Embedding)**.** *If we let*

$$d_i = d_i^{\mathrm{AY}}(m) = \infty \cdot \mathbb{1}_{\{m < b_\mu(K_i)\}} \tag{8.3.9}$$

then we recover the Azéma-Yor embedding τ_{AY} where b_μ denotes the barycenter function of μ.

8.3.3 A Class of Root Barriers

We extend the notion of Root barrier from Root [85, Definition 2.3].

Definition 8.3.6 (Root Barrier). *A subset* $\mathcal{B} \subseteq [0,\infty] \times [0,\infty] \times [M_0,\infty]$ *is called barrier for the Markov process* $(t, X_t, M_t)_{0\le t\le \Xi}$ *if*

(a) $\mathcal{B}$ *is closed,*

(b) if $(t,x,m) \in \mathcal{B}$ *then* $(s,x,m) \in \mathcal{B}$ *whenever* $s > t$*,*

(c) $\infty \times [0,\infty] \times [M_0,\infty] \subseteq \mathcal{B}$*,*

(d) $[0,\infty] \times 0 \times [M_0,\infty] \subseteq \mathcal{B}$*,* $\quad [0,\infty] \times \infty \times [M_0,\infty] \subseteq \mathcal{B}$*,*

(e) $[0,\infty] \times [0,\infty] \times M_0 \subseteq \mathcal{B}$*,* $\quad [0,\infty] \times [0,\infty] \times \infty \subseteq \mathcal{B}$*.*

We denote the set of barriers by $\mathscr{B} := \{\mathcal{B} \text{ is a barrier}\}$*. For convenience set*

$$\mathcal{B}_0 := \infty \times [0,\infty] \times [M_0,\infty] \cup [0,\infty] \times 0 \times [M_0,\infty] \cup [0,\infty] \times \infty \times [M_0,\infty] \cup [0,\infty] \times [0,\infty] \times M_0 \cup [0,\infty] \times [0,\infty] \times \infty.$$

To each $\boldsymbol{d} \in \mathscr{D}$ we can uniquely associate a barrier $\mathcal{B}(\boldsymbol{d}) \in \mathscr{B}$ in the following way,

$$\begin{aligned} \mathcal{B}_i(d_i) &:= \left\{(t,x,m) \in [0,\infty] \times [0,\infty] \times [M_0,\infty] \;:\; x = K_i,\ t \ge d_i(m) \right\}, \\ \mathcal{B}(\boldsymbol{d}) &:= \left(\bigcup_{i=1}^{n} \mathcal{B}_i(d_i)\right) \cup \mathcal{B}_0. \end{aligned} \tag{8.3.10}$$

Inspired by Root [85] we define the following metric which will make the set of barriers compact. We map $H := [0,\infty] \times [0,\infty] \times [M_0,\infty]$ homeomorphically to a bounded cuboid F by

$$(t,x,m) \mapsto \left(\frac{1}{1+t}, \frac{1}{1+x}, \frac{1}{1+m}\right). \tag{8.3.11}$$

Let F be endowed with the standard Euclidean metric $|\cdot|$ and H the induced metric r^{ind}. On the space of closed subsets of H, denoted by $\mathscr{C}$, define a metric by

$$r(C,D) := \max\left(\sup_{y\in C} r^{\mathrm{ind}}(x,D), \sup_{z\in D} r^{\mathrm{ind}}(z,C)\right) \tag{8.3.12}$$

As in Root [85] we obtain: "under r, $\mathscr{C}$ is a separable metric space and $\mathscr{B}$, the space of all barriers is closed in $\mathscr{C}$ and hence compact".

Lemma 8.3.7. *For any $\epsilon > 0$ there exists $\delta > 0$ such that if $\boldsymbol{d}^1, \boldsymbol{d}^2 \in \mathscr{D}$ satisfy $d\big(\mathcal{B}(\boldsymbol{d}^1), \mathcal{B}(\boldsymbol{d}^2)\big) < \delta$ then*

$$\mathbb{P}\left[\left|\tau(\boldsymbol{d}^1)\wedge\Xi - \tau(\boldsymbol{d}^2)\wedge\Xi\right| > \epsilon\right] < \epsilon. \tag{8.3.13}$$

Proof. By definition, both $\tau(\boldsymbol{d}^1)$ and $\tau(\boldsymbol{d}^2)$ only stop at $X = K_i$ for some $i = 1,\ldots,n$. For any $\epsilon' > 0$, by choosing $\delta > 0$ small enough and $\bar{m} > M_0$ sufficiently large, we can achieve

$$\sup_{m\in[K_i,\bar{m}]} |d_i^1(m) - d_i^2(m)| < \epsilon', \qquad \forall i = 1,\ldots,n,$$

$$\mathbb{P}\left[M_\Xi \geq \bar{m}\right] < \epsilon'.$$

From this it follows that (8.3.13) holds. □

8.3.4 Proof of Theorem 8.3.3

Proof of Theorem 8.3.3. For every $\epsilon > 0$ we can approximate G by a non-decreasing, right-continuous, pricewise-constant function G^ϵ such that $0 \leq G^\epsilon - G < \epsilon$. From Proposition 8.2.5 we infer that the optimal $\boldsymbol{\kappa}^\epsilon$ for the robust pricing problem with payoff G^ϵ gives rise to an element $\boldsymbol{d}^\epsilon \in \mathscr{D}$, i.e. $\tau(\boldsymbol{\kappa}^\epsilon)\wedge\Xi = \tau(\boldsymbol{d}^\epsilon)\wedge\Xi$. In Section 8.3.3 we obtained that the space of all barriers is compact. Therefore, the limit of $(\boldsymbol{d}^\epsilon)$ exists along some subsequence and we denote it by $\boldsymbol{d}^\star$. It is clear that

$\boldsymbol{d}^\star \in \mathscr{D}$. Consequently, by dominated convergence,

$$\sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}\left[G(M_\tau)\right] = \lim_{\epsilon \downarrow 0} \sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}\left[G^\epsilon(M_\tau)\right] = \lim_{\epsilon \downarrow 0} \mathbb{E}\left[G^\epsilon\left(M_{\tau(\boldsymbol{\kappa}^\epsilon)\wedge\Xi}\right)\right] = \mathbb{E}\left[G\left(M_{\tau(\boldsymbol{d}^\star)\wedge\Xi}\right)\right]$$

where the last equality sign holds by Lemma 8.3.7 and our assumption that $\mathbb{E}\left[G(M_\Xi)\right] < \infty$. This proves (8.3.5).

As for (8.3.6) we note by Theorem 8.2.1 that

$$\begin{aligned}\sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}\left[G(M_\tau)\right] &\leq \mathbb{A}(G(\bar{S}_T); \mathfrak{P}_\Xi, \mathcal{A}, \mathcal{P}) \leq \mathbb{A}(G^\epsilon(\bar{S}_T); \mathfrak{P}_\Xi, \mathcal{A}, \mathcal{P}) \\ &= \sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}\left[G^\epsilon(M_\tau)\right] \longrightarrow \sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}\left[G(M_\tau)\right]\end{aligned}$$

as $\epsilon \to 0$. □

8.4 Forms of Arbitrage

Theorem 8.4.1 (Arbitrage Relative to $(\mathfrak{P}_\Xi, \mathcal{A}, \mathcal{P})$)**.** *In the setting of Theorem 7.3.2, assume $\mathcal{T}_\Xi^\star \neq \emptyset$. Suppose G is as in Theorem 8.3.3. We add the option $G(M_T)$ to $\mathcal{A}$ and extend $\mathcal{P}\left[G(M_T)\right] := p_G$.*

Then, there is arbitrage relative to $(\mathfrak{P}_\Xi, \mathcal{A}, \mathcal{P})$ if $p_G > \sup_{\tau \in \mathcal{T}_\Xi^\star} \mathbb{E}\left[G(M_\tau)\right]$.

Proof. This can be argued as in the proof of Theorem 7.3.2 by using the results of Section 8.2. □

8.5 Numerical Methods

In this section we explain which numerical methods could be applied to compute the optimal embeddings and static positions, both in the convex Vanilla and

Lookback case. We omit here a rigorous and more complete treatment of important questions regarding well-posedness, existence, uniqueness and numerical implementation as this is not the scope of this thesis.

For a some numerical examples for the convex Vanilla option we refer to Mykland [70, pp. 424–425].

