

# Robust pricing–hedging dualities in continuous time

Zhaoxu Hou<sup>1</sup> · Jan Oblój<sup>1</sup>

Received: 1 September 2015 / Accepted: 16 January 2018 / Published online: 30 May 2018  
© The Author(s) 2018

**Abstract** We pursue a robust approach to pricing and hedging in mathematical finance. We consider a continuous-time setting in which some underlying assets and options, with continuous price paths, are available for dynamic trading and a further set of European options, possibly with varying maturities, is available for static trading. Motivated by the notion of prediction set in Mykland (Ann. Stat. 31:1413–1438, 2003), we include in our setup modelling beliefs by allowing to specify a set of paths to be considered, e.g. superreplication of a contingent claim is required only for paths falling in the given set. Our framework thus interpolates between model-independent and model-specific settings and allows us to quantify the impact of making assumptions or gaining information. We obtain a general pricing–hedging duality result: the infimum over superhedging prices of an exotic option with payoff  $G$  is equal to the supremum of expectations of  $G$  under calibrated martingale measures. Our results include in particular the martingale optimal transport duality of Dolinsky and Soner (Probab. Theory Relat. Fields 160:391–427, 2014) and extend it to multiple dimensions, multiple maturities and beliefs which are invariant under time-changes. In a general setting with arbitrary beliefs and for a uniformly continuous  $G$ , the asserted duality holds between limiting values of perturbed problems.

**Keywords** Robust pricing and hedging · Pricing–hedging duality · Martingale optimal transport · Path space restrictions · Pathwise modelling

**Mathematics Subject Classification (2010)** 91G20 · 91B24 · 60G44

**JEL Classification** C61 · G13

---

✉ J. Oblój  
jan.obloj@maths.ox.ac.uk

<sup>1</sup> Mathematical Institute, University of Oxford, Woodstock Road, Oxford OX2 6GG, UK

## 1 Introduction

**Two approaches to pricing and hedging** The question of pricing and hedging of a contingent claim lies at the heart of mathematical finance. Following Merton's seminal contribution [38], we may distinguish two ways of approaching it. First, one may want to make statements “based on assumptions sufficiently weak to gain universal support”,<sup>1</sup> e.g. market efficiency combined with some broad mathematical idealisation of the market setting. We refer to this perspective as *the model-independent approach*. While very appealing at first, it has been traditionally criticised for producing outputs which are too imprecise to be of practical relevance. This is contrasted with the second, *model-specific approach* which focuses on obtaining explicit statements leading to unique prices and hedging strategies. “To do so, more structure must be added to the problem through additional assumptions at the expense of losing some agreement.”<sup>1</sup> Typically this is done by fixing a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  with risky assets represented by some adapted process  $(S_t)$ .

The model-specific approach, originating from the seminal works of Samuelson [47] and Black and Scholes [8], has revolutionised the financial industry and become the dominating paradigm for researchers in quantitative finance. Accordingly, we refer to it also as the *classical approach*. The original model of Black and Scholes has been extended and generalised, e.g. adding stochastic volatility and/or stochastic interest rates, trying to account for market complexity observed in practice. Such generalisations often lead to market incompleteness and a lack of unique rational warrant prices. Nevertheless, no-arbitrage pricing and hedging was fully characterised in a body of works on the fundamental theorem of asset pricing (FTAP) culminating in Schachermayer [20, 21]. The feasible prices for a contingent claim correspond to expectations of the (discounted) payoff under equivalent martingale measures (EMM) and form an interval. The bounds of the interval are also given by the super- and sub-hedging prices. Put differently, the supremum of expectations of the payoff under EMMs is equal to the infimum of prices of superhedging strategies. We refer to this fundamental result as the *pricing–hedging duality*.

**Short literature review** The ability to obtain unique prices and hedging strategies, which is the strength of the model-specific approach, relies on its primary weakness—the necessity to postulate a fixed probability measure  $\mathbb{P}$  giving a full probabilistic description of future market dynamics. Put differently, this approach captures risks within a given model, but fails to tell us anything about the model uncertainty, also called *Knightian uncertainty*; see Knight [36, Chap. 2]. Accordingly, researchers have extended the classical setup to one where many measures  $\{\mathbb{P}_\alpha : \alpha \in \Lambda\}$  are simultaneously deemed feasible. This can be seen as weakening assumptions and going back from model-specific towards model-independent. The pioneering works considered *uncertain volatility*; see Lyons [37] and Avellaneda et al. [2]. More recently, a systematic approach based on quasi-sure analysis was developed, with stochastic integration based on capacity theory in Denis and Martini [23] and on the aggregation method in Soner et al. [48]; see also Neufeld and Nutz [41]. In discrete time,

<sup>1</sup>Merton [38].

a corresponding generalisation of the FTAP and the pricing–hedging duality was obtained by Bouchard and Nutz [10] and in continuous time by Biagini et al. [6]; see also references therein. We also mention that setups with frictions, e.g. trading constraints, were considered; see Bayraktar and Zhou [3].

In parallel, the model-independent approach has also seen a revived interest. This was mainly driven by the observation that with the increasingly rich market reality, this “universally acceptable” setting may actually provide outputs precise enough to be practically relevant. Indeed, in contrast to when Merton [38] was examining this approach, at present typically not only the underlying is liquidly traded, but so are many European options written on it. Accordingly, these should be treated as *inputs* and hedging instruments, thus reducing the possible universe of no-arbitrage scenarios. Breeden and Litzenberger [11] were the first to observe that if many (all) European options for a given maturity trade, then this is equivalent to fixing the marginal distribution of the stock under any EMM in the classical setting. Hobson [33] in his pioneering work then showed how this can be used to compute model-independent prices and hedges of lookback options. Other exotic options were analysed in subsequent works; see Brown et al. [12], Cox and Wang [18], Cox and Oblój [17]. The resulting no-arbitrage price bounds could still be too wide even for market making, but the associated hedging strategies were shown to perform remarkably well when compared to traditional delta–vega hedging; see Oblój and Ulmer [43]. Note that the superhedging property here is understood in a pathwise sense, and typically the strategies involve buy-and-hold positions in options and simple dynamic trading in the underlying. The universality of the setting and relative insensitivity of the outputs to the (few) assumptions earned this setup the name of *robust approach*.

In the wake of the financial crisis, significant research focus shifted back to the model-independent approach, and many natural questions, such as establishing the pricing–hedging duality and a (robust) version of the FTAP, were pursued. In a one-period setting, the pricing–hedging duality was linked to the Karlin–Isii duality in linear programming by Davis et al. [19]; see also Riedel [45]. Beiglböck et al. [5] re-interpreted the problem as a *martingale optimal transport problem* and established a general discrete-time pricing–hedging duality as an analogue of the Kantorovich duality in optimal transport. Here the primal elements are martingale measures, starting in a given point and having fixed marginal distribution(s) via the Breeden and Litzenberger [11] formula. The dual elements are sub- or superhedging strategies, and the payoff of the contingent claim is the “cost functional”. An analogous result in continuous time, under suitable continuity assumptions, was obtained by Dolinsky and Soner [25], who also more recently considered the discontinuous setting [26]. These topics remain an active field of research. Acciaio et al. [1] considered the pricing–hedging duality and the FTAP with an arbitrary market input in discrete time and under significant technical assumptions. These were relaxed, offering great insights, in a recent work of Burzoni et al. [14]. Galichon et al. [30] applied the methods of stochastic control to deduce the model-independent prices and hedges; see also Henry-Labordère et al. [32]. Several authors considered setups with frictions, e.g. transactions costs in Dolinsky and Soner [24] or trading constraints in Cox et al. [16] and Fahim and Huang [28].

**Main contribution** The present work contributes to the literature on robust pricing and hedging of contingent claims in two ways. First, inspired by Dolinsky and Soner [25], we study the pricing–hedging duality in continuous time and extend their results to multiple dimensions, different market setups and options with uniformly continuous payoffs. Our results are general and obtained in a comprehensive setting. We explicitly specify several important special cases, including the setting when finitely many options are traded, some dynamically and some statically, and the setting when all European call options for  $n$  maturities are traded. The latter gives the martingale optimal transport (MOT) duality with  $n$  marginal constraints which was also recently studied in a discontinuous setup by Dolinsky and Soner [26] and, in parallel to our work, by Guo et al. [31].

Our second main contribution is to propose a *robust approach*, which subsumes the model-independent setting, but allows us to include assumptions and move gradually towards the model-specific setting. In this sense, we strive to provide a setup which connects and interpolates between the two ends of the spectrum considered by Merton [38]. In contrast, all the above works on the model-independent approach stay within Merton [38]’s “universally accepted” setting and analyse the implications of incorporating the ability to trade some options at given market prices for the outputs, namely prices and hedging strategies of other contingent claims. We amend this setup and allow expressing modelling beliefs. These are articulated in a pathwise manner. More precisely, we allow the modeller to deem certain paths *impossible* and exclude them from the analysis; the superhedging property is only required to hold on the remaining set of paths  $\mathfrak{P}$ . This is reflected in the form of the pricing–hedging duality we obtain.

Our framework was inspired by Mykland’s [39] idea of incorporating a prediction set of paths into the pricing and hedging problem. On a philosophical level, we start with the “universally acceptable” setting and proceed by *ruling out* more and more scenarios as *impossible*; see also Cassese [15]. We may proceed in this way until we end up with paths supporting a unique martingale measure, e.g. a geometric Brownian motion, giving us essentially a model-specific setting. The hedging arguments are required to work for all the paths which remain under consideration, and a (strong) arbitrage would be given by a strategy which makes positive profit for all these paths. In discrete time, these ideas were recently explored by Burzoni et al. [13]. This should be contrasted with another way of interpolating between the model-independent and the model-specific, namely one which starts from a given model  $\mathbb{P}$  and proceeds by *adding* more and more *possible* scenarios  $\{\mathbb{P}_\alpha : \alpha \in \Lambda\}$ . This naturally leads to probabilistic (quasi-sure) hedging and different notions of no-arbitrage; see Bouchard and Nutz [10].

Our approach to establishing the pricing–hedging duality involves both discretisation, as in Dolinsky and Soner [25], as well as a variational approach as in Galichon et al. [30]. We first prove an “unconstrained” duality result: (3.4) states that for any derivative with bounded and uniformly continuous payoff function  $G$ , the minimal initial cost of setting up a portfolio consisting of cash and dynamic trading in the risky assets (some of which could be options themselves) which superhedges the payoff  $G$  for every nonnegative continuous path is equal to the supremum of the ex-

pected value of  $G$  over all nonnegative continuous martingale measures.<sup>2</sup> This result is shown through an elaborate discretisation procedure building on ideas in [25, 26]. Subsequently, we develop a variational formulation which allows us to add statically traded options, or the specification of a prediction set  $\mathfrak{P}$ , via Lagrange multipliers. In some cases, this leads to “constrained” duality results, similar to the ones obtained in the works cited above, with superhedging portfolios allowed to trade statically in market options and the martingale measures required to reprice these options. In particular, Theorems 3.6 and 3.12 extend the duality obtained in [19] and in [25], respectively. However, in general, we obtain an *asymptotic* duality result with the dual and primal problems defined through a limiting procedure. The primal value is the limit of superhedging prices on an  $\epsilon$ -neighbourhood of  $\mathfrak{P}$ , and the dual value is the limit of suprema of expectations of the payoff over  $\epsilon$ -(mis)calibrated models; see Definitions 2.1 and 3.11.

The paper is organised as follows. Section 2 introduces our robust framework for pricing and hedging and defines the primal (pricing) and dual (hedging) problems. Section 3 contains all the main results. First, in Sect. 3.1, we outline the unconstrained pricing–hedging duality, displayed in (3.4), and state the constrained (asymptotic) duality results under suitable compactness assumptions. This allows us in particular to treat the case of finitely many traded options. Then, in Sect. 3.2, we apply the previous results to the martingale optimal transport case. All the result except the main unconstrained duality in (3.4) are proved in Sect. 4. The former is stated in Theorem 5.1 and shown in Sect. 5. The proof proceeds via discretisation, of the primal problem in Sect. 5.1 and of the dual problem in Sect. 5.3, with Sect. 5.2 connecting the two via classical duality results. The proofs of two auxiliary results are relegated to the Appendix.

## 2 Robust modelling framework

### 2.1 Traded assets

We consider a financial market with  $d + 1$  primary assets: a numeraire (e.g. the money market account) and  $d$  underlying assets which may be traded at any time  $t \leq T$ . All prices are denominated in the units of the numeraire. In particular, the numeraire’s price is thus normalised and equal to one. The underlying assets’ price path is denoted  $S = ((S_t^{(1)}, \dots, S_t^{(d)}) : t \in [0, T])$ , starts in  $S_0 = (1, \dots, 1)$  and is assumed to be nonnegative and continuous in time. It is thus an element of the canonical space  $\mathcal{C}([0, T], \mathbb{R}_+^d)$  of all  $\mathbb{R}_+^d$ -valued continuous functions on  $[0, T]$ , which we endow with the supremum norm  $\|\cdot\|$ . Throughout, trading is frictionless.

We pursue a robust approach and do not postulate any probability measure which would specify the dynamics for  $S$ . Instead, we incorporate as inputs prices of traded derivative instruments. We assume that there is a set  $\mathcal{X}$  of market-traded options with prices  $\mathcal{P}(X)$ ,  $X \in \mathcal{X}$ , known at time zero. These options can be traded frictionlessly

<sup>2</sup>Note that here and throughout, we assume that all assets are discounted or, more generally, are expressed in terms of some numeraire.

at time zero, but are not assumed to be available for trading at future times. In particular, only buy-and-hold trading in these options is allowed (static trading). An option  $X \in \mathcal{X}$  is just a mapping  $X : \mathcal{C}([0, T], \mathbb{R}_+^d) \rightarrow \mathbb{R}$ , measurable with respect to the  $\sigma$ -field generated by the coordinate process. In the sequel, we only consider continuous payoffs  $X$  and often specialise further to European options, i.e.,  $X(S) = f(S_{T_i})$  for some  $f$  and  $0 < T_i \leq T$ .

Further, in addition to the above, we allow dynamically traded derivative assets. We consider  $K$  continuously traded European options and assume their prices evolve continuously and are strictly positive. The  $j$ th option has initial price  $\mathcal{P}(X_j^{(c)})$  and terminal payoff  $X_j^{(c)}(S_T^{(1)}, \dots, S_T^{(d)})$  at time  $T$ . Since trading is frictionless, we can and do consider traded options with renormalised payoffs  $X_j^{(c)}/\mathcal{P}(X_j^{(c)})$  and initial prices equal to 1. When we need to consider perturbations to options' prices, this then corresponds to a multiplicative perturbation of their payoffs; see e.g. Assumption 3.1 below. We thus have  $d + K$  dynamically traded assets whose price paths belong to

$$\Omega = \{S \in \mathcal{C}([0, T], \mathbb{R}_+^{d+K}) : S_0 = (1, \dots, 1)\}.$$

Their price process is given by the canonical process  $\mathbb{S} = (\mathbb{S}_t)_{0 \leq t \leq T}$  on  $\Omega$ , i.e.,  $\mathbb{S} = (\mathbb{S}^{(1)}, \dots, \mathbb{S}^{(d+K)}) : [0, T] \rightarrow \mathbb{R}_+^{d+K}$ . We let  $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$  be its natural raw filtration.<sup>3</sup> The subset of paths which respect to the *market information* about the future payoff constraints is given by

$$\mathcal{I} := \{S \in \Omega : S_T^{(d+i)} = X_i^{(c)}(S_T^{(1)}, \dots, S_T^{(d)})/\mathcal{P}(X_i^{(c)}), i = 1, \dots, K\}.$$

We sometimes refer to  $\mathcal{I}$  as the information space. For a random variable  $G$  on  $\Omega$ , we clearly have  $G = G \circ \mathbb{S}$ , and we exploit this to write  $G(\mathbb{S})$  instead of simply  $G$  when we want to stress that  $G$  is seen as a function of the assets' path. It is also convenient to think of  $X \in \mathcal{X}$  as functions on  $\Omega$  with  $X(S) = X(S^1, \dots, S^d)$ . We only consider continuous  $X$ , i.e.,  $\mathcal{X} \subseteq \mathcal{C}(\Omega, \mathbb{R})$  with  $\|X\|_\infty := \sup\{|X(S)| : S \in \Omega\}$ . We write  $\mathcal{X} = \emptyset$  to indicate the situation with no statically traded options, and  $K = 0$  to indicate when there are no dynamically traded options.

### 2.2 Beliefs

As argued in the introduction, we allow our agents to express modelling beliefs. These are encoded as restrictions of the path space and may come from time series analysis of past data, or idiosyncratic views about the market in the future. Put differently, we are allowed to rule out paths which we deem *impossible*. The paths which remain are referred to as *prediction set* or *beliefs*. Note that such beliefs may also encode one agent's superior information about the market. As the agent rejects more and more paths, the framework's outputs—the robust price bounds—should get tighter and tighter. This can be seen as a way to quantify the impact of making assumptions or acquiring additional insights or information.