8.5.1 Computation of the Optimal Embedding

8.5.1.1 Convex Vanilla Options

Choose M sufficiently large. In Section 7.1 we have shown that the robust pricing problem for a convex Vanilla option is solved by $\tau_{\mathrm{R}}(\mu^\star_M) \wedge \Xi = \tau(\boldsymbol{b}^M) \wedge \Xi$. We can directly compute $\mu^\star_M$. Then, following Cox and Wang [24] or Oberhauser and dos Reis [72] we first find v such that

$$\min\left\{ v - U_{\mu^\star_M}, \frac{\partial v}{\partial \xi} - \frac{x^2}{2}\frac{\partial^2 v}{\partial x^2} \right\} = 0, \qquad \text{on} \quad [0,\Xi] \times \mathbb{R}_+ \tag{8.5.1}$$

and then $\boldsymbol{b}$ is given by

$$b_i := \inf\left\{ t \leq \Xi \ : \ U_{\mu^\star_M}(K_i) = v(t, K_i) \right\} \wedge \Xi. \tag{8.5.2}$$

8.5.1.2 Simple Barrier Option

Recalling the setting and statement of Theorem 8.1.1, we have

$$\sup_{\tau \in \mathcal{T}^\star_\Xi} \mathbb{E}\left[\mathbb{1}_{\{M_\tau \geq B\}}\right] = \sup_{\boldsymbol{\kappa} \in [0,\Xi]^{2n} \ : \ \tau(\boldsymbol{\kappa}) \wedge \Xi \in \mathcal{T}^\star_\Xi} \mathbb{E}\left[\mathbb{1}_{\left\{M_{\tau(\boldsymbol{\kappa})\wedge\Xi} \geq B\right\}}\right] \tag{8.5.3}$$

$$= \lim_{N\to\infty} \sup_{\boldsymbol{\kappa} \in [0,\Xi]^{2n}} \left\{ \mathbb{P}\left[M_{\tau(\boldsymbol{\kappa})\wedge\Xi} \geq B\right] - N \cdot \sum_{i=1}^n \left(\mathbb{E}\left[\left(X_{\tau(\boldsymbol{\kappa})\wedge\Xi} - K_i\right)^+\right] - c^\star(K_i)\right)^2 \right\}$$

by the standard "penalization argument". Hence, in principle we can use (8.5.3) to compute an optimal $\boldsymbol{\kappa} = (\boldsymbol{\kappa}^0, \boldsymbol{\kappa}^1)$ using any optimization method.

In order to compute $\mathbb{P}\left[M_{\tau(\boldsymbol{\kappa})\wedge\Xi} \geq B\right]$ and $\mathbb{E}\left[\left(X_{\tau(\boldsymbol{\kappa})\wedge\Xi} - K_i\right)^+\right]$ in (8.5.3) one could first compute

$$c_\xi^B(x) := \mathbb{E}\left[\left(X_{\tau(\boldsymbol{\kappa}^0)\wedge H_B\wedge\xi} - x\right)^+\right] \tag{8.5.4}$$

by PDE methods along the lines of Cox and Wang [24] and Oberhauser and dos Reis [72] for $0 \leq \xi \leq \Xi$ and $x \geq 0$. Then this yields

$$\mathbb{P}\left[M_{\tau(\boldsymbol{\kappa})\wedge\xi} \geq B\right] = \mathbb{P}\left[X_{\tau(\boldsymbol{\kappa}^0)\wedge H_B\wedge\xi} = B\right] = -\frac{\partial c_\xi^B}{\partial x}(B-).$$

Further

$$\begin{aligned} \mathbb{E}\left[\left(X_{\tau(\boldsymbol{\kappa})\wedge\Xi} - K_i\right)^+\right] &= c_\Xi^B(K_i) + \mathbb{E}\left[\left(X_{\tau(\boldsymbol{\kappa})\wedge\Xi} - K_i\right)^+ \mathbb{1}_{\{\tau(\boldsymbol{\kappa}^0)\wedge\Xi > H_B\}}\right] \\ = \quad c_\Xi^B(K_i) &+ \int_0^\Xi \left\{\mathbb{P}\left[\tau(\boldsymbol{\kappa}^0)\wedge\Xi = H_B \in \mathrm{d}\xi\right]\mathbb{E}\left[\left(X_{\tau(\boldsymbol{\kappa}^1)\wedge\Xi} - K_i\right)^+ \middle| X_\xi = B\right]\right\} \\ -(B - K_i)^+ \mathbb{P}\left[\tau(\boldsymbol{\kappa})\wedge\Xi = H_B\right]. \end{aligned} \tag{8.5.5}$$

The point about (8.5.5) is that all terms can be computed by PDE methods as outlined above.

8.5.2 Computation of the Static Position

By means of dynamic programming and the well-known link between optimal stopping problems and variational inequalities, see Bensoussan and Lions [11], $V^{\boldsymbol{\lambda}}$

from (7.2.5) and (8.2.4) is expected to solve[5]

$$\min\left\{V^{\boldsymbol{\lambda}}(t,\cdot,\cdot)-G^{\boldsymbol{\lambda}},-\left(\frac{\partial V^{\boldsymbol{\lambda}}}{\partial t}+\frac{1}{2}x^2\frac{\partial^2 V^{\boldsymbol{\lambda}}}{\partial x^2}\right)\right\}=0, \tag{8.5.6a}$$

$$\frac{\partial V^{\boldsymbol{\lambda}}}{\partial m}(t,m,m)=0, \tag{8.5.6b}$$

on

$$\Delta:=\left\{(t,x,m)\in(0,\Xi)\times\mathbb{R}_+^2 : x\le m\right\}. \tag{8.5.7}$$

In the Vanilla case $V^{\boldsymbol{\lambda}}(t,x,m)=V^{\boldsymbol{\lambda}}(t,x)$ and hence (8.5.6a)–(8.5.6b) simplify. Recall (7.2.2), (8.2.2) and observe that

$$\begin{aligned}\boldsymbol{\lambda}\mapsto\quad & V^{\boldsymbol{\lambda}}(0,X_0,M_0)+\sum_{i=1}^{n}\lambda_i p^\star(K_i) \\ &=\sup_{\tau\in\mathcal{T}_\Xi}\mathbb{E}\left[G(X_\tau,M_\tau)-\sum_{i=1}^{n}\lambda_i\left(K_i-X_\tau\right)^+\right]+\sum_{i=1}^{n}\lambda_i p^\star(K_i)\end{aligned} \tag{8.5.8}$$

is convex. This suggests that any optimization algorithm can be applied to solve the problem. More precisely, in each iteration we would need to compute $V^{\boldsymbol{\lambda}}$ by solving the free-boundary problem of the form (8.5.6a)–(8.5.6b).

In this context we refer to the related work of Tan and Touzi [93, Sec. 5].

[5] We omit a rigorous treatment of existence and uniqueness of solutions to (8.5.6a)–(8.5.6b)

Appendix A

Appendix: Part I

A.1 Appendix: Proof of Lemma 2.1.14

In order to prove Lemma 2.1.14 we require to prove, inductively, several auxiliary results along the way. We now state and prove a Lemma which contains the statement of Lemma 2.1.14.

Lemma A.1.1. *Let $n \in \mathbb{N}$ and let Assumption $\circledast$ hold.*

Then

$$y \mapsto K_n(y) \qquad \text{is absolutely continuous and non-increasing.} \tag{A.1.1}$$

If we assume in addition that the embedding property of Theorem 2.1.5 is valid for the first $n-1$ marginals then for almost all $y \geq 0$ we have:

If $\xi_n(y) < y$ then

$$K_n'(y) + \frac{K_n(y)}{y - \xi_n(y)} = K_{\jmath_n(y)}'(y) + \frac{K_{\jmath_n(y)}(y)}{y - \xi_n(y)} \tag{A.1.2}$$

where K'_j *denotes the derivative of* K_j *which exists for almost all* $y \geq 0$ *and* $j = 1, \ldots, n$.

If $\xi_n(y) = y$ *then*

$$K_n(y+) = K_{\jmath_n(y)}(y+). \tag{A.1.3}$$

For $x \in \mathbb{R}$ *the mapping*

$$c^n(x, \cdot) : [0, \infty) \to \mathbb{R}, \quad y \mapsto c^n(x, y) \tag{A.1.4}$$

is locally Lipschitz continuous, non-decreasing and for almost all $y \geq 0$

$$\left.\frac{\partial c^n}{\partial y}(x, y)\right|_{x=\xi_n(y)} = K_{\jmath_n(y)}(y) + (y - \xi_n(y))K'_{\jmath_n(y)}(y). \tag{A.1.5}$$

The mapping $c^n(\cdot, y)$ *is locally Lipschitz continuous and if it is differentiable at* $\xi_n(y)$ *and* $\xi'_n(y) > 0$ *then for almost all* $y \geq 0$

$$K_n(y) + c'_n(\xi_n(y)) - c'_j(\xi_n(y)) - K_j(y) = 0 \tag{A.1.6}$$

for $j = \jmath_n(y)$ *and* j *such that* $n > j > \jmath_n(y)$ *and* $\xi_n(y) = \xi_j(y)$.