<sup>3</sup>All of our results remain valid if one takes the right-continuous version of the filtration instead.

The choice of paths is expressed by the prediction set  $\mathfrak{P} \subseteq \mathcal{I}$ . Our arguments are required to work pathwise on  $\mathfrak{P}$ , while paths in the complement of  $\mathfrak{P}$  are ignored in our considerations. This binary way of specifying beliefs is motivated by the fact that in the end, we only see one path and hence are interested in arguments which work pathwise. Nevertheless, the approach is very comprehensive, and as  $\mathfrak{P}$  changes from all paths in  $\mathcal{I}$  to the support of a given model, we essentially interpolate between model-independent and model-specific setups. It also allows incorporating the information from time series of data coherently into the option pricing setup, as no probability measure is fixed and hence no distinction between real-world and risk-neutral measures is made. The idea of such a prediction set first appeared in Mykland [39]; see also Naddochiy and Oblój [40] and Cox et al. [16] for an extended discussion.

### 2.3 Trading strategies and superreplication

We consider two types of trading: buy-and-hold strategies in options in  $\mathcal{X}$  and dynamic trading in assets  $S^{(i)}$ ,  $i \leq d + K$ . The gains from the latter take the integral form  $\int \gamma_u(S) dS_u$ , and to define this integral pathwise, we need to impose suitable restrictions on  $\gamma$ . We may, following Dolinsky and Soner [25], take  $\gamma = (\gamma_t)_{0 \leq t \leq T}$  to be an  $\mathbb{F}$ -progressively measurable process of finite variation and use the integration by parts formula to define

$$\int_0^t \gamma_u(S) dS_u := \gamma_t \cdot S_t - \gamma_0 \cdot S_0 - \int_0^t S_u d\gamma_u, \quad S \in \Omega,$$

where we write  $a \cdot b$  to denote the usual scalar product for any  $a, b \in \mathbb{R}^{d+K}$  and the last term on the right-hand side is a Stieltjes integral. However, for our duality, it is sufficient to consider a smaller class of processes. Namely, we say that  $\gamma$  is *admissible* if it is  $\mathbb{F}$ -adapted,  $\gamma(S)$  is a simple, i.e., right-continuous and piecewise constant, function for all  $S \in \Omega$ , and

$$\int_0^t \gamma_u(S) dS_u \geq -M, \quad \forall S \in \mathcal{I}, t \in [0, T], \text{ for some } M > 0. \tag{2.1}$$

We denote by  $\mathcal{A}$  the set of such integrands  $\gamma$ . To define static trading, we consider

$$\text{Lin}_N(\mathcal{X}) = \left\{ a_0 + \sum_{i=1}^m a_i X_i : m \in \mathbb{N}, X_i \in \mathcal{X}, a_i \in \mathbb{R}, \sum_{i=0}^m |a_i| \leq N \right\}, \tag{2.2}$$

$$\text{Lin}(\mathcal{X}) = \bigcup_{N \geq 1} \text{Lin}_N(\mathcal{X}).$$

An *admissible (semi-static) trading strategy* is a pair  $(X, \gamma)$  with  $X \in \text{Lin}(\mathcal{X})$  and  $\gamma \in \mathcal{A}$ . We denote the class of such trading strategies by  $\mathcal{A}_{\mathcal{X}}$ . The cost of following  $(X, \gamma) \in \mathcal{A}_{\mathcal{X}}$  is equal to the cost of setting up its static part, i.e., of buying the options at time zero, and is given by

$$\mathcal{P}(X) := a_0 + \sum_{i=1}^m a_i \mathcal{P}(X_i), \quad \text{for } X = a_0 + \sum_{i=1}^m a_i X_i. \tag{2.3}$$

**Throughout, we assume that the above defines  $\mathcal{P}$  uniquely as a linear operator on  $\text{Lin}(\mathcal{X})$ .** This is in particular true if the elements in  $\mathcal{X}$  are linearly independent. Further, to eliminate obvious arbitrages, we assume that for  $X \in \mathcal{X}$ , we have  $\mathcal{P}(X) \leq \|X\|_\infty$ , which by (2.3) then holds for all  $X \in \text{Lin}(\mathcal{X})$ . It follows that  $\mathcal{P}$  is bounded linear, and hence continuous, on  $(\text{Lin}(\mathcal{X}), \|\cdot\|_\infty)$ . Note that with our definitions,  $\text{Lin}(\emptyset) = \mathbb{R}$  and  $\mathcal{P}(a) = a$  for  $a \in \mathbb{R}$ . We also note that dynamically traded assets can be traded statically; so our previous notation  $\mathcal{P}(X_j^{(c)})$  is consistent.

Our prime interest is in understanding robust pricing and hedging of a non-liquidity traded derivative with payoff  $G : \Omega \rightarrow \mathbb{R}$ . Our main results consider bounded payoffs  $G$ , and since the setup is frictionless and there are no trading restrictions, without any loss of generality, we may consider only the superhedging price. The subhedging follows by considering  $-G$ .

**Definition 2.1** Consider  $G : \Omega \rightarrow \mathbb{R}$ . A portfolio  $(X, \gamma) \in \mathcal{A}_{\mathcal{X}}$  is said to *superreplicate*  $G$  on  $\mathfrak{F}$  if

$$X(S) + \int_0^T \gamma_u(S) dS_u \geq G(S), \quad \forall S \in \mathfrak{F}.$$

The (minimal) *superreplication cost or superhedging price* of  $G$  on  $\mathfrak{F}$  is defined as

$$V_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}(G) := \inf\{\mathcal{P}(X) : \exists(X, \gamma) \in \mathcal{A}_{\mathcal{X}} \text{ which superreplicates } G \text{ on } \mathfrak{F}\}.$$

The *approximate superreplication cost* of  $G$  on  $\mathfrak{F}$  is defined as

$$\begin{aligned} \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}(G) := \inf\{\mathcal{P}(X) : \exists(X, \gamma) \in \mathcal{A}_{\mathcal{X}} \text{ which} \\ \text{superreplicates } G \text{ on } \mathfrak{F}^\epsilon \text{ for some } \epsilon > 0\}, \end{aligned}$$

where  $\mathfrak{F}^\epsilon = \{\omega \in \mathcal{I} : \inf_{\nu \in \mathfrak{F}} \|\omega - \nu\| \leq \epsilon\}$ .

As we shall see below, the approximate superreplicating cost appears naturally as the correct object to obtain a duality with general  $\mathfrak{F}$  and  $\mathcal{X}$ . We note, however, that ex post, it is also a natural object from the financial point of view: It requires the superreplication to be robust with respect to an arbitrarily small perturbation of the beliefs.

Note that by definition,  $\mathcal{I}^\epsilon = \mathcal{I}$  and consequently  $V_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G) = \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G)$ . Finally, we denote by  $\mathbf{V}_{\mathcal{I}}(G) = V_{\emptyset, \mathcal{P}, \mathcal{I}}(G)$  the superreplication cost of  $G$  in the absence of constraints.

### 2.4 Market models

Our aim is to relate the robust superhedging price as introduced above to the classical pricing-by-expectation arguments. To this end, we look at all classical models which reprice market-traded options.



**Definition 2.2** We denote by  $\mathcal{M}$  the set of probability measures  $\mathbb{P}$  on  $(\Omega, \mathcal{F}_T)$  such that  $\mathbb{S}$  is an  $(\mathbb{F}, \mathbb{P})$ -martingale and let  $\mathcal{M}_{\mathcal{I}}$  be the set of probability measures  $\mathbb{P} \in \mathcal{M}$  such that  $\mathbb{P}[\mathcal{I}] = 1$ . A probability measure  $\mathbb{P} \in \mathcal{M}_{\mathcal{I}}$  is called an  $(\mathcal{X}, \mathcal{P}, \mathfrak{P})$ -market model or simply a calibrated model if  $\mathbb{P}[\mathfrak{P}] = 1$  and  $\mathbb{E}_{\mathbb{P}}[X] = \mathcal{P}(X)$  for all  $X \in \mathcal{X}$ . The set of such measures is denoted by  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}$ . More generally, a probability measure  $\mathbb{P} \in \mathcal{M}_{\mathcal{I}}$  is called an  $\eta$ - $(\mathcal{X}, \mathcal{P}, \mathfrak{P})$ -market model if  $\mathbb{P}[\mathfrak{P}^\eta] > 1 - \eta$  and  $|\mathbb{E}_{\mathbb{P}}[X] - \mathcal{P}(X)| < \eta$  for all  $X \in \mathcal{X}$ . The set of such measures is denoted by  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^\eta$ .

Any  $\mathbb{P} \in \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}$  provides us with a feasible no-arbitrage price  $\mathbb{E}_{\mathbb{P}}[G]$  for a derivative with payoff  $G$ . The robust price for  $G$  is given as

$$P_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) := \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}} \mathbb{E}_{\mathbb{P}}[G], \tag{2.4}$$

where **throughout**, the expectation is defined with the **convention** that  $\infty - \infty = -\infty$ . In the cases of particular interest,  $(\mathcal{X}, \mathcal{P})$  uniquely determines the marginal distributions of  $\mathbb{S}$  at given maturities, and  $P_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$  is then the value of the corresponding martingale optimal transport problem. We often use this terminology, even in the case of an arbitrary  $\mathcal{X}$ . Finally, in the special case where there are no constraints, i.e.,  $\mathcal{X} = \emptyset$  and  $\mathfrak{P} = \mathcal{I}$ , we write  $\mathbf{P}_{\mathcal{I}}(G)$  to represent the corresponding maximal modelling value, i.e.,

$$\mathbf{P}_{\mathcal{I}}(G) = P_{\emptyset, \mathcal{P}, \mathcal{I}}(G) = \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})].$$

We shall see below that with a general  $\mathcal{X}$  and  $\mathfrak{P}$ , we do not obtain a duality using  $P_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$  in (2.4), but rather have to consider its approximate value given as

$$\tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) := \lim_{\eta \searrow 0} \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^\eta} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})].$$

Ex post, and similarly to the approximate superhedging above, this may be seen as a natural robust object to consider: instead of requiring a perfect calibration, the concept of  $\eta$ -market model allows a controlled degree of mis-calibration. This seems practically relevant since the market prices  $\mathcal{P}$  are an idealised concept obtained e.g. via averaging of bid–ask spreads.

### 3 Main results

Our prime interest, as discussed in the introduction, is in establishing a general robust pricing–hedging duality. Given a non-traded derivative with payoff  $G$ , we have two candidate robust prices for it. The first one,  $V_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$ , is obtained through pricing-by-hedging arguments. The second one,  $P_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$ , is obtained by pricing-via-expectation arguments. In a classical setting, by fundamental results, see e.g. Delbaen and Schachermayer [20, Theorem 5.7], the analogous two prices are equal.

Within the present pathwise robust approach, the pricing–hedging duality was obtained for specific payoffs  $G$  in the literature linking the robust approach with the Skorokhod embedding problem; see Hobson [33] or Oblój [42] for a discussion. Subsequently, an abstract result was established in Dolinsky and Soner [25]. For  $d = 1$ ,  $K = 0$ ,  $\mathfrak{P} = \mathcal{I} = \Omega$  and  $\mathcal{X}$  the set of all call (or put) options with a common maturity  $T$  and with  $\mathcal{P}(X) = \int_0^\infty X(x)\mu(dx)$ ,  $\forall X \in \mathcal{X}$ , where  $\mu$  is a probability measure on  $\mathbb{R}_+$  with mean equal to 1, they showed that

$$V_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G) = P_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G) \quad \text{for a “strongly continuous” class of bounded } G.$$

The result was extended to unbounded claims by broadening the class of admissible strategies and imposing a technical assumption on  $\mu$ . We extend this duality to a much more general setting of abstract  $\mathcal{X}$ , possibly involving options with multiple maturities, a multidimensional setting, and with an arbitrary prediction set  $\mathfrak{P}$ . However, in this generality, the duality may only hold between approximate values. We first give the statements, illustrated with examples, and all the proofs are postponed to Sect. 4.

Note that, for any Borel  $G : \Omega \rightarrow \mathbb{R}$ , the inequality

$$V_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) \geq P_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) \tag{3.1}$$

is true as long as there is at least one  $\mathbb{P} \in \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}$  and at least one  $(X, \gamma) \in \mathcal{A}_{\mathcal{X}}$  which superreplicates  $G$  on  $\mathfrak{P}$ . Indeed, since  $\gamma$  is progressively measurable, the integral  $\int_0^\cdot \gamma_u(S) dS_u$ , defined pathwise via integration by parts, agrees a.s. with the stochastic integral under  $\mathbb{P}$ . Then by (2.1), the stochastic integral is a  $\mathbb{P}$ -supermartingale and hence  $\mathbb{E}_{\mathbb{P}}[\int_0^T \gamma_u(S) dS_u] \leq 0$ . This in turn implies that  $\mathbb{E}_{\mathbb{P}}[G] \leq \mathcal{P}(X)$ . The result follows since  $(X, \gamma)$  and  $\mathbb{P}$  were arbitrary. The converse inequality, however, is very involved. The fundamental difficulty lies in the fact that even in the martingale optimal transport setting of [25], the set  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathcal{I}}$  is not compact. This is in contrast to the discrete-time case; see [5]. In our general setting, the converse inequality to (3.1) may fail, see Example 3.7 below, making it necessary to look at the duality between the approximate values.

### 3.1 General duality

We start with a general duality between the approximate values  $\tilde{V}, \tilde{P}$ . But first, we give our **standing assumption** which states that the prices of dynamically traded options are not “on the boundary of the no-arbitrage region”, i.e., calibrated martingale measures exist under arbitrarily small perturbation of the initial prices.

**Assumption 3.1** Either  $K = 0$  or  $X_1^{(c)}, \dots, X_K^{(c)}$  are bounded and uniformly continuous with market prices  $\mathcal{P}(X_1^{(c)}), \dots, \mathcal{P}(X_K^{(c)})$  satisfying that there exists an  $\epsilon > 0$  such that for any  $(p_i)_{1 \leq i \leq K}$  with  $|\mathcal{P}(X_i^{(c)}) - p_i| \leq \epsilon$  for all  $i \leq K$ , we have  $\mathcal{M}_{\tilde{\mathcal{I}}} \neq \emptyset$ , where

$$\tilde{\mathcal{I}} := \{S \in \Omega : S_T^{(d+i)} = X_i^{(c)}(S_T^{(1)}, \dots, S_T^{(d)})/p_i \text{ for all } 1 \leq i \leq K\}.$$

**Theorem 3.2** *Assume that  $\mathfrak{P}$  is a measurable subset of  $\mathcal{I}$ , Assumption 3.1 holds, all  $X \in \mathcal{X}$  are uniformly continuous and bounded, and  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^\eta \neq \emptyset$  for any  $\eta > 0$ . Then for any uniformly continuous and bounded  $G : \Omega \rightarrow \mathbb{R}$ , we have*

$$\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) \geq \tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G), \tag{3.2}$$

and if  $\text{Lin}_1(\mathcal{X})$  defined in (2.2) is a compact subset of  $(\mathcal{C}(\Omega, \mathbb{R}), \|\cdot\|_\infty)$ , then equality holds, i.e.,

$$\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = \tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G). \tag{3.3}$$

*Remark 3.3* The above remains true if instead of martingale measures in  $\mathcal{M}$ , we restrict to Brownian martingales; see Sect. 4.1.

*Example 3.4* **Finite  $\mathcal{X}$**  Consider  $\mathcal{X} = \{X_1, \dots, X_m\}$ , where the  $X_i$  are bounded and uniformly continuous. In this case,  $\text{Lin}_1(\mathcal{X})$  is a convex and compact subset of  $\mathcal{C}(\Omega, \mathbb{R})$ . Therefore, if  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^\eta \neq \emptyset$  for any  $\eta > 0$ , we can apply Theorem 3.2 to conclude that  $\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = \tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$ .

Let us outline the proof of Theorem 3.2. The first inequality in (3.2) is relatively easy to obtain, and the main effort is in establishing the converse inequality which yields (3.3). This is done in two steps. First, we consider the case without constraints, i.e.,  $\mathcal{X} = \emptyset$  and  $\mathfrak{P} = \mathcal{I}$ . The approximate values  $\tilde{V}, \tilde{P}$  then reduce to  $V$  and  $P$  respectively; so we need to show that for any bounded and uniformly continuous  $G : \Omega \rightarrow \mathbb{R}$ , we have

$$\mathbf{V}_{\mathcal{I}}(G) = \mathbf{P}_{\mathcal{I}}(G). \tag{3.4}$$

This result is a special case of Theorem 5.1, which is shown in Sects. 5.1 and 5.3. Our proof proceeds through discretisation of both the primal and the dual problem and is inspired by the methods in [25, 26] but involves significant technical differences which are necessary to obtain our more general results. The first key difference, when comparing with [25], is that the discretisation therein entangles discretisation of the dynamic hedging part and static hedging part, while we develop a “clean” decoupled discretisation of the dynamic hedging part only. Second, we have to improve the discretisation to deal with payoff functions which are uniformly continuous. This is crucial for the subsequent use of a variational approach to generalise the pricing–hedging duality results to include static hedging in options with different maturities. The time-continuity assumptions on the payoff made in [25] are much stronger, and their results could not be applied directly in our framework.