In the case of non-smoothness of $c^n(\cdot, y)$ *at* $\xi_n(y)$ *we have either*

$$\xi'_n(y) = 0 \tag{A.1.7}$$

or

$$K_n(y) + c'_n(\xi_n(y)+) - c'_j(\xi_n(y)+) - K_j(y) = 0 \tag{A.1.8}$$

for $j = \jmath_n(y)$ *and* j *such that* $n > j > \jmath_n(y)$ *and* $\xi_n(y) = \xi_j(y)$.

Proof. We prove the claim by induction over n. The induction basis $n = 1$ holds by definition and Lemma 2.6 of Brown et al. [18].

Now assume that the claim holds for all $i = 1, \dots, n-1$.

Induction step for c^n. We have

$$\begin{aligned} c^n(x, y+\delta) - c^n(x, y) = &- \left[c_{\imath_n(x;y+\delta)}(x) - (y+\delta-x)K_{\imath_n(x;y+\delta)}(y+\delta)\right] \\ &+ \left[c_{\imath_n(x;y)}(x) - (y-x)K_{\imath_n(x;y)}(y)\right]. \end{aligned} \tag{A.1.9}$$

Firstly, consider the case when there exists a $\delta' > 0$ such that for all $|\delta| < \delta'$ we have $\imath_n(x; y) = \imath_n(x; y+\delta)$. Equation (A.1.9) simplifies and we have

$$\begin{aligned} & c^n(x, y+\delta) - c^n(x, y) = (y+\delta-x)K_{\imath_n(x;y)}(y+\delta) - (y-x)K_{\imath_n(x;y)}(y) \\ = \;& (y+\delta-\xi_{\imath_n(x;y)}(y))K_{\imath_n(x;y)}(y+\delta) - (y-\xi_{\imath_n(x;y)}(y))K_{\imath_n(x;y)}(y) \\ & +(x-\xi_{\imath_n(x;y)}(y))\left[K_{\imath_n(x;y)}(y) - K_{\imath_n(x;y)}(y+\delta)\right] \\ \overset{(2.1.12)}{\leq} \;& c^{\imath_n(x;y)}(\xi_{\imath_n(x;y)}(y), y+\delta) - c^{\imath_n(x;y)}(\xi_{\imath_n(x;y)}(y), y) \\ & +(y-\xi_{\imath_n(x;y)}(y))\left[K_{\imath_n(x;y)}(y) - K_{\imath_n(x;y)}(y+\delta)\right] \\ \leq \;& \max_{i<n}\left\{c^i(\xi_i(y), y+\delta) - c^i(\xi_i(y), y) + (y-\xi_i(y))\left[K_i(y) - K_i(y+\delta)\right]\right\} \\ \leq \;& \mathrm{const}(y)\cdot|\delta| \end{aligned} \tag{A.1.10}$$

by induction hypothesis and where const(y) denotes a constant depending on y. A similar computation shows for $|\delta|$ small enough

$$\begin{aligned}
&c^n(x,y) - c^n(x,y+\delta) \\
\overset{(2.1.12)}{\leq}\ & c^{\imath_n(x;y)}(\xi_{\imath_n(x;y)}(y+\delta), y) - c^{\imath_n(x;y)}(\xi_{\imath_n(x;y)}(y+\delta), y+\delta) \\
&+(x - \xi_{\imath_n(x;y)}(y+\delta)) \left[K_{\imath_n(x;y)}(y+\delta) - K_{\imath_n(x;y)}(y)\right] \\
=\ & (y - \xi_{\imath_n(x;y)}(y+\delta)) \left[K_{\imath_{\imath_n(x;y)}(x;y)}(y) - K_{\imath_{\imath_n(x;y)}(x;y)}(y+\delta)\right] - \delta K_{\imath_{\imath_n(x;y)}(x;y)}(y+\delta) \\
&+(x - \xi_{\imath_n(x;y)}(y+\delta)) \left[K_{\imath_n(x;y)}(y+\delta) - K_{\imath_n(x;y)}(y)\right] \\
\leq\ & \begin{cases} \text{const}(y) \cdot |\delta| & \text{if } \delta < 0, \\ 0 & \text{if } \delta \geq 0, \end{cases}
\end{aligned} \tag{A.1.11}$$

again by induction hypothesis and continuity of ξ_i which indeed allows the constant to be chosen independently of δ. Monotonicity of $c^n(x,\cdot)$ follows. Equation (A.1.11) together with (A.1.10) imply the local Lipschitz continuity. Plugging $x = \xi_n(y)$ into (A.1.9), direct computation shows that (A.1.5) holds.

Secondly, consider the case when $\imath_n(x;\cdot)$ jumps at y. Recall (2.1.46). Note that this is only possible when x satisfies

$$x = \xi_{\imath_n(x;y-\delta)}(y) \qquad \text{for } \delta > 0 \text{ small enough,} \tag{A.1.12}$$

i.e. when $x = \xi_k(y)$ for the index $k = \imath_n(x; y-\delta) > \jmath_n(y)$. By (2.1.46) there exists a $\delta' > 0$ such that $\imath_n(x; y+\delta) = \imath_n(x;y)$ for all $0 \leq \delta < \delta'$. Hence, for $\delta > 0$ small enough, $|c^n(x,y+\delta) - c^n(x,y)|$ has the same upper bound as in the first case. Monotonicity of $c^n(x,\cdot)$ follows.

Furthermore, for $\delta > 0$ we have for the x from (A.1.12) that $\imath_n(x; y-\delta) > \imath_n(x; y)$ holds. For notational simplicity we only consider the case

$$\imath_{\imath_n(x;y-\delta)}(x; y-\delta) = \imath_n(x; y).$$

The general case follows by the same arguments. Observing

$$\left|c^n(x,y) - c^n(x,y-\delta)\right| \overset{\text{(A.1.12)}}{=} \left| -(y-\delta-x)K_{\imath_n(x;y-\delta)}(y-\delta) + (y-x)K_{\imath_n(x;y-\delta)}(y)\right|$$

and using the induction hypothesis establishes the claim regarding monotonicity of $c^n(x, \cdot)$ and local Lipschitz continuity.

We prove (A.1.5) by computing the required right- and left-derivative of $c^n(x, \cdot)$ at $x = \xi_n(y)$. The right-derivative is simply, using (2.1.46) and (A.1.9),

$$K_{\jmath_n(y)}(y) + (y - \xi_n(y))K'_{\jmath_n(y)}(y) \tag{A.1.13}$$

and the left-derivative is, writing $k = \imath_n(\xi_n(y); y-) > \imath_n(\xi_n(y); y) = \jmath_n(y)$,

$$\begin{aligned}
&\lim_{\delta\uparrow 0} \frac{1}{\delta}\Big(- c_k(\xi_n(y)) + (y+\delta-\xi_n(y))K_k(y+\delta) \\
&\qquad + c_{\jmath_n(y)}(\xi_n(y)) - (y-\xi_n(y))K_{\jmath_n(y)}(y)\Big) \\
\overset{\xi_n(y)=\xi_k(y)}{=} \quad & \lim_{\delta\uparrow 0} \frac{1}{\delta}\Big((y+\delta-\xi_n(y))K_k(y+\delta) - (y-\xi_n(y))K_k(y)\Big) \\
= \quad & K_k(y) + (y-\xi_n(y))K'_k(y) \overset{\text{(A.1.2)}}{=} K_{\jmath_n(y)}(y) + (y-\xi_n(y))K'_{\jmath_n(y)}(y)
\end{aligned} \tag{A.1.14}$$

by induction hypothesis. So the two coincide for almost all $y \geq 0$.

Induction step for K_n. A straightforward computation shows that the mapping $y \mapsto \frac{c^n(x,y)}{y-x}$ is non-increasing and hence for $\delta > 0$

$$K_n(y+\delta) = \inf_{\zeta \leq y+\delta} \frac{c^n(\zeta, y+\delta)}{y+\delta-\zeta} \leq \inf_{\zeta \leq y} \frac{c^n(\zeta, y+\delta)}{y+\delta-\zeta} \leq \inf_{\zeta \leq y} \frac{c^n(\zeta, y)}{y-\zeta} = K_n(y)$$

proving that K_n is non-increasing.

Using again that $c^n(x, \cdot)$ is non-decreasing and that ξ_n is continuous, local Lipschitz continuity of K_n now follows from

$$K_n(y) \leq \frac{c^n(\xi_n(y+\delta), y)}{y - \xi_n(y+\delta)} \leq \frac{c^n(\xi_n(y+\delta), y+\delta)}{y - \xi_n(y+\delta)} = K_n(y+\delta)\left(1 + \frac{\delta}{y - \xi_n(y+\delta)}\right)$$

if $\xi_n(y) < y$ and if $\xi_n(y) = y$, recalling (2.1.13), we have

$$\begin{aligned} K_n(y+) = & \inf_{\zeta \leq (y+)} \left\{ \frac{c_n(\zeta) - c_{\imath_n(\zeta; y+)}(\zeta)}{(y+) - \zeta} + K_{\imath_n(\zeta; y+)}(y) \right\} \\ \overset{\text{Lemma 2.1.12}}{=} & \inf_{y \leq \zeta \leq (y+)} \left\{ \frac{c_n(\zeta) - c_{\imath_n(\zeta; y+)}(\zeta)}{(y+) - \zeta} + K_{\imath_n(\zeta; y+)}(y+) \right\} = K_{\imath_n(y;y)}(y+) \\ = & K_{\jmath_n(y)}(y+), \end{aligned}$$

and local Lipschitz continuity of K_n follows by induction hypothesis. Equation (A.1.3) is also proven.

Local Lipschitz continuity of $c^n(\cdot, y)$ follows from the properties of $\imath_n$, cf. (2.1.45), the fact that the functions $c_i, i = 1, \dots, n$, are locally Lipschitz and a similar expansion of terms in the case when $\xi_n(y) = \xi_i(y)$ for some $i < n$.

In order to prove (A.1.6) we first exclude all $y \geq 0$ such that $\xi_n(y)$ is an atom of any $c_1, \dots, c_n$. Amongst all $y \in \{\xi_n' > 0\}$ this is a null-set. By assumption $c^n(\cdot, y)$ is differentiable at $\xi_n(y)$. Using the optimality of ξ_n, direct computation of the left- and right-derivatives proves (A.1.6) for $j = \jmath_n(y)$ and $k = \imath_n(\xi_n(y)+; y)$. Now we want to apply the induction hypothesis to c^k. By choice of y we have that c_k is differentiable at $\xi_n(y) = \xi_k(y)$, i.e. μ_k does not have an atom at $\xi_n(y)$. Hence, by the assumption that the embedding for the first $n-1$ marginals is valid we cannot have $\xi_k'(y) = 0$ (except on a null-set because otherwise the embedding would fail). By (A.1.7) and choice of y, $c^k(\cdot, y)$ therefore has to be differentiable at $\xi_k(y)$. This shows that we can indeed apply (A.1.6) to deduce for $j = \jmath_n(y) = \jmath_k(y)$ and j such that $n > k > j > \jmath_n(y)$ and $\xi_n(y) = \xi_k(y) = \xi_j(y)$ the following equation,

$$\begin{aligned}
0 &= K_n(y) + c_n'(\xi_n(y)) - c_k'(\xi_n(y)) - K_k(y) \\
&= K_n(y) + c_n'(\xi_n(y)) - c_k'(\xi_n(y)) + c_k'(\xi_n(y)) - c_j'(\xi_n(y)) - K_j(y) \\
&= K_n(y) + c_n'(\xi_n(y)) - c_j'(\xi_n(y)) - K_j(y).
\end{aligned}$$

Equation (A.1.6) is proven.

For later use we note the equation

$$\frac{c_n(\xi_n(y)) - c_{\jmath_n(y)}(\xi_n(y))}{y - \xi_n(y)} + K_{\jmath_n(y)}(y) = K_n(y) = \frac{c_n(\xi_n(y)) - c_k(\xi_n(y))}{y - \xi_n(y)} + K_k(y) \tag{A.1.15}$$

for k such that $n > k > \jmath_n(y)$ and $\xi_n(y) = \xi_k(y)$.

Finally, we prove the claimed ODE for K_n in the case $\xi_n(y) < y$. For almost all $y \geq 0$ we have

$$\begin{aligned}
K_n'(y) &= \lim_{\delta \to 0} \frac{1}{\delta} \left[\frac{c^n(\xi_n(y+\delta), y+\delta)}{y + \delta - \xi_n(y+\delta)} - \frac{c^n(\xi_n(y), y)}{y - \xi_n(y)} \right] \\
&= \lim_{\delta \to 0} \frac{1}{\delta} \left[\left(\frac{1}{y + \delta - \xi_n(y+\delta)} - \frac{1}{y - \xi_n(y)} \right) c^n(\xi_n(y+\delta), y+\delta) \right. \\
&\qquad \left. + \frac{c^n(\xi_n(y+\delta), y+\delta) - c^n(\xi_n(y), y)}{y - \xi_n(y)} \right] \\
&= \frac{\xi_n'(y) - 1}{y - \xi_n(y)} K_n(y) + \frac{1}{y - \xi_n(y)} \left(\lim_{\delta \to 0} \frac{c^n(\xi_n(y+\delta), y+\delta) - c^n(\xi_n(y), y)}{\delta} \right).
\end{aligned}$$

The main technical difficulty comes from the possibility that $\xi_n(y) = \xi_k(y)$ for some $k < n$. We present the arguments for this case and leave the other (much easier) case, to the reader.

By assumption the last limit exists and hence we can compute it using some "convenient" sequence $\delta_m \downarrow 0$ where δ_m is such that $\jmath_n(y + \delta_m) = l$ for all $m \in \mathbb{N}$. Note that by continuity of $\xi_1, \ldots, \xi_n$ at y we have that either $l = \jmath_n(y)$ or l is such that $\xi_l(y) = \xi_n(y)$. This will enable us to apply (A.1.15). Recall (2.1.46). For δ_m small enough such that $\imath_n(\xi_n(y); y + \delta_m) = \jmath_n(y)$ we obtain

$$\begin{aligned}
& c^n(\xi_n(y+\delta_m), y+\delta_m) - c^n(\xi_n(y), y+\delta_m) \\
= \;& c_n(\xi_n(y+\delta_m)) - c_l(\xi_n(y+\delta_m)) + (y+\delta_m - \xi_n(y+\delta_m))K_l(y+\delta_m) \\
& -c_n(\xi_n(y)) + c_{\jmath_n(y)}(\xi_n(y)) - (y+\delta_m - \xi_n(y))K_{\jmath_n(y)}(y+\delta_m) \\
\overset{(A.1.15)}{=} & c_n(\xi_n(y+\delta_m)) - c_l(\xi_n(y+\delta_m)) + (y+\delta_m - \xi_n(y+\delta_m))K_l(y+\delta_m) \\
& -c_n(\xi_n(y)) + c_l(\xi_n(y)) - (y - \xi_n(y))(K_l(y) - K_{\jmath_n(y)}(y)) \\
& -(y+\delta_m - \xi_n(y))K_{\jmath_n(y)}(y+\delta_m).
\end{aligned}$$

From this we obtain for almost all $y \geq 0$ by using the induction hypothesis

$$\begin{aligned}
& \lim_{m\to\infty} \frac{c^n(\xi_n(y+\delta_m), y+\delta_m) - c^n(\xi_n(y), y+\delta_m)}{\delta_m} \\
= \;& \xi_n'(y+)\Big[c_n'(\xi_n(y)+) - c_l'(\xi_n(y)+) - K_l(y)\Big] \\
& +K_l(y) + (y-\xi_n(y))K_l'(y) - K_{\jmath_n(y)}(y) - (y-\xi_n(y))K_{\jmath_n(y)}'(y) \\
\underset{(A.1.2)}{\overset{(A.1.6)}{=}} & -\xi_n'(y+)K_n(y), \qquad \text{(A.1.16)}
\end{aligned}$$

in the case when $c^n(\cdot, y)$ is differentiable at $\xi_n(y)$. Together with (A.1.5) this yields

$$\begin{aligned}
K_n'(y) &= \frac{\xi_n'(y) - 1}{y - \xi_n(y)}K_n(y) + \frac{1}{y-\xi_n(y)}\left(-K_n(y)\xi_n'(y) + \frac{\partial c^n}{\partial y}(\xi_n(y), y)\right) \\
&= -\frac{K_n(y)}{y-\xi_n(y)} + \frac{1}{y-\xi_n(y)}\Big(K_{\jmath_n(y)}(y) + (y-\xi_n(y))K_{\jmath_n(y)}'(y)\Big). \qquad \text{(A.1.17)}
\end{aligned}$$

In order to finish the proof we just have to establish that (A.1.17) also holds in the case when $c^n(\cdot, y)$ is not differentiable at $\xi_n(y)$.