We note also that in a quasi-sure setting, an analogue of (3.4) was obtained in Possamai et al. [44] and earlier papers, as discussed therein. However, while similar in spirit, there is no immediate link between our results or proofs and those in [44]. Here, we consider a comparatively smaller set of admissible trading strategies and require a pathwise superhedging property. Consequently, we also need to impose stronger regularity constraints on  $G$ .

In the second step of the proof, we use a variational approach combined with a minimax argument to obtain duality under all the constraints. Specifically, in analogy

to e.g. Proposition 5.2 in Henry-Labordère et al. [32], for any uniformly continuous and bounded  $G : \Omega \rightarrow \mathbb{R}$ , we can write

$$\begin{aligned} \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) &= \inf_{X \in \text{Lin}_N(\mathcal{X}), N \geq 0} (\mathbf{V}_{\mathcal{I}}(G - X - N\lambda_{\mathfrak{P}}) + \mathcal{P}(X)) \\ &= \inf_{X \in \text{Lin}_N(\mathcal{X}), N \geq 0} (\mathbf{P}_{\mathcal{I}}(G - X - N\lambda_{\mathfrak{P}}) + \mathcal{P}(X)), \end{aligned} \tag{3.5}$$

where  $\lambda_{\mathfrak{P}}(\omega) := \inf_{v \in \mathfrak{P}} \|\omega - v\| \wedge 1$ . An application of the minimax theorem then yields

$$\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{I}}} \inf_{X \in \text{Lin}_N(\mathcal{X})} (\mathbb{E}_{\mathbb{P}}[G - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X)).$$

The final, and somewhat technical, argument is to show that the above is dominated by  $\tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$ .

We end this section with a study of the relation between  $V, P$  and their approximate values  $\tilde{V}, \tilde{P}$ . As already noted, by definition,  $\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) \geq V_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$  and  $\tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) \geq P_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$ . Therefore, when  $\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = \tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$ , the duality  $V_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = P_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$  follows if we can show that  $P_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = \tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$ . We establish this equality for an important family of market setups, but also provide examples when it fails.

Consider first the case with no specific beliefs,  $\mathfrak{P} = \mathcal{I}$ , and finitely many traded put options with maturities  $0 < T_1 < \dots < T_n = T$ , i.e.,

$$\mathcal{X} = \{(K_{k,j}^{(i)} - \mathbb{S}_{T_j}^{(i)})^+ : 1 \leq i \leq d, 1 \leq j \leq n, 1 \leq k \leq m(i, j)\}, \tag{3.6}$$

where  $0 < K_{k,j}^{(i)} < K_{k',j}^{(i)}$  for any  $k < k'$  and  $m(i, j) \in \mathbb{N}$ . To simplify the notation, we write

$$\mathcal{P}((K_{k,j}^{(i)} - \mathbb{S}_{T_j}^{(i)})^+) = p_{k,i,j}, \quad \forall i, j, k.$$

In analogy to Assumption 3.1, we need to impose that the put prices are in the interior of the no-arbitrage region.

**Assumption 3.5** Market put prices are such that there exists an  $\epsilon > 0$  such that for any  $(\tilde{p}_{k,i,j})_{i,j,k}$  with  $|\tilde{p}_{k,i,j} - p_{k,i,j}| \leq \epsilon$  for all  $i, j, k$ , there exists a  $\tilde{\mathbb{P}} \in \mathcal{M}_{\mathcal{I}}$  with

$$\tilde{p}_{k,i,j} = \mathbb{E}_{\tilde{\mathbb{P}}}[(K_{k,j}^{(i)} - \mathbb{S}_{T_j}^{(i)})^+], \quad \forall i, j, k.$$

**Theorem 3.6** Let  $\mathcal{X}$  be given by (3.6) and assume that the market prices satisfy Assumptions 3.1 and 3.5. Then for any uniformly continuous and bounded  $G : \Omega \rightarrow \mathbb{R}$ , we have

$$V_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G) = P_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G).$$

The above result establishes a general robust pricing–hedging duality when finitely many put options are traded. It extends in many ways the duality obtained in Davis

et al. [19] for  $d = n = 1$  and  $K = 0$ . Note that in general,  $\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G) = V_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G)$ ; so it follows from Example 3.4 that we also have  $\tilde{P}_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G) = P_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G)$  in Theorem 3.6. These equalities may still hold, but may also fail, when nontrivial beliefs are specified. We present now three examples to highlight various possible scenarios. In the first two examples, for different reasons, the pricing–hedging duality fails, i.e.,  $V_{\mathfrak{F}} > P_{\mathfrak{F}}$ , while the approximate duality (3.3) holds. In the last example, all quantities are equal.

*Example 3.7* Consider the case when there are no traded options,  $\mathcal{X} = \emptyset$ ,  $K = 0$  and  $d = 1$ , and let

$$\mathfrak{F} = \{S \in \Omega : S_T \leq 2\}.$$

Define  $G : \Omega \rightarrow \mathbb{R}$  by  $G(S) = (\max_{0 \leq t \leq T} S_t - 4)^+ \wedge 1$ . Theorem 3.2 implies that  $\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}(G) = \tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}(G)$ . On the other hand, it is straightforward to see that  $P_{\mathfrak{F}}(G) = 0$ . By letting  $\mathcal{M}_{\mathfrak{F}}^{\text{loc}}$  be the set of  $\mathbb{P}$  such that  $\mathbb{S}$  is a  $\mathbb{P}$ -local martingale and  $\mathbb{P}[\mathfrak{F}] = 1$ , we further have

$$V_{\mathfrak{F}}(G) \geq \sup_{\mathbb{P} \in \mathcal{M}_{\mathfrak{F}}^{\text{loc}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] > 0 = P_{\mathfrak{F}}(G).$$

*Example 3.8* In this example, we consider  $\mathfrak{F}$  corresponding to the Black–Scholes model. For simplicity, consider the case without any traded options, i.e.,  $K = 0$ ,  $\mathcal{X} = \emptyset$ ,  $d = 1$ , and let<sup>4</sup>

$$\mathfrak{F} = \{S \in \Omega : S \text{ admits a quadratic variation and } d(S)_t = \sigma^2 S_t^2 dt, 0 \leq t \leq T\}.$$

Then  $\mathcal{M}_{\mathfrak{F}} = \{\mathbb{P}_{\sigma}\}$ , where  $\mathbb{S}$  is a geometric Brownian motion with constant volatility  $\sigma$  under  $\mathbb{P}_{\sigma}$ . The duality in Theorem 3.2 then gives that for any bounded and uniformly continuous  $G$ ,

$$\begin{aligned} \tilde{V}_{\mathfrak{F}}(G) &= \inf\{x : \exists \gamma \in \mathcal{A} \text{ which superreplicates } G - x \text{ on } \mathfrak{F}^{\epsilon} \text{ for some } \epsilon > 0\} \\ &= \lim_{\eta \searrow 0} \sup_{\mathbb{P} \in \mathcal{M}_{\mathfrak{F}}^{\eta}} \mathbb{E}_{\mathbb{P}}[G]. \end{aligned}$$

However, in this case,  $\mathbb{P}_{\sigma}$  has full support on  $\Omega$  so that  $\mathfrak{F}^{\epsilon} = \Omega$  and  $\mathcal{M}_{\mathfrak{F}}^{\epsilon} = \mathcal{M}$  for any  $\epsilon > 0$ . The above then boils down to the case with no beliefs, and we have

$$V_{\mathcal{I}}(G) = \tilde{V}_{\mathfrak{F}}(G) = \tilde{P}_{\mathfrak{F}}(G) = \sup_{\mathbb{P} \in \mathcal{M}} \mathbb{E}_{\mathbb{P}}[G] \geq \mathbb{E}_{\mathbb{P}_{\sigma}}[G] = P_{\mathfrak{F}}(G),$$

where for most  $G$  the inequality is strict.

*Example 3.9* Consider again the case with no traded options,  $K = 0$  and  $\mathcal{X} = \emptyset$ , and let

$$\mathfrak{F} = \{S \in \Omega : \|S\| \leq b\} \quad \text{for some } b \geq 1.$$

<sup>4</sup>A rigorous pathwise definition of such  $\mathfrak{F}$  is possible; see Step 4 in the proof of Theorem 3.16 in Sect. 4.

Given a bounded and uniformly continuous payoff function  $G$ , consider the duality in Theorem 3.2. For each  $N \in \mathbb{N}$ , we pick  $\mathbb{P}^{(N)} \in \mathcal{M}_{\mathfrak{P}}^{1/N}$  such that

$$\mathbb{E}_{\mathbb{P}^{(N)}}[G] \geq \sup_{\mathbb{P} \in \mathcal{M}_{\mathfrak{P}}^{1/N}} \mathbb{E}_{\mathbb{P}}[G] - 1/N.$$

Let  $\tau$  be the first hitting time of  $b + 1/N$  by  $\mathbb{S}$  and define  $\tilde{\mathbb{S}}^{(N)}$  by

$$\tilde{\mathbb{S}}_t^{(N)} = \mathbb{S}_0 + \frac{b}{b + 1/N} (\mathbb{S}_{t \wedge \tau} - \mathbb{S}_0).$$

By definition,  $\mathbb{P}^{(N)} \circ (\tilde{\mathbb{S}}^{(N)})^{-1} \in \mathcal{M}_{\mathfrak{P}}$ . Also note that  $\mathbb{P}^{(N)}[\tau < T] \leq 1/N$ . Hence by uniform continuity of  $G$ , it is straightforward to see that

$$|\mathbb{E}_{\mathbb{P}^{(N)}}[G(\tilde{\mathbb{S}}^{(N)})] - \mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S})]| \longrightarrow 0 \quad \text{as } N \rightarrow \infty,$$

which leads to  $\tilde{P}_{\mathfrak{P}}(G) = P_{\mathfrak{P}}(G)$ . As  $V_{\mathfrak{P}}(G) \leq \tilde{V}_{\mathfrak{P}}(G) = \tilde{P}_{\mathfrak{P}}(G)$ , we then conclude that

$$\tilde{V}_{\mathfrak{P}}(G) = \tilde{P}_{\mathfrak{P}}(G) = P_{\mathfrak{P}}(G) = V_{\mathfrak{P}}(G).$$

### 3.2 Martingale optimal transport duality

We focus now on the case when  $(\mathcal{X}, \mathcal{P})$  determines the marginal distributions of  $\mathbb{S}_{T_j}^{(i)}$  for  $i \leq d$  and given maturities  $0 < T_1 < \dots < T_n = T$ . For concreteness, let us consider the case when put options are traded, i.e.,

$$\mathcal{X} = \{(\kappa - \mathbb{S}_{T_j}^{(i)})^+ : i = 1, \dots, d, j = 1, \dots, n, \kappa \in \mathbb{R}_+\}. \tag{3.7}$$

Arbitrage considerations, see e.g. Cox and Oblój [17] and Cox et al. [16], show that absence of (a weak type of) arbitrage is equivalent to  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathcal{I}} \neq \emptyset$ . Note that the latter is equivalent to market prices  $\mathcal{P}$  being encoded by a vector  $\boldsymbol{\mu}$  of probability measures  $(\mu_j^{(i)})$  with

$$p_{i,j}(\kappa) = \mathcal{P}((\kappa - \mathbb{S}_{T_j}^{(i)})^+) = \int (\kappa - s)^+ \mu_j^{(i)}(ds), \tag{3.8}$$

where for each  $i = 1, \dots, d$ ,  $\mu_1^{(i)}, \dots, \mu_n^{(i)}$  have finite first moments, mean 1 and increase in *convex order* (written as  $\mu_1^{(i)} \leq \mu_2^{(i)} \leq \dots \leq \mu_n^{(i)}$ ), i.e., we have  $\int \phi(x) \mu_1^{(i)}(dx) \leq \dots \leq \int \phi(x) \mu_n^{(i)}(dx)$  for any convex function  $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}$ . In fact, as noted already by Breeden and Litzenberger [11], the  $\mu_j^{(i)}$  are defined by

$$\mu_j^{(i)}([0, \kappa]) = p'_{i,j}(\kappa+) \quad \text{for } \kappa \in \mathbb{R}_+.$$

We may think of  $(\mu_j^{(i)})$  and  $\mathfrak{P}$  as the modelling inputs. The set of calibrated market models  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}$  is simply the set of probability measures  $\mathbb{P} \in \mathcal{M}$  such

that  $\mathbb{S}_{T_j}^{(i)}$  is distributed according to  $\mu_j^{(i)}$  and  $\mathbb{P}[\mathfrak{P}] = 1$ . Accordingly, we write  $\mathcal{M}_{\boldsymbol{\mu}, \mathfrak{P}} = \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}$  and  $P_{\boldsymbol{\mu}, \mathfrak{P}}(G) = P_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$ .

*Remark 3.10* It follows, see Strassen [49], that  $\mathcal{M}_{\boldsymbol{\mu}, \mathcal{I}}$  is nonempty if and only if  $\mu_1^{(i)}, \dots, \mu_n^{(i)}$  have finite first moments, mean 1 and increase in convex order, for any  $i = 1, \dots, d$ . However, in general, the additional constraints associated with a nontrivial  $\mathfrak{P} \subsetneq \mathcal{I}$  are much harder to understand.

In this context, we can improve Theorem 3.2 and narrow down the class of approximate market models by requiring that they match exactly the marginal distributions at the last maturity.

**Definition 3.11** Let  $\mathcal{M}_{\boldsymbol{\mu}, \mathfrak{P}}^\eta$  be the set of all measures  $\mathbb{P} \in \mathcal{M}$  such that  $\mathcal{L}_{\mathbb{P}}(\mathbb{S}_{T_j}^{(i)})$ , the law of  $\mathbb{S}_{T_j}^{(i)}$  under  $\mathbb{P}$ , satisfies

$$\mathcal{L}_{\mathbb{P}}(\mathbb{S}_{T_n}^{(i)}) = \mu_n^{(i)} \text{ and } d_p(\mathcal{L}_{\mathbb{P}}(\mathbb{S}_{T_j}^{(i)}), \mu_j^{(i)}) \leq \eta, \quad \text{for } j = 1, \dots, n - 1, i = 1, \dots, d,$$

and furthermore  $\mathbb{P}[\mathfrak{P}^\eta] \geq 1 - \eta$ , where  $d_p$  is the Lévy–Prokhorov metric on probability measures. Finally, let

$$\tilde{P}_{\boldsymbol{\mu}, \mathfrak{P}}(G) := \lim_{\eta \searrow 0} \sup_{\mathbb{P} \in \mathcal{M}_{\boldsymbol{\mu}, \mathfrak{P}}^\eta} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})].$$

Note that  $\mathcal{M}_{\boldsymbol{\mu}, \mathfrak{P}} = \mathcal{M}_{\boldsymbol{\mu}, \mathfrak{P}}^0 \subseteq \mathcal{M}_{\boldsymbol{\mu}, \mathfrak{P}}^\eta \subseteq \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^{\epsilon(\eta)}$  for a suitable choice<sup>5</sup> of  $\epsilon(\eta)$  converging to zero as  $\eta \rightarrow 0$ . It follows that  $P_{\boldsymbol{\mu}, \mathfrak{P}}(G) \leq \tilde{P}_{\boldsymbol{\mu}, \mathfrak{P}}(G) \leq \tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$ . The following result extends and sharpens the duality obtained in Theorem 3.2 to the current setting.

**Theorem 3.12** Under Assumption 3.1, let  $\mathfrak{P}$  be a measurable subset of  $\mathcal{I}, \mathcal{X}$  given by (3.7) and  $\mathcal{P}$  such that for any  $\eta > 0$ ,  $\mathcal{M}_{\boldsymbol{\mu}, \mathfrak{P}}^\eta \neq \emptyset$ , where  $\boldsymbol{\mu}$  is defined via (3.8). Then for any uniformly continuous and bounded  $G$ , the robust pricing–hedging duality holds between the approximate values, i.e.,

$$\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = \tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = \tilde{P}_{\boldsymbol{\mu}, \mathfrak{P}}(G).$$

*Remark 3.13* Theorem 3.12 readily extends to unbounded exotic options, e.g. look-back options, following the approach of Dolinsky and Soner [25]. Fix  $p > 1$  and relax the admissibility in (2.1) to

$$\int_0^t \gamma_u(S) dS_u \geq -M \left( 1 + \sup_{0 \leq s \leq t} |S_s|^p \right), \quad \forall S \in \mathcal{I}, t \in [0, T], \text{ for some } M > 0.$$

<sup>5</sup>One can take  $\epsilon(\eta) = \sqrt{\eta} + 2f(1/\sqrt{\eta})$  with  $f(x) = \max_{1 \leq i \leq d} (p_i, n(x) - x + 1)$ .

Likewise, assume that all  $\mu_j^{(i)}$  admit a finite  $p$ th moment and allow static trading in European options with payoffs which grow at most as  $|x|^p$ . Then the duality in Theorem 3.12 extends to uniformly continuous  $G$  with  $|G(S)| \leq \text{const}(1 + \sup_{0 \leq t \leq T} |S_t|^p)$ .