To this end, we first argue that (A.1.16), and hence (A.1.17), remains true in the case when $c^n(\cdot, y)$ is not differentiable at $\xi_n(y)$, but when the slope of the supporting tangent to $c^n(\cdot, y)$ at $\xi_n(y)$ which passes the x-axis at y equals the right-derivative of $c^n(\cdot, y)$ at $\xi_n(y)$. In that case, denoting $k = \imath_n(\xi_n(y)+; y)$,

$$c_n'(\xi_n(y)+) - c_k'(\xi_n(y)+) - K_k(y) = -K_n(y), \tag{A.1.18}$$

holds by optimality of $\xi_n(y)$. Recall the sequence (δ_m). We do not necessarily have $k = \jmath_n(y + \delta_m) = l$. Nevertheless, we argue that

$$c_n'(\xi_n(y)+) - c_l'(\xi_n(y)+) - K_l(y) = -K_n(y) \tag{A.1.19}$$

holds. We can safely assume that $\xi_n'(y+) > 0$ (in the other case the conclusion of (A.1.17) remains true). Also it is enough to consider the case when $k > \jmath_n(y + \delta)$ for all $\delta > 0$ sufficiently small (otherwise we may consider an alternative sequence (δ_m) where $k = \jmath_n(y + \delta_m)$ for all m, i.e. $l = k$ and (A.1.19) would coincide with (A.1.18)). Consequently, $\xi_k'(y+) \geq \xi_n'(y+) > 0$. Then, since by induction hypothesis (A.1.6) or (A.1.8) holds true for k, we must have, for almost all y that

$$c_k'(\xi_n(y)+) - c_j'(\xi_n(y)+) - K_j(y) = -K_k(y) \tag{A.1.20}$$

for $j = \jmath_k(y) = \jmath_n(y)$ and j such that $n > k > j > \jmath_n(y)$ and $\xi_n(y) = \xi_k(y) = \xi_j(y)$. Then, combining these results we conclude by (A.1.18) and (A.1.20) that indeed (A.1.19) holds.

Now we consider the case of non-smoothness of $c^n(\cdot, y)$ at $\xi_n(y)$ and where y is such that the slope of the supporting tangent to $c^n(\cdot, y)$ at $\xi_n(y)$ which crosses the x-axis at y does not equal the right-derivative of $c^n(\cdot, y)$ at $\xi_n(y)$. We show that in this case we have for sufficiently small $\delta > 0$,

$$\xi_n(y) = \xi_n(y + \delta) \qquad \text{and hence} \qquad \xi_n'(y+) = 0, \tag{A.1.21}$$

which implies that (A.1.17) holds as well.

To achieve this we place a suitable tangent to $c^n(\cdot, y)$ at $\xi_n(y)$. Since, by assumption, $c^n(\cdot, y)$ has a kink at $\xi_n(y)$ we have some flexibility to do that. Recalling the tangent interpretation of (2.1.12) we know by choice of $\xi_n(y)$ that we can place a supporting tangent to $c^n(\cdot, y)$ at $\xi_n(y)$ which passes through the x-axis at y. Alternatively, by choice of y, we can place a tangent to $c^n(\cdot, y)$ at $\xi_n(y)$ which crosses the x-axis at some $y + \delta > y$. This implies that

$$\arg\min_{\zeta \le y} \frac{c^n(\zeta, y)}{y + \delta - \zeta} = \xi_n(y). \tag{A.1.22}$$

Assume first that $\xi_n(y) < y$. Denote $k = \imath_n(\xi_n(y)+; y)$. For simplicity of the argument let us also assume that $\xi_n(y) = \xi_k(y) \neq \xi_i(y)$ for all $i \neq k, n$. Also denote $j = \jmath_n(y) = \imath_n(\xi_n(y); y)$.

Now we will use (A.1.22) to deduce (A.1.21). By continuity and monotonicity of ξ_n we have for $\delta > 0$ small enough that $\xi_n(y) \le \xi_n(y + \delta) < \xi_n(y) + \epsilon < y$ for some $\epsilon = \epsilon(\delta) > 0$. By taking δ small enough we can also assume that $k = \max_{\zeta \le \xi_n(y)+\epsilon} \imath_n(\zeta; y + \delta)$. Then we have

$$\inf_{\zeta \le y+\delta} \frac{c^n(\zeta, y+\delta)}{y + \delta - \zeta} \ge \inf_{\xi_n(y) \le \zeta < \xi_n(y)+\epsilon} \frac{c^n(\zeta, y)}{y + \delta - \zeta} + \inf_{\xi_n(y) \le \zeta < \xi_n(y)+\epsilon} \frac{c^n(\zeta, y+\delta) - c^n(\zeta, y)}{y + \delta - \zeta}. \tag{A.1.23}$$

As for the first infimum in (A.1.23) we know from (A.1.22) that it is attained at $\zeta = \xi_n(y)$. Now we will show that the second infimum in (A.1.23) is also attained at $\zeta = \xi_n(y)$. To this end consider the following estimate for $\zeta \in (\xi_n(y), \xi_n(y) + \epsilon)$,

$$\begin{aligned}
&c^n(\zeta, y+\delta) - c^n(\zeta, y) \\
&= - c_{\imath_n(\zeta; y+\delta)}(\zeta) + c_{\imath_n(\zeta; y)}(\zeta) - (y - \zeta) K_{\imath_n(\zeta; y)}(y) + (y + \delta - \zeta) K_{\imath_n(\zeta; y+\delta)}(y + \delta) \\
&\overset{(2.1.12)}{\ge} - (y - \zeta) K_{\imath_{\imath_n(\zeta;y)}(\zeta; y)}(y) + (y + \delta - \zeta) K_{\imath_n(\zeta; y+\delta)}(y + \delta) \\
&\ge - (y - \zeta) K_j(y) + (y + \delta - \zeta) K_j(y + \delta) =: l(\zeta, y; \delta).
\end{aligned} \tag{A.1.24}$$

Since $l(\cdot, y; \delta)$ is non-decreasing and non-negative, we deduce that

$$\operatorname*{arg\,min}_{\xi_n(y) \le \zeta < \xi_n(y)+\epsilon} \frac{l(\zeta, y; \delta)}{y + \delta - \zeta} = \xi_n(y).$$

Finally, because at $\zeta = \xi_n(y)$ there is equality in (A.1.24) we can conclude

$$\operatorname*{arg\,min}_{\zeta \le y+\delta} \frac{c^n(\zeta, y+\delta)}{y+\delta-\zeta} = \xi_n(y)$$

as required.

In the case when $\xi_n(y) = y$ we obtain by (2.1.46) for $\delta > 0$ sufficiently small that $\imath_n(y; y+\delta) = \imath_n(y; y)$ and hence, using $c_n(y) = c_{\imath_n(y;y)}(y)$,

$$\begin{aligned}\operatorname*{arg\,min}_{\zeta \le y+\delta} \frac{c^n(\zeta, y+\delta)}{y+\delta-\zeta} &= \operatorname*{arg\,min}_{y \le \zeta \le y+\delta} \left(\underbrace{\frac{c_n(\zeta) - c_{\imath_n(\zeta; y+\delta)}(\zeta)}{y+\delta-\zeta}}_{\ge 0} + \underbrace{K_{\imath_n(\zeta; y+\delta)}(y+\delta)}_{\ge K_{\imath_n(y;y+\delta)}(y+\delta)} \right) \\ &\overset{(2.1.13)}{=} \xi_n(y) = y.\end{aligned}$$

The proof is complete. □

A.2 When Does Assumption ⊛(ii) Fail?

Although Assumption ⊛(ii) was convenient in the construction and proof of our main result, it has the disadvantage that it is not very explicit.

Let (X_1, X_2) be a martingale with marginals ν_1 and ν_2. Hobson and Klimmek [53] consider optimal lower bounds on $\mathbb{E}\left[|X_1 - X_2|\right]$ of the form $\int g \mathrm{d}\nu_1 + \int h \mathrm{d}\nu_2$. For their construction they require some condition on the difference of the distribution functions of ν_1 and ν_2. In our construction we have the additional state variable of the continuously sampled maximum and our construction depends on this variable. Hence, the fact that Assumption ⊛(ii) features quantities which relate to this variable should not surprise.

The following result sheds more light on the question when Assumption ⊛(ii) fails and can be seen as an extension of Brown et al. [18, Section 3.5].

Proposition A.2.1 (Sufficient Condition). *Let Assumption ⊛(i) holds and assume that $c_1, \dots, c_n$ are twice continuously differentiable.*

Set $\mu_0 \equiv 0$ and recall the definition of l_{μ_n} and r_{μ_n} in (2.1.8). Suppose that for all $n \in \mathbb{N}$ and all $\alpha = \zeta_0 \le \zeta_1 \le \dots \le \zeta_n = \beta$, where either $\alpha > l_{\mu_n}$ or $\beta < r_{\mu_n}$, the measures

$$\lambda_1 := \mu_n\big|_{[\alpha,\beta]} \qquad \textit{and} \qquad \lambda_2 := \sum_{i=0}^{n-1} \mu_i\big|_{[\zeta_i,\zeta_{i+1}]} \tag{A.2.1}$$

do not have the same mass or mean or are not in convex order.

Then $\mu_1, \dots, \mu_n$ satisfy Assumption ⊛(ii).

Proof. The case $n = 2$ follows from Brown et al. [18, Section 3.5]. For notational convenience we will only spell out the proof in the case $n = 3$. It should be clear that similar arguments apply for general $n \in \mathbb{N}$.

It follows immediately that

$$c_i < c_{i+1} \qquad \text{for } i = 1, \dots, n-1.$$

We will proceed by proving that if $\alpha < \beta$ are minimizers in (2.1.16) then

$$\lambda_1\big|_{[\alpha,\beta]} \qquad \text{and} \qquad \lambda_2 = \mu_1\big|_{[\alpha,\xi_2]} + \mu_2\big|_{[\xi_2,\beta]} \tag{A.2.2}$$

from (A.2.1) have the same mass and mean and are in convex order. This situation however is ruled out by our assumption.

We restrict to the following case,

$$\imath_3(\beta; y) = 2 \qquad \text{and} \qquad \imath_3(\alpha; y) = 1. \tag{A.2.3}$$

The other cases follow similarly.

Note that under our assumptions (A.1.6) holds for $\xi_3(y) = \alpha$ and $\xi_3(y) = \beta$.

Masses. The mass of λ_1 is $c_3'(\beta) - c_3'(\alpha)$. The mass of λ_2 is

$$c_2'(\beta) - c_2'(\xi_2) + c_1'(\xi_2) - c_1'(\alpha) \overset{\text{(A.1.6)}}{=} c_2'(\beta) + K_2(y) - K_1(y) - c_1'(\alpha)$$
$$\overset{\text{(A.1.6)}}{=} c_3'(\beta) - c_3'(\alpha)$$

where at the last equality we applied (A.1.6) for α and β.

Means. The mean of λ_1 is

$$\int_\alpha^\beta x\mu_3(\mathrm{d}x) = c_3(\alpha) - c_3(\beta) + (\beta - \alpha)c_3'(\beta) + \underbrace{\alpha\mu_3([\alpha, \beta])}_{=\xi_2\mu_3([\alpha,\beta]) - (\xi_2 - \alpha)(c_3'(\beta) - c_3'(\alpha))}.$$

The mean of λ_2 is

$$c_2(\xi_2) - c_2(\beta) + (\beta - \xi_2)c_2'(\beta) + \xi_2\mu_2([\xi_2, \beta])$$
$$+ c_1(\alpha) - c_1(\xi_2) + (\xi_2 - \alpha)\underbrace{c_1'(\xi_2)}_{=\mu_1([\alpha,\xi_2]) + c_1'(\alpha)} + \alpha\mu_1([\alpha, \xi_2]).$$

Observe that by optimality of α and β and smoothness we have

$$\frac{\partial c^3}{\partial x}(\alpha, y) = \frac{c^3(\beta, y) - c^3(\alpha, y)}{\beta - \alpha} = \frac{\partial c^3}{\partial x}(\beta, y).$$

Now we compute the difference of the means as

$$\begin{aligned}
& c^3(\alpha, y) - c^3(\beta, y) + (\xi_2 - \alpha)\frac{\partial c^3}{\partial x}(\alpha, y) + (\beta - \xi_2)\frac{\partial c^3}{\partial x}(\beta, y) \\
& -(y-\alpha)K_1(y) + (y-\beta)K_2(y) \\
& -(\xi_2 - \alpha)c_3'(\alpha) + (\xi_2 - \alpha)K_1(y) - (\alpha - \xi_2)c_3'(\beta) + (\beta - \xi_2)K_2(y) \\
& -c_2(\xi_2) + c_1(\xi_2) + \alpha\mu_3([\alpha, \beta]) - \xi_2\mu_2([\xi_2, \beta]) - \xi_2\mu_1([\alpha, \xi_2]) \\
= \;& (y - \xi_2)K_2(y) - (y - \xi_2)K_1(y) - c_2(\xi_2) + c_1(\xi_2) \\
& -(\xi_2 - \alpha)c_3'(\alpha) - (\alpha - \xi_2)c_3'(\beta) - (\xi_2 - \alpha)(c_3'(\beta) - c_3'(\alpha)) \\
& +\xi_2\big(\mu_3([\alpha, \beta] - \mu_2([\xi_2, \beta]) - \mu_1([\alpha, \xi_2]))\big) \\
= \;& 0
\end{aligned}$$

by definition of $K_2(y)$ and the same mass of λ_1 and λ_2.

Convex Order. We have to show

$$\int (x-z)^+ \lambda_1(\mathrm{d}x) \leq \int (x-z)^+ \lambda_2(\mathrm{d}x) \qquad \text{for all } z \in [\alpha, \beta].$$

We compute for $z \in [\alpha, \beta]$

$$\int (x-z)^+ \lambda_1(\mathrm{d}x) = c_3(z) - c_3(\beta) - (\beta - z)\underbrace{\mu_3([\beta, \infty])}_{=-c_3'(\beta)}$$

and for $z \in [\alpha, \xi_2]$

$$\begin{aligned}
\int (x-z)^+ \lambda_2(\mathrm{d}x) = \;& c_1(z) - c_1(\xi_2) - (\xi_2 - z)\underbrace{\mu_1([\xi_2, \infty])}_{=-c_1'(\xi_2)} \\
& + c_2(\xi_2) - c_2(\beta) - (\beta - \xi_2)\underbrace{\mu_2([\beta, \infty])}_{=-c_2'(\beta)} + (\xi_2 - z)\underbrace{\mu_2([\xi_2, \beta])}_{=c_2'(\beta) - c_2'(\xi_2)}.
\end{aligned}$$

For $z \in [\alpha, \xi_2]$, the difference of these two expression is

$$\begin{aligned}
&\overbrace{c^3(z,y) - c^3(\beta,y) + (\beta - z)(c_3'(\beta) - c_2'(\beta) - K_2(y))}^{\geq 0} \\
&+(y-\beta)K_2(y) - (y-\xi_2)K_1(y) + (\beta - z)K_2(y) \\
&-c_2(\xi_2) + c_1(\xi_2) + (\xi_2 - z)\underbrace{(c_2'(\xi_2) - c_1'(\xi_2))}_{=-K_2(y)+K_1(y)} \geq 0
\end{aligned}$$

by definition of $K_2(y)$.

The case $z \in [\xi_2, \beta]$ follows similarly. □

Appendix B

Appendix: Part II

B.1 The Fan [39] and Kneser [61] Min-Max Theorem

The following definition extends the concept of concavity and convexity to spaces without linear structure.

Definition B.1.1 (Concave-Convex-Like)**.** *A function $f : \mathscr{X} \times \mathscr{Y} \to \mathbb{R}$ is said to be concave-convex-like on $\mathscr{X} \times \mathscr{Y}$ if, for $0 \leq a \leq 1$ the following two properties hold:*

(i) For $x_1, x_2 \in \mathscr{X}$ there exists $x_3 \in \mathscr{X}$ such that

$$f(x_3, y) \geq af(x_1, y) + (1 - a)f(x_2, y)$$

for all $y \in \mathscr{Y}$.