In the case of one maturity,  $n = 1$ , we have  $\tilde{P}_{\mu, \mathcal{I}}(G) = P_{\mu, \mathcal{I}}(G)$ . In particular, Theorem 3.12 extends the duality of Dolinsky and Soner [25] by allowing arbitrary dimension  $d$ . It is also possible to consider a multidimensional extension where the whole marginal distribution  $\mathcal{L}_{\mathbb{P}}(\mathbb{S}_T)$  is fixed, or equivalently  $\mathcal{X}$  is large enough, e.g. dense in the Lipschitz-continuous functions on  $\mathbb{R}^d$ . For  $n = 1$  and  $\mathfrak{F} = \mathcal{I}$ , such an extension follows via Theorem 3.2 and Lemma 4.3; see Hou [34, Sect. 3.2 of Chap. 5].

**Assumption 3.14**  $G$  is bounded and uniformly continuous, and such that there exists a constant  $L > 0$  such that for all  $\mathbb{R}_+^{d+K}$ -valued functions  $v, \hat{v}$  of the form

$$v_t = \sum_{i=1}^n \sum_{j=0}^{m_i-1} v_{i,j} \mathbb{1}_{[t_{i,j}, t_{i,j+1})}(t) + v_{n,m_n-1} \mathbb{1}_{T_n}(t),$$

$$\hat{v}_t = \sum_{i=1}^n \sum_{j=0}^{m_i-1} v_{i,j} \mathbb{1}_{[\hat{t}_{i,j}, \hat{t}_{i,j+1})}(t) + v_{n,m_n-1} \mathbb{1}_{T_n}(t),$$

where  $t_{1,0} = \hat{t}_{1,0} = 0, t_{i,0} = \hat{t}_{i,0} = T_{i-1}, 2 \leq i \leq n, t_{i,m_i} = \hat{t}_{i,m_i} = T_i, 1 \leq i \leq n$ , we have

$$|G(v) - G(\hat{v})| \leq L \|v\| \sum_{i=1}^n \sum_{j=1}^{m_i} |\Delta t_{i,j} - \Delta \hat{t}_{i,j}|, \tag{3.9}$$

where  $\Delta t_{i,j} := t_{i,j} - t_{i,j-1}$  and  $\Delta \hat{t}_{i,j} := \hat{t}_{i,j} - \hat{t}_{i,j-1}$ .

Note that Assumption 3.14 is close in spirit to Assumption 2.1 in [25], but is weaker and, unlike the latter, is satisfied by European options with intermediate maturities  $T_1, \dots, T_{n-1}$ . Next we introduce a particular class of prediction sets. Our definition is closely related to time-invariant sets in Vovk [51], also recently used in Beiglböck et al. [4], but slightly different as we work with all continuous functions and also require that maturities  $T_i$  are preserved.

**Definition 3.15** We say  $\mathfrak{F}$  is *time-invariant* if  $(S_t)_{t \in [0, T]} \in \mathfrak{F}$  implies that we have  $(S_{f(t)})_{t \in [0, T]} \in \mathfrak{F}$  for any nondecreasing and continuous function  $f : [0, T] \rightarrow [0, T]$  with  $f(0) = 0$  and  $f(T_i) = T_i$  for  $i = 1, \dots, n$ .

We note that many natural path restrictions are time-invariant. Particular examples include  $\{S \in \mathcal{I} : \|S\| \leq b\}$  for a given bound  $b$ , or the set of paths which satisfy a drawdown constraint for a selection of assets  $J \subseteq \{1, \dots, d + K\}$ , i.e.,

$$\left\{ S \in \mathcal{I} : S_t^{(i)} \geq \alpha_i \sup_{u \leq t} S_u^{(i)}, t \in [0, T], i \in J \right\}, \quad \text{for some fixed } \alpha_i \in [0, 1].$$



**Theorem 3.16** *Under Assumption 3.1, let  $\mathfrak{P}$  be a closed and time-invariant subset of  $\mathcal{I}$ ,  $\mathcal{X}$  given by (3.7) and  $\mathcal{P}$  such that  $\mathcal{M}_{\mu, \mathfrak{P}} \neq \emptyset$ , where  $\mu$  is defined via (3.8). Assume there exists  $p > 1$  for which  $\int |x|^p \mu_n^{(i)}(dx) < \infty$  for  $i = 1, \dots, d$ . Then for any  $G$  which satisfies Assumption 3.14, we have*

$$\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = V_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = P_{\mu, \mathfrak{P}}(G) = \tilde{P}_{\mu, \mathfrak{P}}(G).$$

### 4 Auxiliary results and proofs

We present now the proofs of all the results in Sect. 3. As noted before, we use the unconstrained duality (3.4) which follows from Theorem 5.1 stated and proved in Sect. 5 below. We start by describing a discretisation of a continuous path, often referred to as the ‘‘Lebesgue discretisation’’, a term we also use. The discretisation is a crucial tool in Sect. 5.1, but is also employed in the proofs of Lemmas 4.4, 4.3, 4.5 and Theorem 3.16 below.

**Definition 4.1** For a positive integer  $N$  and any  $S \in \Omega$ , we set  $\tau_0^{(N)}(S) = 0$ , then define

$$\tau_k^{(N)}(S) := \inf \left\{ t \geq \tau_{k-1}^{(N)}(S) : |S_t - S_{\tau_{k-1}^{(N)}(S)}| = \frac{1}{2^N} \right\} \wedge T$$

and let  $m^{(N)}(S) := \min\{k \in \mathbb{N} : \tau_k^{(N)}(S) = T\}$ . We write  $m^{(N)}$  for the measurable map  $\Omega \ni S \mapsto m^{(N)}(S)$  and note that by definition,  $m^{(N)} = m^{(N)}(\mathbb{S})$ .

Following the observation that  $m^{(N)}(S) < \infty$  for all  $S \in \Omega$ , we say that the sequence of stopping times  $0 = \tau_0^{(N)} < \tau_1^{(N)} < \dots < \tau_{m^{(N)}}^{(N)} = T$  forms a *Lebesgue partition* of  $[0, T]$  on  $\Omega$ . Similar partitions were studied previously; see e.g. Bichteler [7] and Vovk [51]. Their main appearances have been as tools to build a pathwise version of the Itô integral. They can also be interpreted, from a financial point of view, as candidate times for rebalancing portfolio holdings; see Whalley and Wilmott [52].

*Remark 4.2* Consider  $N \geq 3$  and two paths  $S, \tilde{S} \in \Omega$  such that  $\|S - \tilde{S}\| < 2^{-N}$ . Then for each  $i < m^{(N-2)}(S)$ ,  $\{|S_t| : t \in (\tau_{i-1}^{(N-2)}(S), \tau_i^{(N-2)}(S))\} \cap \{k/2^N : k \in \mathbb{N}_+\}$  has at least four elements, which implies that there exists at least one  $j < m^{(N)}(\tilde{S})$  such that  $\tau_j^{(N)}(\tilde{S}) \in (\tau_{i-1}^{(N-2)}(S), \tau_i^{(N-2)}(S))$ . Consequently,  $m^{(N-2)}(S) \leq m^{(N)}(\tilde{S})$  and hence for any weakly converging sequence of probability measures  $\mathbb{P}^{(k)} \rightarrow \mathbb{P}$  and any bounded nonincreasing function  $\phi : \mathbb{N} \rightarrow \mathbb{R}$ ,

$$\mathbb{E}_{\mathbb{P}}[\phi(m^{(N)}(\mathbb{S}))] \leq \liminf_{k \rightarrow \infty} \mathbb{E}_{\mathbb{P}^{(k)}}[\phi(m^{(N-2)}(\mathbb{S}))]. \tag{4.1}$$

#### 4.1 Proof of Theorem 3.2 and Remark 3.3

To establish (3.2), we consider an  $(X, \gamma) \in \mathcal{A}_{\mathcal{X}}$  that superreplicates  $G$  on  $\mathfrak{P}^\epsilon$  for some  $\epsilon > 0$ . Since  $X$  is bounded and  $\gamma$  is admissible, we can find suitable  $M > 0$

such that

$$X(\mathbb{S}) + \int_0^T \gamma_u \, d\mathbb{S}_u \geq G(\mathbb{S}) - M\lambda_{\mathfrak{P}}(\mathbb{S}), \tag{4.2}$$

where we recall that  $\lambda_{\mathfrak{P}}(\omega) = \inf_{\nu \in \mathfrak{P}} \|\omega - \nu\| \wedge 1$ . Next, for each  $N \geq 1$ , we pick  $\mathbb{P}^{(N)} \in \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^{1/N}$  such that

$$\mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S})] \geq \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^{1/N}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] - \frac{1}{N}.$$

Since  $\gamma$  is progressively measurable, the integral  $\int_0^T \gamma_u(\mathbb{S}) \, d\mathbb{S}_u$ , defined pathwise via integration by parts, agrees a.s. with the stochastic integral under any  $\mathbb{P}^{(N)}$ . Then by (2.1), the stochastic integral is a  $\mathbb{P}^{(N)}$ -supermartingale and hence  $\mathbb{E}_{\mathbb{P}^{(N)}}[\int_0^T \gamma_u(\mathbb{S}) \, d\mathbb{S}_u] \leq 0$ . Therefore, from (4.2),

$$\mathbb{E}_{\mathbb{P}^{(N)}}[X(\mathbb{S})] \geq \mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S}) - M\lambda_{\mathfrak{P}}(\mathbb{S})] \geq \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^{1/N}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] - \frac{1}{N} - \frac{2M}{N}. \tag{4.3}$$

Also note that  $X$  takes the form  $a_0 + \sum_{i=1}^m a_i X_i$ ,  $X_i \in \mathcal{X}$ , and hence by the definition of  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^{1/N}$ ,

$$|\mathcal{P}(X) - \mathbb{E}_{\mathbb{P}^{(N)}}[X(\mathbb{S})]| \longrightarrow 0 \quad \text{as } N \rightarrow \infty.$$

Together with (4.3), this yields  $\mathcal{P}(X) \geq \tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$  and (3.2) follows because  $(X, \gamma) \in \mathcal{A}_{\mathcal{X}}$  was arbitrary.

To establish (3.3), we show the converse inequality in three steps.

*Step 1: Duality without constraints.* This is the crucial and also the most technical part of the proof which we defer to Sect. 5. The duality in (3.4) follows as a special case of Theorem 5.1, which is stated and proved in Sect. 5.

*Step 2: Calculus of variation approach.* Fix  $G$ . Note that any  $(X, \gamma)$  that superreplicates  $G - N\lambda_{\mathfrak{P}}$  on  $\mathcal{I}$  also superreplicates  $G - N/M$  on  $\mathfrak{P}^M$ . It follows that for any fixed  $M, N \geq 1$ ,

$$\begin{aligned} \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) &= \inf\{\mathcal{P}(X) : \exists (X, \gamma) \in \mathcal{A}_{\mathcal{X}}, \epsilon > 0 \text{ such that} \\ &\quad (X, \gamma) \text{ superreplicates } G \text{ on } \mathfrak{P}^\epsilon\} \\ &\leq \frac{N}{M} + \inf\{\mathcal{P}(X) : \exists (X, \gamma) \in \mathcal{A}_{\mathcal{X}} \text{ which} \\ &\quad \text{superreplicates } G - N\lambda_{\mathfrak{P}} \text{ on } \mathcal{I}\}. \end{aligned}$$

Taking the infimum over  $M$  and then over  $N$ , we obtain

$$\begin{aligned} \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) &\leq \inf_{N \geq 0} \inf\{\mathcal{P}(X) : \exists(X, \gamma) \in \mathcal{A}_{\mathcal{X}} \text{ which} \\ &\quad \text{superreplicates } G - N\lambda_{\mathfrak{P}} \text{ on } \mathcal{I}\} \\ &= \inf_{N \geq 0} V_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G - N\lambda_{\mathfrak{P}}) = \inf_{N \geq 0} \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G - N\lambda_{\mathfrak{P}}). \end{aligned} \tag{4.4}$$

On the other hand, given any  $(X, \gamma) \in \mathcal{A}_{\mathcal{X}}$  and  $\epsilon > 0$  such that  $(X, \gamma)$  superreplicates  $G$  on  $\mathfrak{P}^\epsilon$ , by the admissibility of  $(X, \gamma)$  and boundedness of  $X$  and  $G$ , if  $N > 0$  is sufficiently large, then

$$X(S) + \int_0^T \gamma_u(S) dS_u \geq G(S) - N\lambda_{\mathfrak{P}}, \quad S \in \mathcal{I},$$

that is,  $(X, \gamma)$  superreplicates  $G - N\lambda_{\mathfrak{P}}$  on  $\mathcal{I}$ . It follows that we have equality in (4.4). We also have

$$\begin{aligned} \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathcal{I}}(G - N\lambda_{\mathfrak{P}}) &= \inf_{X \in \text{Lin}(\mathcal{X})} (\mathcal{P}(X) + \inf\{x \in \mathbb{R} : \exists \gamma \in \mathcal{A} \text{ such that} \\ &\quad (x, \gamma) \text{ superreplicates} \\ &\quad G - N\lambda_{\mathfrak{P}} - X \text{ on } \mathcal{I}\}) \\ &= \inf_{X \in \text{Lin}(\mathcal{X})} (\mathcal{P}(X) + \mathbf{V}_{\mathcal{I}}(G - N\lambda_{\mathfrak{P}} - X)) \\ &= \inf_{X \in \text{Lin}(\mathcal{X})} (\mathcal{P}(X) + \mathbf{P}_{\mathcal{I}}(G - N\lambda_{\mathfrak{P}} - X)), \end{aligned}$$

where the last equality is justified by Theorem 5.1 as  $\lambda_{\mathfrak{P}}$  and  $X$  are bounded and uniformly continuous. Combining the above with (4.4), we conclude that (3.5) holds.

*Step 3: Application of the minimax theorem.* We rewrite (3.5) and apply a minimax argument to get

$$\begin{aligned} \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) &= \inf_{X \in \text{Lin}(\mathcal{X}), N \geq 0} (\mathbf{P}_{\mathcal{I}}(G - X - N\lambda_{\mathfrak{P}}) + \mathcal{P}(X)) \\ &= \lim_{N \rightarrow \infty} \inf_{X \in \text{Lin}_N(\mathcal{X})} \left( \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X) \right) \\ &= \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{I}}} \inf_{X \in \text{Lin}_N(\mathcal{X})} (\mathbb{E}_{\mathbb{P}}[G - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X)) \tag{4.5} \\ &\leq \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^{\eta_N}} \mathbb{E}_{\mathbb{P}}[G] = \tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G), \tag{4.6} \end{aligned}$$

for  $\eta_N = 2\kappa/\sqrt{N}$  with  $\kappa = 1 + \|G\|_\infty$ , where  $\|G\|_\infty = \sup_{S \in \Omega} |G(S)|$ . The crucial third equality follows by a minimax theorem (see e.g. Terkelsen [50, Corollary 2]) by

observing that the mapping

$$\text{Lin}_N(\mathcal{X}) \times \mathcal{M}_T \ni (X, \mathbb{P}) \mapsto \mathbb{E}_{\mathbb{P}}[G(\mathbb{S}) - X(\mathbb{S}) - N\lambda_{\mathfrak{P}}(\mathbb{S})] + \mathcal{P}(X) \in \mathbb{R}$$

is bilinear and  $\text{Lin}_N(\mathcal{X})$  is convex and compact. To justify the inequality between (4.5) and (4.6), consider  $\mathbb{P} \in \mathcal{M}_T \setminus \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^{\eta_N}$ . Then in particular, either there exists  $X^* \in \mathcal{X}$  such that  $|\mathbb{E}_{\mathbb{P}}[X^*] - \mathcal{P}(X^*)| > \eta_N \frac{1}{\sqrt{N}}$  or  $\mathbb{P}[\mathbb{S} \notin \mathfrak{P}^{\eta_N}] \geq \eta_N$ . In the former case, since  $\pm NX^* \in \text{Lin}_N(\mathcal{X})$ , we obtain

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[G - NX^* - N\lambda_{\mathfrak{P}}] + \mathcal{P}(NX^*) &\leq \mathbb{E}_{\mathbb{P}}[G] - N(\mathbb{E}_{\mathbb{P}}[X^*] - \mathcal{P}(X^*)) \\ &< \kappa - 2\kappa\sqrt{N} \leq -\kappa, \end{aligned}$$

where, without loss of generality, we assume  $\mathbb{E}_{\mathbb{P}}[X^*] < \mathcal{P}(X^*)$ . In the latter case, we have  $\mathbb{E}_{\mathbb{P}}[N\lambda_{\mathfrak{P}}] \geq N \frac{2\kappa}{\sqrt{N}} \frac{2\kappa}{\sqrt{N}} = 4\kappa^2 \geq 4\kappa$ , while  $|\mathbb{E}_{\mathbb{P}}[X] - \mathcal{P}(X)| \leq N \frac{2\kappa}{N} = 2\kappa$  for any  $X \in \text{Lin}_N(\mathcal{X})$ . It follows that

$$\mathbb{E}_{\mathbb{P}}[G - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X) \leq \kappa - 4\kappa + 2\kappa = -\kappa.$$

On the other hand, since (3.2) implies  $\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(0) = 0$ , we have

$$\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G + \|G\|_{\infty}) - \|G\|_{\infty} \geq \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(0) - \|G\|_{\infty} = -\kappa + 1,$$

and hence we may restrict to measures in  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^{\eta_N}$  in (4.5). Dropping nonpositive terms, we obtain (4.6) which completes the proof of Theorem 3.2.  $\square$

For Remark 3.3, it remains to argue that Theorem 3.2 remains true when we restrict to Brownian martingales. Specifically, given  $T$  and a probability space  $(\Omega^W, \mathbb{F}^W, P^W)$  with a  $\tilde{d}$ -dimensional Brownian motion  $W$  on  $[0, T]$ , where  $\mathbb{F}^W = (\mathcal{F}_t^W)_{0 \leq t \leq T}$  is the  $P^W$ -completion of the natural filtration of  $W$ , consider

$$\mathbb{P} := P^W \circ (Z^\alpha)^{-1}, \quad \text{where } Z^\alpha := \int_0^\cdot \alpha_u \, dW_u \tag{4.7}$$

for some  $\mathbb{F}^W$ -progressively measurable process  $\alpha$  with values in the  $(d + K) \times \tilde{d}$  matrices such that the above vector integral is well defined. Let  $\underline{\mathcal{M}}_T$  be the family of all  $\mathbb{P} \in \mathcal{M}_T$  which admit such a representation. From (3.5), as argued above, and Remark 5.2 below, we have

$$\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = \inf_{X \in \text{Lin}(\mathcal{X}), N \geq 0} \left( \sup_{\mathbb{P} \in \underline{\mathcal{M}}_T} \mathbb{E}_{\mathbb{P}}[G - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X) \right).$$

Then by following the same argument as in Step 3 above, we can show that we have  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^{\eta_N} \cap \underline{\mathcal{M}}_T \neq \emptyset$  when  $N$  is sufficiently large and

$$\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^{\eta_N} \cap \underline{\mathcal{M}}_T} \mathbb{E}_{\mathbb{P}}[G].$$

### 4.2 Proof of Theorem 3.6

The set  $\mathcal{X}$  is finite and as discussed in Example 3.4, we can apply Theorem 3.2. Together with  $V_{\mathcal{X},\mathcal{P},\mathcal{I}} = \tilde{V}_{\mathcal{X},\mathcal{P},\mathcal{I}}$ , this yields

$$V_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = \tilde{P}_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) = \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}}^{1/N}} \mathbb{E}_{\mathbb{P}}[G]. \tag{4.8}$$

Now for every positive integer  $N$ , we pick  $\mathbb{P}^{(N)} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}}^{1/N}$  such that

$$\mathbb{E}_{\mathbb{P}^{(N)}}[G] + 1/N \geq \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}}^{1/N}} \mathbb{E}_{\mathbb{P}}[G].$$

We let

$$p_{k,i,j}^{(N)} := \mathbb{E}_{\mathbb{P}^{(N)}}[(K_{k,j}^{(i)} - S_{k,j}^{(i)})^+], \quad \tilde{p}_{k,i,j}^{(N)} := \sqrt{N}(p_{k,i,j} - (1 - 1/\sqrt{N})p_{k,i,j}^{(N)}),$$

for any  $i = 1, \dots, d, j = 1, \dots, n, k = 1, \dots, m(i, j)$ . Note that

$$|\tilde{p}_{k,i,j}^{(N)} - p_{k,i,j}| = (\sqrt{N} - 1)|p_{k,i,j} - p_{k,i,j}^{(N)}| \leq \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}}, \quad \forall i, j, k.$$

Then it follows from Assumption 3.5 that when  $N$  is large, there exists a  $\tilde{\mathbb{P}}^{(N)} \in \mathcal{M}_{\mathcal{I}}$  such that

$$\tilde{p}_{k,i,j}^{(N)} = \mathbb{E}_{\tilde{\mathbb{P}}^{(N)}}[(K_{k,j}^{(i)} - S_{k,j}^{(i)})^+], \quad \forall i, j, k.$$

Now we consider  $\mathbb{Q} := (1 - 1/\sqrt{N})\mathbb{P}^{(N)} + \tilde{\mathbb{P}}^{(N)}/\sqrt{N}$ . It follows that

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}[(K_{k,j}^{(i)} - S_{k,j}^{(i)})^+] &= (1 - 1/\sqrt{N})\mathbb{E}_{\mathbb{P}^{(N)}}[(K_{k,j}^{(i)} - S_{k,j}^{(i)})^+] \\ &\quad + \frac{1}{\sqrt{N}}\mathbb{E}_{\tilde{\mathbb{P}}^{(N)}}[(K_{k,j}^{(i)} - S_{k,j}^{(i)})^+] \\ &= (1 - 1/\sqrt{N})p_{k,i,j}^{(N)} + \tilde{p}_{k,i,j}^{(N)}/\sqrt{N} = p_{k,i,j} \end{aligned}$$

and hence  $\mathbb{Q} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}}$ . In addition,

$$|\mathbb{E}_{\mathbb{Q}}[G] - \mathbb{E}_{\mathbb{P}^{(N)}}[G]| \leq \frac{1}{\sqrt{N}}(\mathbb{E}_{\mathbb{P}^{(N)}}[|G|] + \mathbb{E}_{\tilde{\mathbb{P}}^{(N)}}[|G|]) \leq \frac{2\|G\|_{\infty}}{\sqrt{N}}.$$

Therefore, we have

$$\begin{aligned} \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}}^{1/N}} \mathbb{E}_{\mathbb{P}}[G] &\leq \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X},\mathcal{P},\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G] + \frac{2\|G\|_{\infty}}{\sqrt{N}} + \frac{1}{N} \\ &= P_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) + \frac{2\|G\|_{\infty}}{\sqrt{N}} + \frac{1}{N}, \end{aligned}$$

and taking limits as  $N \rightarrow \infty$  yields  $\tilde{P}_{\mathcal{X},\mathcal{P},\mathcal{I}}(G) \leq P_{\mathcal{X},\mathcal{P},\mathcal{I}}(G)$ . Together with (4.8) and (3.1), this completes the proof.  $\square$

### 4.3 Proof of Theorem 3.12

From Theorem 3.2, we know that  $\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) \geq \tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G)$  and by definition,  $\tilde{P}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) \geq \tilde{P}_{\mu, \mathfrak{P}}(G)$ . Hence, to establish Theorem 3.12, it suffices to show that

$$\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) \leq \tilde{P}_{\mu, \mathfrak{P}}(G).$$

This is a special case ( $\alpha = \beta = 0$ ) of Proposition 4.5 below, which is a crucial technical result also used to prove Theorem 3.16 below. We recall that  $\underline{\mathcal{M}}_{\mathcal{I}} = \mathcal{M}_{\mathcal{I}} \cap \underline{\mathcal{M}}$ , where  $\underline{\mathcal{M}} := \{\mathbb{P} \in \mathcal{M} : \mathbb{P} \text{ satisfies (4.7)}\}$ . We also require additional notation for sets of martingale measures with a constraint on the final marginal only. For a probability measure  $\pi$  on  $\mathbb{R}^d$ , we let  $\mathcal{M}_{\pi, \mathcal{I}} := \{\mathbb{P} \in \mathcal{M}_{\mathcal{I}} : \mathcal{L}_{\mathbb{P}}(S_{T_n}) = \pi\}$  and likewise, for a  $d$ -tuple  $\mu_n = (\mu_n^{(1)}, \dots, \mu_n^{(d)})$  of probability measures on  $\mathbb{R}$ , we let  $\mathcal{M}_{\mu_n, \mathcal{I}} := \{\mathbb{P} \in \mathcal{M}_{\mathcal{I}} : \mathcal{L}_{\mathbb{P}}(S_{T_n}^{(i)}) = \mu_n^{(i)}, i = 1, \dots, d\}$ . These notations are used in statements and proofs below, and it will be clear from the context if we work with the former or the latter object. Finally, we allow a perturbation by defining

$$\mathcal{M}_{\mu_n, \mathcal{I}, \eta} := \{\mathbb{P} \in \mathcal{M}_{\mathcal{I}} : d_{\mathbb{P}}(\mathcal{L}_{\mathbb{P}}(S_{T_n}^{(i)}, \mu_n^{(i)}) \leq \eta, i = 1, \dots, d\}.$$

We note that these sets are different from the main objects introduced in Definition 3.11 and are only needed for some technical arguments below.

We start with two lemmas leading to Proposition 4.5.

**Lemma 4.3** *Consider probability measures  $\pi^{(N)}, \pi$  on  $\mathbb{R}_+^d$  with mean vectors  $\mathbf{1}$  and  $(\pi^{(N)})$  converging weakly to  $\pi$ . Then, for any  $\alpha, \beta \geq 0, D \in \mathbb{N}$  and a bounded uniformly continuous  $G$ ,*

$$\limsup_{N \rightarrow \infty} \sup_{\mathbb{P} \in \underline{\mathcal{M}} \cap \mathcal{M}_{\pi^{(N)}, \mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G - \beta \sqrt{m^{(D)}} \wedge \alpha] \leq \sup_{\mathbb{P} \in \underline{\mathcal{M}} \cap \mathcal{M}_{\pi, \mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G - \beta \sqrt{m^{(D-2)}} \wedge \alpha],$$

where  $m^{(D)}$  is given in Definition 4.1.

*Proof* See Sect. A.1 in the Appendix. □

**Lemma 4.4** *Under Assumption 3.1, let  $\mathfrak{P}$  be a measurable subset of  $\mathcal{I}$ ,  $\text{Lin}_1(\mathcal{X})$  a compact subset of  $(\mathcal{C}(\Omega, \mathbb{R}), \|\cdot\|_{\infty})$  and  $\mathcal{M}_s$  a nonempty convex subset of  $\mathcal{M}_{\mathcal{I}}$  such that  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^{\eta} \cap \mathcal{M}_s \neq \emptyset$  for all  $\eta > 0$ . Then for any  $\alpha, \beta \geq 0, D \in \mathbb{N}$  and a bounded uniformly continuous  $G$ ,*

$$\begin{aligned} & \inf_{X \in \text{Lin}(\mathcal{X}), N \geq 0} \left( \sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \beta \sqrt{m^{(D)}} \wedge \alpha - X - N\lambda_{\mathfrak{P}}] + \mathcal{P}(X) \right) \\ & \leq \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}^{1/N} \cap \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \beta \sqrt{m^{(D-2)}} \wedge \alpha], \end{aligned} \tag{4.9}$$

with equality when  $\alpha = \beta = 0$ , where  $m^{(D)}$  is defined in Definition 4.1.

*Proof* See Sect. A.2 in the Appendix. □

**Proposition 4.5** *Under Assumption 3.1, let  $\mathfrak{X}$  be a measurable subset of  $\mathcal{I}$ ,  $\mathcal{X}$  given by (3.7) and  $\mathcal{P}$  such that for any  $\eta > 0$ ,  $\mathcal{M}_{\mu, \mathfrak{X}}^\eta \neq \emptyset$ , where  $\mu$  is defined via (3.8). Then for any uniformly continuous and bounded  $G$  and  $\alpha, \beta \geq 0, D \in \mathbb{N}$ ,*

$$\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{X}}(G - \beta\sqrt{m^{(D)}} \wedge \alpha) \leq \tilde{P}_{\mu, \mathfrak{X}}(G - \beta\sqrt{m^{(D-8)}} \wedge \alpha),$$

where  $m^{(D)}$  is defined in Definition 4.1.

*Proof* Define

$$\mathfrak{G}_M(\mathbb{R}_+^d) := \left\{ f \in \mathcal{C}(\mathbb{R}_+^d, \mathbb{R}) : \sup_{\mathbf{x} \neq \mathbf{y}} \frac{|f(\mathbf{x}) - f(\mathbf{y})|}{|\mathbf{x} - \mathbf{y}|} \leq M, \|f\|_\infty \leq M \text{ and} \right. \\ \left. f(x_1, \dots, x_d) = f(x_1 \wedge M^2, \dots, x_d \wedge M^2) \right. \\ \left. \text{for all } (x_1, \dots, x_d) \in \mathbb{R}_+^d \right\}$$

and  $\mathfrak{G}(\mathbb{R}_+^d) = \bigcup_{M>0} \mathfrak{G}_M(\mathbb{R}_+^d)$ . Let  $\mathcal{Z}_M = \{f(\mathbb{S}_{T_n}^{(i)}) : f \in \mathfrak{G}_M(\mathbb{R}_+), i = 1, \dots, d\}$ ,  $\mathcal{Y}_M = \{f(\mathbb{S}_{T_j}^{(i)}) : f \in \mathfrak{G}_M(\mathbb{R}_+), i = 1, \dots, d, j = 1, \dots, n - 1\}$  and write

$$\mathcal{Z} = \bigcup_{M \geq 0} \mathcal{Z}_M, \quad \mathcal{Y} = \bigcup_{M \geq 0} \mathcal{Y}_M.$$

Note that given any  $f \in \mathcal{C}_b(\mathbb{R}_+, \mathbb{R})$ ,  $\epsilon > 0$  and a measure  $\mu$  on  $\mathbb{R}_+$  with a finite first moment, there is some  $u : \mathbb{R}_+ \rightarrow \mathbb{R}$  of the form  $u(s) = a_0 + \sum_{i=1}^n a_i(\kappa_i - s)^+$  such that  $u \geq f$  and  $\int(u - f) d\mu < \epsilon$ . This gives the first inequality in the following:

$$\begin{aligned} &\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{X}}(G - \beta\sqrt{m^{(D)}} \wedge \alpha) \\ &\leq \tilde{V}_{\mathcal{Z} \cup \mathcal{Y}, \mathcal{P}, \mathfrak{X}}(G - \beta\sqrt{m^{(D)}} \wedge \alpha) \\ &= \inf_{X \in \text{Lin}(\mathcal{Z} \cup \mathcal{Y}), N \geq 0} (\mathbf{V}_{\mathcal{I}}(G - X - \beta\sqrt{m^{(D)}} \wedge \alpha - N\lambda_{\mathfrak{X}}) + \mathcal{P}(X)) \\ &\leq \inf_{X \in \text{Lin}(\mathcal{Z} \cup \mathcal{Y}), N \geq 0} \left( \sup_{\mathbb{P} \in \underline{\mathcal{M}}_{\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G - X - \beta\sqrt{m^{(D-2)}} \wedge \alpha - N\lambda_{\mathfrak{X}}] + \mathcal{P}(X) \right) \\ &= \inf_{Y \in \text{Lin}(\mathcal{Y}), N \geq 0, M \geq 0} \inf_{Z \in \text{Lin}(\mathcal{Z}_M)} \left( \sup_{\mathbb{P} \in \underline{\mathcal{M}}_{\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G - Y - Z - \beta\sqrt{m^{(D-2)}} \wedge \alpha - N\lambda_{\mathfrak{X}}] \right. \\ &\quad \left. + \mathcal{P}(Y + Z) \right) \\ &\leq \inf_{Y \in \text{Lin}(\mathcal{Y}), M \geq 0, N \geq 0} \lim_{L \rightarrow \infty} \sup_{\mathbb{P} \in \underline{\mathcal{M}}_{\mathcal{I}} \cap \mathcal{M}_{\mathcal{Z}_M, \mathcal{P}, \mathcal{I}}^{1/L}} \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D-4)}} \wedge \alpha - Y \\ &\quad - N\lambda_{\mathfrak{X}} + \mathcal{P}(Y)] \\ &\leq \inf_{Y \in \text{Lin}(\mathcal{Y}), M \geq 0, N \geq 0} \sup_{\mathbb{P} \in \underline{\mathcal{M}}_{\mathcal{I}} \cap \mathcal{M}_{\mathcal{Z}_M, \mathcal{P}, \mathcal{I}}^{1/M}} \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D-4)}} \wedge \alpha - Y \\ &\quad - N\lambda_{\mathfrak{X}} + \mathcal{P}(Y)]. \end{aligned} \tag{4.10}$$

Above, the first equality follows from the previously argued equality in (4.4). The second inequality is justified by Theorem 5.1 and Remark 5.2 below. Both the ensuing equality and the last inequality are clear. It remains to observe that the third inequality follows from Lemma 4.4. To justify this, note that  $\mathfrak{G}_M(\mathbb{R}_+)$  is a convex and compact subset of  $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ , and it follows that  $\text{Lin}_1(\mathcal{Z}_M) = \mathcal{Z}_M$  is a convex compact subset of  $(\mathcal{C}(\Omega, \mathbb{R}), \|\cdot\|_\infty)$ . In addition,  $\mathcal{M}_{\mu, \mathfrak{F}}^\eta \neq \emptyset$  for all  $\eta > 0$  implies that  $\mathcal{M}_{\mathcal{Z}_M, \mathcal{P}, \mathcal{I}}^{1/M} \neq \emptyset$  for all  $M$ , and clearly we can obtain such measures on a Wiener space so that  $\mathcal{M}_{\mathcal{Z}_M, \mathcal{P}, \mathcal{I}}^{1/M} \cap \underline{\mathcal{M}}_{\mathcal{I}} \neq \emptyset$  for all  $M$ , as required.