(ii) For $y_1, y_2 \in \mathscr{Y}$ there exists $y_3 \in \mathscr{Y}$ such that

$$f(x, y_3) \leq af(x, y_1) + (1 - a)f(x, y_2)$$

for all $x \in \mathscr{X}$.

We refer to Sion [89] for a collection of Min-Max theorems.

Theorem B.1.2 (Fan [39] and Kneser [61] Min-Max Theorem, cf. also Theorem 4.2 of Sion [89]). *Suppose that* $\mathscr{X}$ *and* $\mathscr{Y}$ *are non-empty sets with* f *concave-convex like on* $\mathscr{X} \times \mathscr{Y}$. *Suppose that* $\mathscr{X}$ *is compact and* $f(\cdot, y)$ *is lower semi-continuous on* $\mathscr{X}$ *for each* $y \in \mathscr{Y}$. *Then*

$$\inf_{y\in\mathscr{Y}} \sup_{x\in\mathscr{X}} f(x,y) = \sup_{x\in\mathscr{X}} \inf_{y\in\mathscr{Y}} f(x,y). \tag{B.1.1}$$

Bibliography

[1] B Acciaio, M Beiglböck, F Penkner, and W Schachermayer. A model free version of the Fundamental Theorem of Asset Pricing and the Superreplication Theorem. *Mathematical Finance*, December 2013.

[2] B Acciaio, M Beiglböck, F Penkner, W Schachermayer, and J Temme. A trajectorial interpretation of Doob's martingale inequalities. *The Annals of Applied Probability*, 23(4):1494–1505, 2013.

[3] Torben G Andersen, Tim Bollerslev, Francis X Diebold, and Paul Labys. The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96:42–55, 2000.

[4] J Azéma and M Yor. Une solution simple au problème de Skorokhod. *Séminaire de Probabilités XIII: 06, 90-115, LNM 721.*, 1979.

[5] J Azéma, R F Gundy, and M Yor. Sur l'intégrabilité uniforme des martingales continues. In *Seminar on Probability, XIV (Paris, 1978/1979) (French)*, pages 53–61. Springer, Berlin, 1980.

[6] Ole E. Barndorff-Nielsen and Neil Shephard. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2):253–280, 2002.

[7] Mathias Beiglböck and Martin Huesmann. Optimal Transport and Skorokhod Embedding. *arXiv.org*, (1307.3656v1), July 2013.

[8] Mathias Beiglböck and Marcel Nutz. Martingale Inequalities and Deterministic Counterparts. *arXiv.org*, (1401.4698v1), January 2014.

[9] Mathias Beiglböck and Pietro Siorpaes. Pathwise versions of the Burkholder-Davis-Gundy inequality. *arXiv.org*, (1305.6188v1), May 2013.

[10] Mathias Beiglböck, Pierre Henry-Labordère, and Friedrich Penkner. Model-independent bounds for option prices—a mass transport approach. *Finance and Stochastics*, 17(3):477–501, 2013.

[11] Alain Bensoussan and Jacques-Louis Lions. *Applications of variational inequalities in stochastic control*, volume 12 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam-New York, 1982.

[12] Avi Bick and Walter Willinger. Dynamic spanning without probabilities. *Stochastic Processes and their Applications*, 50(2):349–374, 1994.

[13] F Black and M Scholes. The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654, 1973.

[14] David Blackwell and Lester E Dubins. A converse to the dominated convergence theorem. *Illinois Journal of Mathematics*, 7:508–514, 1963.

[15] Douglas T Breeden and Robert H Litzenberger. Prices of State-Contingent Claims Implicit in Option Prices. *The Journal of Business*, 51(4):621–651, October 1978.

[16] H Brown, D G Hobson, and L C G Rogers. Robust hedging of barrier options. *Mathematical Finance. An International Journal of Mathematics, Statistics and Financial Economics*, 11(3):285–314, 2001.

[17] H Brown, D G Hobson, and L C G Rogers. Robust Hedging of Barrier Options. *Mathematical Finance*, 11(3):285–314, 2001.

[18] Haydyn Brown, David Hobson, and L C G Rogers. The maximum maximum of a martingale constrained by an intermediate law. *Probability Theory and Related Fields*, 119(4):558–578, 2001.

[19] P Carr. Local variance gamma option pricing model. *presentation Bloomberg/Courant Institute*, 2009.

[20] L Carraro, N El Karoui, and J Obłój. On Azéma-Yor processes, their optimal properties and the Bachelier-Drawdown equation. *Annals of Probability*, 40 (1):372–400, jan 2012.

[21] Rama Cont. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance*, 16(3):519–547, 2006.

[22] A M G Cox and Jan Obłój. Robust hedging of double touch barrier options. *SIAM Journal on Financial Mathematics*, 2(1):141–182, 2011.

[23] A M G Cox and G Peskir. Embedding Laws in Diffusions by Functions of Time. *arXiv.org*, (1201.5321v3), January 2012.

[24] A M G Cox and J Wang. Root's barrier: Construction, optimality and applications to variance options. *Annals of Applied Probability*, 23(3):859–894, March 2013.

[25] A M G Cox, David Hobson, and Jan Obłój. Pathwise inequalities for local time: Applications to Skorokhod embeddings and optimal stopping. *The Annals of Applied Probability*, 18(5):1870–1896, October 2008.

[26] Alexander Cox, Zhaoxu Hou, and Jan Obłój. Robust pricing and hedging under trading restrictions and the emergence of local martingale models. *arXiv.org*, (arXiv:1406.0551), 2014.

[27] Alexander M G Cox and David G Hobson. Local martingales, bubbles and option prices. *Finance and Stochastics*, 9(4):477–492, 2005.

[28] Alexander M G Cox and Jan Obłój. Robust pricing and hedging of double no-touch options. *Finance and Stochastics*, 15(3):573–605, 2011.

[29] Alexander M G Cox and Jiajie Wang. Optimal robust bounds for variance options. *arXiv.org*, (1308.4363v1), August 2013.

[30] Alexander M G Cox, David Hobson, and Jan Obłój. Time-homogeneous diffusions with a given marginal at a random time. *ESAIM: Probability and Statistics*, 15:S11–S24, October 2010.

[31] K Dambis. On the decomposition of continuous submartingales. *Theory of Probability and its Applications*, 10(3):401–410, 1965.

[32] M Davis, J Obłój, and V Raval. Arbitrage bounds for prices of weighted variance swaps. *Mathematical Finance*, 2013.

[33] Mark H A Davis and David G Hobson. The range of traded option prices. *Mathematical Finance. An International Journal of Mathematics, Statistics and Financial Economics*, 17(1):1–14, 2007.

[34] Laurent Denis and Claude Martini. A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. *Ann. Appl. Probab.*, 16(2):827–852, 2006.

[35] Yan Dolinsky and H Mete Soner. Martingale optimal transport and robust hedging in continuous time. *Probability Theory and Related Fields*, pages 1–37, 2013.

[36] Yan Dolinsky and H Mete Soner. Martingale optimal transport in the Skorokhod space. *arXiv.org*, (1404.1516v1), 2014.

[37] Lester E Dubins and Gideon Schwarz. On Continuous Martingales. *Proceedings of the National Academy of Sciences of the United States of America*, 53 (5):913–916, May 1965.

[38] B Dupire. Arbitrage bounds for volatility derivatives as free boundary problem. *Presentation at PDE and Mathematical Finance*, 2005.

[39] K Fan. Minimax theorems. *National Academy of Sciences, Washington, DC, Proceedings USA*, (39):42–47, 1953.

[40] H Föllmer. Calcul d'Itô sans probabilités. In *Séminaire de Probabilités XV (Univ. Strasbourg, 1979/1980)*, volume 850 of *Lecture Notes in Math.*, pages 143–150. Springer, Berlin, 1981.

[41] Martin Forde. A diffusion-type process with a given joint law for the terminal level and supremum at an independent exponential time. *Stochastic Processes and their Applications*, 121(12):2802–2817, 2011.

[42] Martin Forde, Andrey Pogudin, and Hongzhong Zhang. Hitting times, occupation times, trivariate laws and the forward Kolmogorov equation for a one-dimensional diffusion with memory. *Advances in Applied Probability*, 45 (3):860–875, 2013.

[43] Avner Friedman. *Stochastic differential equations and applications. Vol. 1.* Academic Press, London, 1975.