For any  $\mathbb{P} \in \mathcal{M}_{\mathcal{Z}_M, \mathcal{P}, \mathcal{I}}^{1/M}$ , let  $\epsilon_{\mathbb{P}} = \max\{d_p(\mu_n^{(i)}, \mathcal{L}_{\mathbb{P}}(\mathbb{S}_{T_n}^{(i)})) : i = 1, \dots, d\}$ , where the Lévy–Prokhorov metric  $d_p$  on probability measures on  $\mathbb{R}_+^d$  is given by

$$d_p(\mu, \nu) := \sup_{f \in \mathfrak{G}_1^b(\mathbb{R}_+^d)} \left| \int f \, d\nu - \int f \, d\mu \right|,$$

where

$$\mathfrak{G}_1^b(\mathbb{R}_+^d) := \{f \in \mathcal{C}(\mathbb{R}_+^d, \mathbb{R}) : \|f\| \leq 1 \text{ and } |f(\mathbf{x}) - f(\mathbf{y})| \leq |\mathbf{x} - \mathbf{y}|, \forall \mathbf{x} \neq \mathbf{y}\}$$

(see e.g. Bogachev [9, Theorem 8.3.2]). Pick  $g \in \mathfrak{G}_1^b(\mathbb{R}_+)$  such that

$$\left| \int_{\mathbb{R}_+} g(x) \mu_n^{(i)}(dx) - \mathbb{E}_{\mathbb{P}}[g(\mathbb{S}_{T_n}^{(i)})] \right| > \epsilon_{\mathbb{P}}/2 \quad \text{for some } i = 1, \dots, d,$$

and define  $\hat{g} \in \mathfrak{G}_M(\mathbb{R}_+)$  via  $\hat{g}(x) = Mg(x \wedge M^2)$ . Then by the definition of  $\mathcal{M}_{\mathcal{Z}_M, \mathcal{P}, \mathcal{I}}^{1/M}$  and  $\hat{g}$ ,

$$\begin{aligned} \frac{1}{M} &\geq \left| \int_{\mathbb{R}_+} \hat{g}(x) \mu_n^{(i)}(dx) - \mathbb{E}_{\mathbb{P}}[\hat{g}(\mathbb{S}_{T_n}^{(i)})] \right| \\ &\geq M \left| \int g \, d\mu_n^{(i)} - \mathbb{E}_{\mathbb{P}}[g(\mathbb{S}_{T_n}^{(i)})] \right| - M \mu_n^{(i)}(\{|x| \geq M^2\}) - M \mathbb{P}[|\mathbb{S}_{T_n}^{(i)}| \geq M^2] \\ &\geq M \frac{\epsilon_{\mathbb{P}}}{2} - \frac{2}{M}. \end{aligned}$$

It follows that  $\epsilon_{\mathbb{P}} \leq 3/M^2$  and hence  $\mathcal{M}_{\mathcal{Z}_M, \mathcal{P}, \mathcal{I}}^{1/M} \subseteq \mathcal{M}_{\mu_n, \mathcal{I}, 1/M}$  for  $M \geq 3$ . Fix  $Y \in \text{Lin}(\mathcal{Y})$  and for each  $M \geq 3$ , take  $\mathbb{P}^{(M)} \in \underline{\mathcal{M}} \cap \mathcal{M}_{\mu_n, \mathcal{I}, 1/M}$  such that

$$\begin{aligned} &\mathbb{E}_{\mathbb{P}^{(M)}}[G - \beta\sqrt{m^{(D-4)}} \wedge \alpha - Y - N\lambda_{\mathfrak{F}} + \mathcal{P}(Y)] \\ &\geq \sup_{\mathbb{P} \in \underline{\mathcal{M}} \cap \mathcal{M}_{\mu_n, \mathcal{I}, 1/M}} \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D-4)}} \wedge \alpha - Y - N\lambda_{\mathfrak{F}} + \mathcal{P}(Y)] - \frac{1}{M}. \end{aligned}$$

Let  $\pi_n^{(M)}$  be the law of  $(\mathbb{S}_{T_n}^{(1)}, \dots, \mathbb{S}_{T_n}^{(d)})$  under  $\mathbb{P}^{(M)}$  and note that its marginals have mean 1. The family  $(\pi_n^{(M)})_{M \geq 3}$  is tight and by Prokhorov’s theorem, there exists a



subsequence  $(\pi_n^{(M_k)})_{k \in \mathbb{N}}$  converging to some  $\pi_n$ . Note that the marginal distributions of  $\pi_n$  are  $\mu_n^{(i)}$ ,  $i = 1, \dots, d$ . By the choice of  $\mathbb{P}^{(M)}$  and Lemma 4.3, it follows that

$$\begin{aligned} & \lim_{M \rightarrow \infty} \sup_{\mathbb{P} \in \underline{\mathcal{M}} \cap \mathcal{M}_{\mu_n, \mathcal{I}, 1/M}} \mathbb{E}_{\mathbb{P}}[G - \beta \sqrt{m^{(D-4)}} \wedge \alpha - Y - N\lambda_{\mathfrak{F}} + \mathcal{P}(Y)] \\ & \leq \sup_{\mathbb{P} \in \underline{\mathcal{M}} \cap \mathcal{M}_{\mu_n, \mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G - \beta \sqrt{m^{(D-6)}} \wedge \alpha - Y - N\lambda_{\mathfrak{F}} + \mathcal{P}(Y)]. \end{aligned}$$

With this and using  $\mathcal{M}_{\mathcal{Z}_M, \mathcal{P}, \mathcal{I}}^{1/M} \subseteq \mathcal{M}_{\mu_n, \mathcal{I}, 1/M}$ , we may continue (4.10) by writing

$$\begin{aligned} & \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}(G) \\ & \leq \inf_{Y \in \text{Lin}(\mathcal{Y})} \inf_{M \geq 0, N \geq 0} \sup_{\mathbb{P} \in \underline{\mathcal{M}} \cap \mathcal{M}_{\mathcal{Z}_M, \mathcal{P}, \mathcal{I}}^{1/M}} \mathbb{E}_{\mathbb{P}}[G - \beta \sqrt{m^{(D-4)}} \wedge \alpha - Y \\ & \quad - N\lambda_{\mathfrak{F}} + \mathcal{P}(Y)] \\ & \leq \inf_{Y \in \text{Lin}(\mathcal{Y}), N \geq 0} \sup_{\mathbb{P} \in \mathcal{M}_{\mu_n, \mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G - \beta \sqrt{m^{(D-6)}} \wedge \alpha - Y - N\lambda_{\mathfrak{F}} + \mathcal{P}(Y)] \\ & \leq \inf_{M \geq 0} \inf_{Y \in \text{Lin}(\mathcal{Y}_M), N \geq 0} \sup_{\mathbb{P} \in \mathcal{M}_{\mu_n, \mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G - \beta \sqrt{m^{(D-6)}} \wedge \alpha - Y - N\lambda_{\mathfrak{F}} + \mathcal{P}(Y)] \\ & \leq \inf_{M \geq 0} \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\mu_n, \mathcal{I}} \cap \mathcal{M}_{\mathcal{Y}_M, \mathcal{P}, \mathfrak{F}}^{1/N}} \mathbb{E}_{\mathbb{P}}[G - \beta \sqrt{m^{(D-8)}} \wedge \alpha], \end{aligned}$$

where the last inequality follows from Lemma 4.4 since by analogous arguments to the ones above, we may argue that  $\mathcal{M}_{\mu, \mathfrak{F}}^{1/M^2} \subseteq \mathcal{M}_{\mathcal{Y}_M, \mathcal{P}, \mathfrak{F}}^{1/M} \cap \mathcal{M}_{\mu_n, \mathcal{I}} \subseteq \mathcal{M}_{\mu, \mathfrak{F}}^{1/M}$  when  $M$  is large enough. The first inclusion implies that  $\mathcal{M}_{\mathcal{Y}_M, \mathcal{P}, \mathfrak{F}}^{1/M} \cap \mathcal{M}_{\mu_n, \mathcal{I}} \neq \emptyset$ , justifying the application of Lemma 4.4. The second inclusion allows us to continue the above chain of inequalities to conclude the proof via

$$\begin{aligned} \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}(G) & \leq \lim_{M \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\mu, \mathfrak{F}}^{1/M}} \mathbb{E}_{\mathbb{P}}[G - \beta \sqrt{m^{(D-8)}} \wedge \alpha] \\ & = \tilde{P}_{\mu, \mathfrak{F}}(G - \beta \sqrt{m^{(D-8)}} \wedge \alpha). \quad \square \end{aligned}$$

### 4.4 Proof of Theorem 3.16

We first make two simple observations.

*Remark 4.6* If  $\mathfrak{F}$  is a nonempty closed (with respect to the sup-norm) subset of  $\Omega$ , then

$$\mathfrak{F} = \bigcap_{\epsilon > 0} \mathfrak{F}^\epsilon = \bigcap_{\epsilon > 0} \overline{\mathfrak{F}^\epsilon},$$

where  $\overline{\mathfrak{F}^\epsilon}$  is the closure of  $\mathfrak{F}^\epsilon$ .

**Lemma 4.7** *If  $\mathfrak{P}$  is time-invariant, then for every  $\epsilon > 0$ ,  $\mathfrak{P}^\epsilon$  is also time-invariant.*

*Proof* This follows easily by observing that for two paths  $S, \tilde{S} \in \Omega$  and any nondecreasing continuous function  $f : [0, T_n] \rightarrow [0, T_n]$  with  $f(0) = 0$  and  $f(T_i) = T_i$  for any  $i = 1, \dots, n$ , we have  $\|\tilde{S} - S\| = \|\tilde{S}_{f(\cdot)} - S_{f(\cdot)}\|$ .  $\square$

We now proceed with the proof of Theorem 3.16. Recall that the inequalities  $\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) \geq V_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) \geq P_{\mu, \mathfrak{P}}(G)$  hold in general. In addition, according to Theorem 3.12,  $\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) = P_{\mu, \mathfrak{P}}(G)$ . Therefore, we only need to show that  $\tilde{P}_{\mu, \mathfrak{P}}(G) = P_{\mu, \mathfrak{P}}(G)$ . Our proof of this equality is divided into six steps. First, using Proposition 4.5, we argue that it suffices to consider measures with “good control” on the expectation of  $m^{(D)}(\mathbb{S})$ . Next, we perform three time changes within each trading period  $[T_i, T_{i+1}]$ . The resulting time change of  $\mathbb{S}$ , denoted by  $\tilde{\mathbb{S}}$ , allows a “good control” over its quadratic variation process. At the same time, we keep  $G(\mathbb{S})$  and  $G(\tilde{\mathbb{S}})$  “close”, and given a measure  $\mathbb{P} \in \mathcal{M}_{\mu, \mathfrak{P}}^\eta$  with “good control” on  $\mathbb{E}_{\mathbb{P}}[m^{(D)}(\mathbb{S})]$ , since  $\mathfrak{P}^\eta$  is time-invariant, the law of the time-changed price process  $\tilde{\mathbb{S}}$  remains an element of  $\mathcal{M}_{\mu, \mathfrak{P}}^\eta$ . Then in Step 5, given a sequence of models with improved calibration precisions, we show tightness of the quadratic variation process of the time-changed price process  $\tilde{\mathbb{S}}$  under these measures. This then leads to tightness of the image measures via  $\tilde{\mathbb{S}}$ . In Step 6, we deduce the duality  $\tilde{P}_{\mu, \mathfrak{P}}(G) = P_{\mu, \mathfrak{P}}(G)$  from tightness and conclude.

Recall that  $\mathcal{X}$  is given by (3.7). Let

$$\mathcal{X}_n = \{(\kappa - \mathbb{S}_{T_n}^{(i)})^+ : i = 1, \dots, d, \kappa \in \mathbb{R}_+\}$$

and write  $P_{\mu_n, \mathfrak{P}} := P_{\mathcal{X}_n, \mathcal{P}, \mathfrak{P}}$  for the associated primal problem, where the martingale measures have fixed marginals  $\mu_n^{(i)}$ , given by (3.8), of the distribution of  $\mathbb{S}_{T_n}$ . Note that by definition,  $P_{\mu_n, \mathcal{I}} = \tilde{P}_{\mu_n, \mathcal{I}}$  and that since the  $\mu_n^{(i)}$  have finite  $p$ th moment, we have  $P_{\mu_n, \mathcal{I}}(\|\mathbb{S}\|) < \infty$ .

*Step 1: Reducing to measures  $\mathbb{P}$  with good control on  $\mathbb{E}_{\mathbb{P}}[\sqrt{m^{(D)}(\mathbb{S})}]$ .* Let  $G$  satisfy Assumption 3.14. Choose  $\kappa \geq 1$  such that  $\|G\| \leq \kappa$  and let  $f_e : \mathbb{R}_+^{d+K} \rightarrow \mathbb{R}_+$  be a modulus of continuity of  $G$ , i.e.,

$$|G(\omega) - G(\nu)| \leq f_e(|\omega - \nu|) \quad \text{for any } \omega, \nu \in \Omega$$

with  $\lim_{x \rightarrow 0} f_e(x) = 0$ . Fix  $D \in \mathbb{N}$ . Consider  $X_D : \Omega \rightarrow \mathbb{R}$  given by

$$\begin{aligned} X_D(S) &= \sqrt{m^{(D)}(S) \wedge 2^{6D} \kappa^2} \sum_{j=1}^{d+K} |S_{\tau_j^{(D)}(S)}^{(i)} - S_{\tau_{j-1}^{(D)}(S)}^{(i)}|^2 \\ &\geq 2^{-D} \sqrt{m^{(D)}(S) \wedge (2^{6D} \kappa^2)} - 1 \\ &\geq \left(2^{-D} (\sqrt{m^{(D)}(S) \wedge 2^{6D} \kappa^2} - 1)\right) = \kappa^{2D} \wedge \frac{\sqrt{m^{(D)}(S)}}{2^D} - 2^{-D}, \end{aligned}$$

where the  $\tau_i^{(D)}$  and  $m^{(D)}$  are defined in Definition 4.1. It follows from the proof of Lemma 5.4 in Dolinsky and Soner [26] that there exists a  $\gamma \in \mathcal{A}$  such that

$$\int_0^{\tau_{m^{(D)}(S) \wedge 2^{6D} \kappa^2}} \gamma_u \, dS_u + 3(d + K) \max_{0 \leq j \leq (m^{(D)}(S) \wedge 2^{6D} \kappa^2)} |S_{\tau_j^{(D)}}| \geq \tilde{X}_D(S), \quad S \in \mathcal{I}.$$

Hence  $\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(X_D) \leq 3(d + K) \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(\|\mathbb{S}\| \wedge (\kappa^2 2^{5D} + 1))$ . Reducing  $\mathcal{X}$  to options with maturity  $T_n$  and considering  $\mathcal{I}$  instead of  $\mathfrak{P}$  only increases the superhedging price, and therefore

$$0 \leq \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(X_D) \leq 3(d + K) V_{\mathcal{X}_n, \mathcal{P}, \mathcal{I}}(\|\mathbb{S}\| \wedge (\kappa^2 2^{5D} + 1)) \leq 3(d + K) P_{\mu_n, \mathcal{I}}(\|\mathbb{S}\|),$$

which is finite, and where the last inequality follows from Theorem 3.12 applied to the case of a single maturity. It now follows from sublinearity of  $\tilde{V}$  that

$$\begin{aligned} \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) &\leq \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}\left(G - \kappa 2^D \wedge \frac{X_D}{2^D}\right) + \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(X_D/2^D) \\ &\leq \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}\left(G - \kappa 2^D \wedge \frac{\sqrt{m^{(D)}}}{2^{2D}}\right) + c_2/2^D \\ &\leq \tilde{P}_{\mu, \mathfrak{P}}\left(G - \kappa 2^D \wedge \frac{\sqrt{m^{(D-8)}}}{2^{2D}}\right) + c_2/2^D \\ &= \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\mu, \mathfrak{P}}^{1/N}} \mathbb{E}_{\mathbb{P}}\left[G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D-8)}(\mathbb{S})}}{2^{2D}}\right] + c_2/2^D, \end{aligned} \tag{4.11}$$

where  $c_2$  is a constant independent of  $D$  and the last inequality follows from Proposition 4.5.

Next we denote by  $\widehat{\mathcal{M}}_{\mathcal{I}}^{\kappa}$  the set of  $\mathbb{P} \in \mathcal{M}_{\mathcal{I}}$  such that

$$\mathbb{E}_{\mathbb{P}}\left[\kappa 2^D \wedge \frac{\sqrt{m^{(D-8)}(\mathbb{S})}}{2^{2D}}\right] \leq 2\kappa + 2. \tag{4.12}$$

We notice that if  $\mathbb{P} \notin \widehat{\mathcal{M}}_{\mathcal{I}}^{\kappa}$ , then

$$\mathbb{E}_{\mathbb{P}}\left[G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D-8)}(\mathbb{S})}}{2^{2D}}\right] < \kappa - 2\kappa - 2 = -\kappa - 2,$$

while by the inequalities in (4.11) above, for  $D$  sufficiently large,

$$\sup_{\mathbb{P} \in \mathcal{M}_{\mu, \mathfrak{P}}^{1/N}} \mathbb{E}_{\mathbb{P}}\left[G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D-8)}(\mathbb{S})}}{2^{2D}}\right] \geq \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{P}}(G) - \frac{c_2}{2^D} \geq -\kappa - 1.$$

It follows that in (4.11), it suffices to consider  $\mathbb{P} \in \widehat{\mathcal{M}}_{\mathcal{I}}^{\kappa} \cap \mathcal{M}_{\mu, \mathfrak{P}}^{1/N}$ , which in particular is nonempty.

*Step 2: First time change: “squeezing paths and adding constant paths”.* The first time change squeezes the evolution on  $[T_{i-1}, T_i]$  to  $[T_{i-1}, T_i - 1/D]$  and adds a constant piece to the path on  $[T_i - 1/D, T_i]$ . To achieve this, define an increasing function  $f : [0, T_n] \rightarrow [0, T_n]$  by

$$f(t) = \sum_{i=1}^n \left( T_i \wedge \left( T_{i-1} + \frac{(T_i - T_{i-1})(t - T_{i-1})}{T_i - T_{i-1} - 1/D} \right) \right) \mathbb{1}_{\{T_{i-1} < t \leq T_i\}}$$

and then a process  $(\tilde{S}_t)_{t \in [0, T_n]}$  by a time change of  $\mathbb{S}$  via  $f$ , i.e.,  $\tilde{S}_t = \mathbb{S}_{f(t)}$ . Note that  $f(T_i - 1/D) = T_i$ , as required. We argue below that (3.9) implies that we have  $|G(\mathbb{S}) - G(\tilde{\mathbb{S}})| \rightarrow 0$  as  $D \rightarrow \infty$ .

Now for every  $N \in \mathbb{N}$ , take  $\mathbb{P}^{(N)} \in \widehat{\mathcal{M}}_{\mathcal{I}}^{\kappa} \cap \mathcal{M}_{\mu, \mathfrak{P}}^{1/N}$  such that

$$\mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S})] \geq \sup_{\mathbb{P} \in \tilde{\mathcal{M}}_{\mathcal{I}} \cap \mathcal{M}_{\mu, \mathfrak{P}}^{1/N}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] - 1/N.$$

Since  $\mathbb{S}_{T_i} = \tilde{S}_{T_i}$ , we have in particular  $\mathcal{L}_{\mathbb{P}^{(N)}}(\mathbb{S}_{T_i}) = \mathcal{L}_{\mathbb{P}^{(N)}}(\tilde{S}_{T_i})$  for all  $i \leq n$ . Also, being a time change of  $\mathbb{S}$ , the process  $(\tilde{S}_t)_{t \in [0, T_n]}$  is a martingale (in the time-changed filtration). It follows that its distribution  $\mathbb{P}^{(N)} \circ (\tilde{S}_t)^{-1}$  is an element of  $\mathcal{M}_{\mu, \mathfrak{P}}^{1/N}$  as  $\mathfrak{P}^{1/N}$  is time-invariant, by Lemma 4.7.

*Step 3: Second time change: introducing a lower bound on the time step.* The second time change ensures that we can bound from below the difference between any two consecutive stopping times in the Lebesgue discretisation in Definition 4.1. We want to do this by adding a constancy interval of length  $\delta$  to each step of the discretisation. As we have squeezed the paths above, we have length  $1/D$  to use up while still keeping the time changes to within the intervals  $[T_{i-1}, T_i]$ . Taking suitably small  $\delta$ , this allows us, with high probability, to alter all the steps in the Lebesgue discretisation.

For ease of notation, it is helpful to rename the elements of the set

$$\{\tau_j^{(D)} : j \leq m^{(D)}\} \cup \{T_i : i = 1, \dots, n\}$$

as follows. We define a sequence of stopping times  $\tau_{i,j}^{(D)} : \Omega \rightarrow [T_{i-1}, T_i]$  and  $m_i^{(D)} : \Omega \rightarrow \mathbb{N}_+$  in a recursive manner. Set  $T_0 = m_0^{(D)}(\mathbb{S}) = \tau_{0,-1}^{(D)} = 0$  and for  $i = 1, \dots, n$ , set  $\tau_{i,0}^{(D)}(\mathbb{S}) = T_{i-1}$  and let

$$\begin{aligned} \tau_{i,1}^{(D)}(\mathbb{S}) &= \inf \left\{ t \geq T_{i-1} : \left| \mathbb{S}_t - \mathbb{S}_{\tau_{i-1, m_{i-1}^{(D)}(\mathbb{S})-1}^{(D)}}(\mathbb{S}) \right| = \frac{1}{2D} \right\} \wedge T_i, \\ \tau_{i,k}^{(D)}(\mathbb{S}) &= \inf \left\{ t \geq \tau_{i,k-1}(\mathbb{S}) : \left| \mathbb{S}_t - \mathbb{S}_{\tau_{i,k-1}^{(D)}(\mathbb{S})} \right| = \frac{1}{2D} \right\} \wedge T_i, \\ m_i^{(D)}(\mathbb{S}) &= m_{i-1}^{(D)}(\mathbb{S}) + \min\{k \in \mathbb{N} : \tau_{i,k}^{(D)}(\mathbb{S}) = T_i\}. \end{aligned}$$

It follows that for any  $S \in \mathcal{I}$ ,

$$m^{(D)}(S) \leq m_n^{(D)}(S) \leq m^{(D)}(S) + n - 1. \tag{4.13}$$

Set  $\Theta = 2\lceil \kappa^2 2^{6D} \rceil + n$  and  $\delta = 1/(4D\Theta^2)$ . We now define a sequence of stopping times  $\sigma_{i,j} : \Omega \rightarrow [0, T_n]$  by  $\sigma_{i,0}(S) := T_{i-1}$ ,  $\sigma_{i,\Theta+1}(S) := T_i$ , and for  $j \leq \Theta$ , we put

$$\sigma_{i,j}(S) := (\tau_{i,j}^{(D-8)}(S) + \delta j) \wedge (T_i - 1/(2D)) \quad \text{if } j < m_i^{(D-8)}(S),$$

while  $\sigma_{i,j}(S) := T_i - 1/(2D)$  otherwise, where  $i = 1, \dots, n$ . Then it follows from the definition that

$$T_{i-1} = \sigma_{i,0}(S) \leq \sigma_{i,1}(S) \leq \dots \leq \sigma_{i,\Theta}(S) < \sigma_{i,\Theta+1}(S) = T_i$$

for all  $S \in \Omega$ . Further, since the process  $\tilde{\mathbb{S}}$  is always constant on  $[T_i - 1/D, T_i]$ , we have  $\tau_{i,j}^{(D-8)}(\tilde{\mathbb{S}}) \leq T_i - 1/D$  and hence for  $j \leq \Theta \wedge (m_i^{(D-8)}(\tilde{\mathbb{S}}) - 1)$  that

$$\sigma_{i,j}(\tilde{\mathbb{S}}) \leq \tau_{i,m_i^{(D-8)}-1}^{(D-8)}(\tilde{\mathbb{S}}) + \delta\Theta \leq T_i - \frac{1}{D} + \frac{1}{4D\Theta} < T_i - \frac{1}{2D}.$$

Also, for all  $j = 1, \dots, (\Theta \wedge (m_i^{(D-8)}(\tilde{\mathbb{S}}) - 1))$ ,

$$\sigma_{i,j}(\tilde{\mathbb{S}}) - \sigma_{i,j-1}(\tilde{\mathbb{S}}) = \delta + (\tau_{i,j}^{(D-8)}(\tilde{\mathbb{S}}) - \tau_{i,j-1}^{(D-8)}(\tilde{\mathbb{S}})) \geq \delta.$$

We are now ready to define the time-changed process  $\check{\mathbb{S}}$  by

$$\begin{aligned} \check{\mathbb{S}}_t = & \sum_{i=1}^n \left( \sum_{j=0}^{\Theta-1} \tilde{\mathbb{S}}_{\tau_{i,j}^{(D-8)}(\tilde{\mathbb{S}}) + (t - \sigma_{i,j}(\tilde{\mathbb{S}}) - \delta) + \mathbb{1}_{[\sigma_{i,j}(\tilde{\mathbb{S}}), \sigma_{i,j+1}(\tilde{\mathbb{S}})]}(t)} \right. \\ & \left. + \tilde{\mathbb{S}}_{(\tau_{i,\Theta}^{(D-8)}(\tilde{\mathbb{S}}) + \frac{1}{T_i-1} - \frac{1}{T_i - \sigma_{i,\Theta}(\tilde{\mathbb{S}})}) \wedge T_i} \mathbb{1}_{[\sigma_{i,\Theta}(\tilde{\mathbb{S}}), T_i]}(t) \right). \end{aligned}$$

Observe that  $\check{\mathbb{S}}$  is a (continuous) time change of  $\tilde{\mathbb{S}}$  and  $\check{\mathbb{S}}_{T_i} = \tilde{\mathbb{S}}_{T_i} = \mathbb{S}_{T_i}$  for  $i \leq n$ . As before, this implies that  $\check{\mathbb{S}}$  remains a martingale and  $\mathbb{P}^{(N)} \circ (\check{\mathbb{S}}_t)^{-1} \in \mathcal{M}_{\mu, \mathfrak{P}}^{1/N}$ .

We argue now that  $|G(\mathbb{S}) - G(\check{\mathbb{S}})|$  is small for large  $D$ . To this end, we approximate a path  $S$  with a piecewise constant function  $\tilde{F}^{(D)}(S)$  which jump at the times  $\tau_{i,j}^{(D)}$ . A similar discretisation is used later in Sect. 5; see (5.2). For  $S \in \Omega$ , consider

$$\tilde{F}_t^{(D)}(S) = \sum_{i=1}^n \sum_{j=0}^{m_i^{(D)}-1} S_{\tau_{i,j}^{(D)}} \mathbb{1}_{[\tau_{i,j}^{(D)}, \tau_{i,j+1}^{(D)})}(t) + S_{T_n} \mathbb{1}_{\{T_n\}}(t), \quad t \in [0, T].$$

Then the time-continuity property of  $G$  in (3.9) ensures that

$$\begin{aligned} |G(\mathbb{S}) - G(\check{\mathbb{S}})| & \leq |G(\mathbb{S}) - G(\tilde{F}^{(D)}(\mathbb{S}))| + |G(\check{\mathbb{S}}) - G(\tilde{F}^{(D)}(\check{\mathbb{S}}))| \\ & \quad + |G(\tilde{F}^{(D)}(\mathbb{S})) - G(\tilde{F}^{(D)}(\check{\mathbb{S}}))| \\ & \leq 2f_e(2^{-D+9}) + \frac{2nL\|\mathbb{S}\|}{D}. \end{aligned} \tag{4.14}$$

Similarly, for any  $S \in \Omega$  with  $m_n^{(D-8)}(\check{\mathbb{S}}(S)) = m_n^{(D-8)}(S) \leq \Theta$ , again by (3.9), we have

$$\begin{aligned} |G(\check{\mathbb{S}}(S)) - G(\check{\mathbb{S}}(S))| &\leq \left| G(\check{\mathbb{S}}(S)) - G(\tilde{F}^{(D-8)}(\check{\mathbb{S}}(S))) \right| \\ &\quad + \left| G(\check{\mathbb{S}}(S)) - G(\tilde{F}^{(D-8)}(\check{\mathbb{S}}(S))) \right| \\ &\quad + \left| G(\tilde{F}^{(D-8)}(\check{\mathbb{S}}(S))) - G(\tilde{F}^{(D-8)}(\check{\mathbb{S}}(S))) \right| \\ &\leq 2f_e(2^{-D+9}) + nL\|\check{\mathbb{S}}(S)\|\Theta\delta \\ &\leq 2f_e(2^{-D+9}) + nL\|\mathbb{S}(S)\|/D, \end{aligned} \tag{4.15}$$

when  $D$  is sufficiently large. From (4.12), the Markov inequality gives

$$\mathbb{P}^{(N)}[\{S \in \mathcal{I} : m^{(D-8)}(S) \geq \Theta - n + 2\}] \leq \frac{2\kappa + 2}{\kappa 2^D},$$

and hence by (4.13),

$$\mathbb{P}^{(N)}[\{S \in \mathcal{I} : m_n^{(D-8)}(S) \geq \Theta + 1\}] \leq \frac{2\kappa + 2}{\kappa 2^D}. \tag{4.16}$$

Furthermore, by (4.15) and (4.16),

$$\begin{aligned} |\mathbb{E}_{\mathbb{P}^{(N)}}[G(\check{\mathbb{S}})] - \mathbb{E}_{\mathbb{P}^{(N)}}[G(\check{\mathbb{S}})]| &\leq 2\kappa\mathbb{P}^{(N)}[m_n^{(D-8)}(\check{\mathbb{S}}) > \Theta] + 2f_e(2^{-D+9}) \\ &\quad + nL\mathbb{E}_{\mathbb{P}^{(N)}}[\|\mathbb{S}\|]/D \\ &\leq \frac{4\kappa + 4}{2^D} + 2f_e(2^{-D+9}) \\ &\quad + nLV_{\mathcal{X}_n, \mathcal{P}, \mathcal{I}}(\|\mathbb{S}\|)/D. \end{aligned} \tag{4.17}$$

*Step 4: Third time change: controlling the increments of the quadratic variation.* We say that  $\omega \in \mathcal{C}([0, T], \mathbb{R})$  admits a quadratic variation if

$$\lim_{N \rightarrow \infty} \sum_{k=0}^{m^{(N)}(\omega)-1} \left( \omega_{\tau_k^{(N)}(\omega) \wedge t} - \omega_{\tau_{k+1}^{(N)}(\omega) \wedge t} \right)^2$$

exists and is a continuous function for  $t \in [0, T]$ . In this case, we denote this limit with  $\langle \omega \rangle$  and otherwise we let  $\langle \omega \rangle$  be zero. In addition, for  $S \in \Omega$ , we say  $S$  admits a quadratic variation if  $S^{(i)}$  admits a quadratic variation for any  $i \leq d + K$ .

It follows from Theorem 4.30.1 in Rogers and Williams [46] and its proof that for any  $\mathbb{P} \in \mathcal{M}$ ,  $\langle \mathbb{S} \rangle := (\langle \mathbb{S}^{(1)} \rangle, \dots, \langle \mathbb{S}^{(d+K)} \rangle)$  agrees  $\mathbb{P}$ -a.s. with the classical definition of the quadratic variation of  $\mathbb{S}$  under  $\mathbb{P}$ , i.e.,  $\mathbb{S}^2 - \langle \mathbb{S} \rangle$  is a  $\mathbb{P}$ -martingale. Further, Doob’s inequality gives for all  $i \leq d$  that

$$\mathbb{E}_{\mathbb{P}^{(N)}}[\|\check{\mathbb{S}}^{(i)}\|^p] \leq \left( \frac{p}{p-1} \right)^p \int_{[0, \infty)} x^p \mu_n^{(i)}(dx),$$

and by the BDG inequalities, there exist constants  $c_p, C_p \in (0, \infty)$  such that

$$c_p \mathbb{E}_{\mathbb{P}^{(N)}} [\langle \check{\mathbb{S}}^{(i)} \rangle_{T_n}^{p/2}] \leq \mathbb{E}_{\mathbb{P}^{(N)}} [\|\check{\mathbb{S}}^{(i)}\|^p] \leq C_p \mathbb{E}_{\mathbb{P}^{(N)}} [\langle \check{\mathbb{S}}^{(i)} \rangle_{T_n}^{p/2}].$$

It follows that

$$\mathbb{E}_{\mathbb{P}^{(N)}} \left[ \sum_{i=1}^{d+K} \langle \check{\mathbb{S}}^{(i)} \rangle_{T_n}^{p/2} \right] \leq K_1,$$

where  $K_1 := \frac{1}{c_p} ((\frac{p}{p-1})^p \sum_{i=1}^d \int_{[0, \infty)} x^p \mu_n^{(i)}(dx) + K\kappa^p)$ .