[44] 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(1):312–336, 2014.

[45] David Gilat and Isaac Meilijson. A simple proof of a theorem of Blackwell & Dubins on the maximum of a uniformly integrable martingale. In *Séminaire de Probabilités XXII*, pages 214–216. Springer Berlin Heidelberg, 1988.

[46] L P Hansen and T J Sargent. *Robustness.* Princeton University Press, 2008.

[47] L P Hansen and T J Sargent. Fragile beliefs and the price of uncertainty. *Quantitative Economics*, 1(1):129–162, 2010.

[48] P Henry-Labordère. Automated Option Pricing: Numerical Methods. *Journal of Theoretical and Applied Finance*, 2013.

[49] Pierre Henry-Labordère, Jan Obłój, Peter Spoida, and Nizar Touzi. Maximum Maximum of Martingales given Marginals. *arXiv.org*, (1203.6877v2), March 2012.

[50] D G Hobson. Robust hedging of the lookback option. *Finance Stoch.*, 2(4): 329–347, 1998.

[51] David Hobson. The maximum maximum of a martingale. In *Séminaire de Probabilités XXXII*, volume 1686 of *Lecture Notes in Mathematics*, pages 250–263. Springer Berlin Heidelberg, 1998.

[52] David Hobson. The Skorokhod embedding problem and model-independent bounds for option prices. In *Paris-Princeton Lectures on Mathematical Finance 2010*, pages 267–318. Springer, Berlin, 2011.

[53] David Hobson and Martin Klimmek. Robust price bounds for the forward starting straddle. *arXiv.org*, (1304.2141v1), April 2013.

[54] David Hobson and Martin Klimmek. Maximizing functionals of the maximum in the Skorokhod embedding problem and an application to variance swaps. *The Annals of Applied Probability*, 23(5):2020–2052, 10 2013.

[55] Robert A Jarrow, Philip Protter, and Kazuhiro Shimbo. Asset price bubbles in an incomplete market. *Mathematical Finance*, 20:145–185, 2007.

[56] Yuri Kabanov and Mher Safarian. *Markets with Transaction Costs*. Springer Finance. Springer, 2010.

[57] Robert P Kertz and Uwe Rösler. Martingales with given maxima and terminal distributions. *Israel Journal of Mathematics*, 69(2):173–192, June 1990.

[58] Robert P Kertz and Uwe Rösler. Stochastic and convex orders and lattices of probability measures, with a martingale interpretation. *Israel Journal of Mathematics*, 77(1-2):129–164, February 1992.

[59] Robert P Kertz and Uwe Rösler. Hyperbolic-concave functions and Hardy-Littlewood maximal functions. *Institute of Mathematical Statistics, Lecture Notes–Monograph Series*, 22:196–210, 1992.

[60] John Maynard Keynes. *Treatise on Probability.* London: Macmillan & Co, 1921.

[61] Hellmuth Kneser. Sur un théorème fondamental de la théorie des jeux. *C. R. Acad. Sci. Paris*, 234:2418–2420, 1952.

[62] F H Knight. *Risk, Uncertainty and Profit.* Boston: Houghton Mifflin, 1921.

[63] T J Lyons. Uncertain volatility and the risk-free synthesis of derivatives. *Applied Mathematical Finance*, 1995.

[64] Dilip B Madan and Marc Yor. Making Markov Martingales Meet Marginals: With Explicit Constructions. *Bernoulli*, 8(4):509–536, August 2002.

[65] P A Meyer. *Probability and potentials.* Blaisdell book in pure and applied mathematics. Blaisdell Pub. Co., 1966.

[66] Itrel Monroe. On Embedding Right Continuous Martingales in Brownian Motion. *Ann. Math. Statist.*, 43(4):1293–1311, August 1972.

[67] P A Mykland. Combining statistical intervals and market prices: The worst case state price distribution. *Dept of Statistics, The Univ. of Chicago, (TR 553)*, 2005.

[68] Per Aslak Mykland. Conservative Delta Hedging. *The Annals of Applied Probability*, 10(2):664–683, May 2000.

[69] Per Aslak Mykland. Financial Options and Statistical Prediction Intervals. *The Annals of Statistics*, 31(5):1413–1438, October 2003.

[70] Per Aslak Mykland. The interpolation of options. *Finance and Stochastics*, 7(4):417–432, 2003.

[71] Marcel Nutz and Bruno Bouchard. Arbitrage and Duality in Nondominated Discrete-Time Models. *arXiv.org*, (arXiv:1305.6008v2), 2013.

[72] Harald Oberhauser and Goncalo dos Reis. Root's barrier, viscosity solutions of obstacle problems and reflected FBSDEs. *arXiv.org*, (1301.3798v1), January 2013.

[73] Jan Obłój. The Skorokhod embedding problem and its offspring. *Probability Surveys*, 1:321–390 (electronic), 2004.

[74] Jan Obłój. On some aspects of Skorokhod Embedding Problem and its applications in Mathematical Finance. *lecture notes*, 2013. URL `http://people.maths.ox.ac.uk/obloj/LectureNotes_v3.pdf`.

[75] Jan Obłój and Peter Spoida. An Iterated Azéma-Yor Type Embedding for Finitely Many Marginals. *arXiv.org*, (1304.0368v2), April 2013.

[76] Jan Obłój and Peter Spoida. Robust Pricing and Hedging with Beliefs about Realized Variance. *In preparation*, 2014.

[77] Jan Obłój and Marc Yor. On local martingale and its supremum: harmonic functions and beyond. In *From Stochastic Calculus to Mathematical Finance*, pages 517–534. Springer-Verlag, 2006.

[78] Jan Obłój, Peter Spoida, and Nizar Touzi. Martingale Inequalities via Pathwise Arguments. *arXiv.org*, (1409.6255), 2014.

[79] Shige Peng. *G-Expectation, G-Brownian Motion and Related Stochastic Calculus of Itô Type*, volume 2 of *Abel Symposia*. Springer Berlin Heidelberg, 2007.

[80] Goran Peskir. Optimal Stopping of the Maximum Process: The Maximality Principle. *The Annals of Probability*, 26(4):1614–1640, October 1998.

[81] Goran Peskir and Albert Shiryaev. *Optimal Stopping and Free-Boundary Problems*. Springer, November 2006.

[82] Vladimir Piterbarg. Derivatives pricing - Spread options, Farkas's lemma and linear programming. *Risk*, 2011.

[83] L C G Rogers. A guided tour through excursions. *The Bulletin of the London Mathematical Society*, 21(4):305–341, 1989.

[84] L C G Rogers. The joint law of the maximum and terminal value of a martingale. *Probability Theory and Related Fields*, 95(4):451–466, December 1993.

[85] D H Root. The Existence of Certain Stopping Times on Brownian Motion. *Ann. Math. Statist.*, 40(2):715–718, April 1969.

[86] H Rost. Skorokhod stopping times of minimal variance. In *Séminaire de Probabilités, X (Première partie, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975)*, pages 194–208. Lecture Notes in Math., Vol. 511. Springer, Berlin, 1976.

[87] Hermann Rost. The stopping distributions of a Markov Process. *Inventiones Mathematicae*, 14:1–16, 1971.

[88] Moshe Shaked and J George Shanthikumar. *Stochastic orders and their applications.* Academic Pr, 1994.

[89] Maurice Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8:171–176, 1958.

[90] Anatolij Vladimirovič Skorokhod. Studies in the Theory of Random Processes. *Translated from the Russian by Scripta Technica, Inc. Addison-Wesley Publishing Co., Inc., Reading, Mass.*, 1965.

[91] M Soner, N Touzi, and J Zhang. Martingale representation theorem for the G-expectation. *Stochastic Processes and Their Applications*, 121:265–287, 2011.

[92] V Strassen. The Existence of Probability Measures with Given Marginals. *Ann. Math. Statist.*, 36(2):423–439, April 1965.

[93] Xiaolu Tan and Nizar Touzi. Optimal transportation under controlled stochastic dynamics. *The Annals of Probability*, 41(5):3201–3240, 2013.

[94] Pierre Vallois. On the joint distribution of the supremum and terminal value of a uniformly integrable martingale. In *Stochastic processes and optimal control (Friedrichroda, 1992)*, pages 183–199. Gordon and Breach, Montreux, 1993.

[95] Cedric Villani. *Optimal transport - Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften.* 2009.

[96] Vladimir Vovk. Ito calculus without probability in idealized financial markets. *arXiv.org*, (1108.0799v1), August 2011.