In the following, we want to modify  $\check{\mathbb{S}}$  on

$$\begin{aligned} \tilde{\mathcal{I}} &:= \{S \in \mathcal{I} : \check{\mathbb{S}}(S) \text{ admits a quadratic variation}\} \\ &= \{S \in \mathcal{I} : S \text{ admits a quadratic variation}\} \end{aligned}$$

to obtain another process  $\check{\check{\mathbb{S}}}$  with a better control of the quadratic variation, while its law remains in  $\mathcal{M}_{\mu, \mathfrak{P}}^{1/N}$ . In fact,  $\check{\check{\mathbb{S}}}$  will be obtained as a time change of  $\check{\mathbb{S}}$  on each interval  $[\sigma_{i,j}(\check{\mathbb{S}}), \sigma_{i,j+1}(\check{\mathbb{S}})]$ . Then by the continuity of  $G$ , it follows that

$$|G(\check{\mathbb{S}}(S)) - G(\check{\check{\mathbb{S}}}(S))| \leq f_e(2^{-D+9}), \quad \forall S \in \tilde{\mathcal{I}} \cap \{h \in \mathcal{I} : m_n^{(D-8)}(\check{\mathbb{S}}(h)) \leq \Theta\}.$$

This together with (4.16) and the fact that  $\mathbb{P}[\tilde{\mathcal{I}}] = 1$  for any  $\mathbb{P} \in \mathcal{M}_{\mathcal{I}}$  yields

$$\begin{aligned} |\mathbb{E}_{\mathbb{P}^{(N)}}[G(\check{\mathbb{S}}) - G(\check{\check{\mathbb{S}}})] &\leq f_e(2^{-D+9}) + 2\kappa \mathbb{P}^{(N)}[\{S \in \mathcal{I} : m_n^{(D-8)}(\check{\mathbb{S}}(S)) \geq \Theta + 1\}] \\ &\leq f_e(2^{-D+9}) + \frac{4\kappa + 4}{2^D}. \end{aligned}$$

Hence, by (4.14) and (4.17),

$$\begin{aligned} |\mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S}) - G(\check{\check{\mathbb{S}}})] &\leq 5f_e(2^{-D+9}) + \frac{2nL\|\mathbb{S}\|}{D} + \frac{8\kappa + 8}{2^D} \\ &\quad + \frac{2nLV\chi_{n, \mathcal{P}, \mathcal{I}}(\|\mathbb{S}\|)}{D}. \end{aligned} \tag{4.18}$$

First, for every  $i, j, k$ , define  $\rho^{(i,j,k)} : \Omega \rightarrow [T_{i-1}, T_i]$  by

$$\rho^{(i,j,k)}(S) = \sigma_{i,j}(\check{\mathbb{S}}(S)) + \delta(1 - 2^{-k+1}).$$

Then for  $i = 1, \dots, n, j = 0, 1, \dots$ , let  $\theta_t^{(i,j,0)} = \sigma_{i,j}$  and define recursively for  $k = 1, 2, \dots$  a change of time  $\theta^{(i,j,k)} : \mathcal{I} \times [\rho^{i,j,k}, \rho^{i,j,k+1}] \rightarrow [T_{i-1}, T_i]$  by

$$\begin{aligned} \theta_t^{(i,j,k)}(S) &= \inf \left\{ u \geq \theta_{\rho^{i,j,k}}^{(i,j,k-1)}(S) : \sum_{\ell=1}^{d+K} (\langle \check{\mathbb{S}}^{(\ell)}(S) \rangle_u - \langle \check{\mathbb{S}}^{(\ell)}(S) \rangle_{\theta_{\rho^{i,j,k}}^{(i,j,k-1)}}}) \right. \\ &\quad \left. > 2^k(t - \rho^{(i,j,k)})/\delta \right\} \wedge \sigma_{i,j+1}(\check{\mathbb{S}}(S)) \end{aligned}$$

$$\text{for } t \in [\rho^{i,j,k}, \rho^{i,j,k+1}], S \in \tilde{\mathcal{I}}.$$

For  $S \in \Omega \setminus \tilde{\mathcal{I}}$ , set  $\theta_t^{(i,j,k)}(S) = t, 0 \leq t \leq T_n$ .

We consider a time change of  $\check{S}$  via the  $\theta^{(i,j,k)}$ , defined by  $\check{S}_t := \check{S}_{\theta_t^{(i,j,k)}(\mathbb{S})}$  for  $t \in [\rho^{(i,j,k)}(\mathbb{S}), \rho^{(i,j,k+1)}(\mathbb{S})]$  for all  $i, j, k$  as above. Note that  $\theta_{\rho^{i,j,k}}^{(i,j,k-1)} = \theta_{\rho^{i,j,k}}^{(i,j,k)}$  so that the resulting process is continuous. Consider  $S \in \tilde{\mathcal{I}}$  and  $i, j$  such that we have  $\sigma_{i,j+1}(\check{S}(S)) - \sigma_{i,j}(\check{S}(S)) > 0$ , as otherwise everything collapses to one point. Then the quadratic variation of  $\check{S}(S)$  grows on  $[\rho^{(i,j,k)}(S), \rho^{(i,j,k+1)}(S)]$  linearly at the rate  $2^k/\delta$ , and  $\rho^{(i,j,k+1)}(S) - \rho^{(i,j,k)}(S) = 2^{-k}\delta$ . In particular,  $\check{S}$  accumulates one unit of quadratic variation over each interval  $[\rho^{(i,j,k)}(S), \rho^{(i,j,k+1)}(S)]$  for  $k$  increasing until the total quadratic variation of  $\check{S}$  on  $[\sigma_{i,j+1}(\check{S}(S)) - \sigma_{i,j}(\check{S}(S))]$  is exhausted. Trivially bounding the quadratic variation of  $\check{S}$  over a small interval by its quadratic variation over  $[0, T_n]$ , we see that

$$\sum_{\ell=1}^{d+K} (\check{S}^{(\ell)}(S))_t - \check{S}^{(\ell)}(S)_s \leq 2^{k_0}|t-s|/\delta \quad \text{for } \sigma_{i,j}(\check{S}(S)) \leq s \leq t \leq \sigma_{i,j+1}(\check{S}(S)),$$

whenever  $S \in \tilde{\mathcal{I}}$  is such that  $\sum_{i=1}^{d+K} \langle \check{S}^{(i)}(S) \rangle_{T_n} \leq k_0$ . Therefore, for such  $S$ , we have

$$\sum_{\ell=1}^{d+K} (\check{S}^{(\ell)})_t - \langle \check{S}^{(\ell)} \rangle_s \leq 2^{k_0+1}|t-s|/\delta, \quad \forall s, t \in [0, T_n] \text{ with } |t-s| \leq \delta. \tag{4.19}$$

We can ensure this happens with large probability since by Markov’s inequality,

$$\begin{aligned} \mathbb{P}^{(N)} \left[ \sum_{i=1}^{d+K} \langle \check{S}^{(i)} \rangle_{T_n} > k_0 \right] &= \mathbb{P}^{(N)} \left[ \sum_{i=1}^{d+K} \langle \check{S}^{(i)} \rangle_{T_n} > k_0 \right] \\ &\leq \frac{\mathbb{E}_{\mathbb{P}^{(N)}} [\sum_{i=1}^{d+K} \langle \check{S}^{(i)} \rangle_{T_n}^{p/2}]}{k_0^{p/2}} \leq K_1 k_0^{-p/2}. \end{aligned}$$

Finally, we observe that each  $\theta_t^{(i,j,k)}(\mathbb{S})$  is a stopping time relative to the natural filtration of  $\check{S}$ , and hence  $\check{S}$  is a continuous  $\mathbb{P}^{(N)}$ -martingale.

*Step 5: Tightness of the measures through tightness of the quadratic variation processes.* Together with (4.19), by the Arzelà–Ascoli theorem, the above implies that the family  $\{\mathbb{P}^{(N)} \circ (\check{S})^{-1} : N \in \mathbb{N}\}$  is tight in  $\mathcal{C}([0, T_n], \mathbb{R}^{d+K})$ . Then by Theorem VI.4.13 in Jacod and Shiryaev [35],  $\{\mathbb{P}^{(N)} \circ \check{S}^{-1}\}_{N \in \mathbb{N}}$  is tight in  $\mathbb{D}([0, T_n], \mathbb{R}^{d+K})$ , the space of right-continuous functions with left limits. By Theorem VI.3.21 in Jacod and Shiryaev [35], this implies that for all  $\epsilon > 0, \eta > 0$ , there are  $N_0 \in \mathbb{N}$  and  $\theta > 0$  with

$$N \geq N_0 \implies \mathbb{P}^{(N)}[w'_{T_n}(\check{S}, \theta) \geq \eta] \leq \epsilon,$$

where  $w'_{T_n}$  is defined by

$$w'_{T_n}(S, \theta) = \inf \left\{ \max_{i \leq r} \sup_{t_{i-1} \leq s \leq t < t_i} |S_t - S_s| : r \in \mathbb{N}, 0 = t_0 < \dots < t_r = T_n, \inf_{i < r} (t_i - t_{i-1}) \geq \theta \right\}.$$



Clearly, for  $S \in \Omega$ , continuity of  $S$  implies that

$$w_{T_n}(S, \theta) := \sup\{|S_t - S_s| : 0 \leq s < t \leq T_n, t - s \leq \theta\} \leq 2w'_{T_n}(S, \theta).$$

Then we have

$$N \geq N_0 \implies \mathbb{P}^{(N)}[w_{T_n}(\ddot{S}, \theta) \geq 2\eta] \leq \epsilon,$$

which then by Theorem VI.1.5 in Jacod and Shiryaev [35] implies that the family  $\{\mathbb{P}^{(N)} \circ \ddot{S}^{-1} : N \in \mathbb{N}\}$  is tight, now in  $\mathcal{C}([0, T_n], \mathbb{R}^{d+K})$ .

*Step 6: Tightness gives exact duality.* By tightness, there exists a converging subsequence  $\{\mathbb{P}^{(N_k)} \circ \ddot{S}^{-1}\}$  such that  $\mathbb{P}^{(N_k)} \circ \ddot{S}^{-1} \rightarrow \mathbb{P}$  weakly for some probability measure  $\mathbb{P}$  on  $\Omega$ . Consequently,

$$\lim_{k \rightarrow \infty} \mathbb{E}_{\mathbb{P}^{(N_k)}}[G(\ddot{S})] = \mathbb{E}_{\mathbb{P}}[G(S)].$$

In addition, if  $\mathbb{P}$  is an element of  $\mathcal{M}_{\mu, \mathfrak{F}}$ , then

$$\begin{aligned} V_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}(G) &\leq \tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}(G) \\ &\leq \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\mu, \mathfrak{F}}^{1/N}} \mathbb{E}_{\mathbb{P}} \left[ G(\mathbb{S}) - \kappa 2^D \wedge \frac{\sqrt{m^{(D-8)}(\mathbb{S})}}{2^{2D}} \right] + \frac{c_2}{2^D} \\ &\leq \liminf_{N \rightarrow \infty} \mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S})] + c_2/2^D \leq \liminf_{N \rightarrow \infty} \mathbb{E}_{\mathbb{P}^{(N)}}[G(\ddot{S})] + e(D) \\ &\leq \lim_{k \rightarrow \infty} \mathbb{E}_{\mathbb{P}^{(N_k)}}[G(\ddot{S})] + e(D) \leq \mathbb{E}_{\mathbb{P}}[G(S)] + e(D) \\ &\leq P_{\mu, \mathfrak{F}}(G) + e(D), \end{aligned}$$

where  $e(x) := 5f_e(2^{-x+9}) + \frac{2nL\|\mathbb{S}\|}{x} + \frac{c_2+8\kappa+8}{2^x} + \frac{2nLV_{\mathcal{X}_n, \mathcal{P}, \mathcal{I}}(\|\mathbb{S}\|)}{2^x}$  and the third inequality follows from (4.18). Recalling that  $\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{F}} = \tilde{P}_{\mu, \mathfrak{F}}^x$  and letting  $D \rightarrow \infty$ , we obtain the desired equality  $\tilde{P}_{\mu, \mathfrak{F}} = P_{\mu, \mathfrak{F}}$  and conclude that

$$\tilde{V}_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}(G) = V_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}(G) = P_{\mu, \mathfrak{F}}(G) = \tilde{P}_{\mu, \mathfrak{F}}(G).$$

It remains to argue that  $\mathbb{P}$  is an element of  $\mathcal{M}_{\mu, \mathfrak{F}}$ . First, it is straightforward to see that  $\mathbb{S}$  is a  $\mathbb{P}$ -martingale and  $\mathcal{L}_{\mathbb{P}}(S_{T_i}) = \mu_i$  for any  $i \leq n$ . To show that  $\mathbb{P}[\mathbb{S} \in \mathfrak{F}] = 1$ , notice that by the Portemanteau theorem, for every  $\epsilon > 0$ ,

$$\mathbb{P}[\mathbb{S} \in \overline{\mathfrak{F}}^\epsilon] \geq \limsup_{k \rightarrow \infty} \mathbb{P}^{(N_k)}[\mathbb{S} \in \overline{\mathfrak{F}}^\epsilon] \geq \limsup_{k \rightarrow \infty} \mathbb{P}^{(N_k)}[\mathbb{S} \in \mathfrak{F}^{1/N_k}] = 1.$$

Therefore, it follows from Remark 4.6 and monotone convergence that

$$\mathbb{P}[\mathbb{S} \in \mathfrak{F}] = \lim_{\epsilon \searrow 0} \mathbb{P}[\mathbb{S} \in \overline{\mathfrak{F}}^\epsilon] = 1,$$

and hence  $\mathbb{P} \in \mathcal{M}_{\mu, \mathfrak{F}}$ . □

### 5 Pricing–hedging duality without constraints

This and the subsequent section are devoted to establishing the crucial pricing–hedging duality result in the absence of constraints, which was exploited in all the proofs above.

**Theorem 5.1** *Under Assumption 3.1, for any  $\alpha, \beta \geq 0$  and  $D \in \mathbb{N}$ ,*

$$\mathbf{V}_{\mathcal{I}}(G - \beta\sqrt{m^{(D)}} \wedge \alpha) \leq \mathbf{P}_{\mathcal{I}}(G - \beta\sqrt{m^{(D-2)}} \wedge \alpha), \tag{5.1}$$

where  $m^{(D)}$  is defined in Definition 4.1.

*Remark 5.2* As a by-product of the proof of Theorem 5.1, (5.1) still holds true when the probabilistic models  $\mathbb{P}$  are restricted to those which arise within a Brownian setup, i.e.,  $\mathbb{P}$  satisfies (4.7).

The strategy of the proof is inspired by Dolinsky and Soner [25] and proceeds via discretisation, of the dual side in Sect. 5.1 and of the primal side in Sect. 5.3. The duality between the discrete counterparts is obtained by using classical probabilistic results of Föllmer and Kramkov [29].

#### 5.1 Discretisation of the dual

##### 5.1.1 A discrete-time approximation through simple strategies

The proof of Theorem 5.1 is based on a discretisation method involving a discretisation of the path space into a countable set of piecewise constant functions. These are obtained as a “shift” of the “Lebesgue discretisation” of a path. Recall from Definition 4.1 that for a positive integer  $N$  and any  $S \in \Omega$ ,  $\tau_0^{(N)}(S) = 0$ ,  $m_0^{(N)}(S) = 0$ ,

$$\tau_k^{(N)}(S) = \inf \left\{ t \geq \tau_{k-1}^{(N)}(S) : |S_t - S_{\tau_{k-1}^{(N)}(S)}| = \frac{1}{2^N} \right\} \wedge T$$

and  $m^{(N)}(S) = \min\{k \in \mathbb{N} : \tau_k^{(N)}(S) = T\}$ . Now denote by  $\mathcal{A}_N$  the set of  $\gamma \in \mathcal{A}$  with  $|\gamma| \leq N$  and for which trading in the risky assets only takes place at the moments  $0 = \tau_0^{(N)}(S) < \tau_1^{(N)}(S) < \dots < \tau_{m^{(N)}(S)}^{(N)}(S) = T$ . Set

$$\mathbf{V}_{\mathcal{I}}^{(N)}(G) := \inf\{x : \exists \gamma \in \mathcal{A}_N \text{ which superreplicates } G - x\}.$$

Then it is obvious from the definition of  $\mathbf{V}_{\mathcal{I}}^{(N)}$  that  $\mathbf{V}_{\mathcal{I}}^{(N_1)}(G) \geq \mathbf{V}_{\mathcal{I}}^{(N_2)}(G) \geq \mathbf{V}_{\mathcal{I}}(G)$  for any  $N_2 \geq N_1$ , and in fact, the following result states that  $\mathbf{V}_{\mathcal{I}}^{(N)}(G)$  converges to  $\mathbf{V}_{\mathcal{I}}(G)$  asymptotically.

**Corollary 5.3** *Under the assumptions of Theorem 5.1,*

$$\lim_{N \rightarrow \infty} \mathbf{V}_{\mathcal{I}}^{(N)}(G) = \mathbf{V}_{\mathcal{I}}(G).$$

### 5.1.2 A countable class of piecewise constant functions

In this section, we construct a countable set of piecewise constant functions which can approximate any continuous function  $S$  to a certain degree. It is achieved in three steps. The first step is to use the Lebesgue partition defined in the last section to discretise a continuous function into a piecewise constant function whose jump times are the stopping times. Due to the arbitrary nature of jump times and jump sizes, the set of piecewise constant functions  $F^{(N)}(S)$ , generated through this procedure over all  $S$ , is uncountable. To overcome this, in the subsequent two steps, we restrict the jump times and sizes to a countable set and hence define a class of approximating schemes. As explained in Sect. 3.1, our methods are closely inspired by [25], but in order to deal with payoff functions which are uniformly continuous, so that in applications we can include static hedging in options with different maturities, we had to devise an improved discretisation scheme.

We denote by  $\mathcal{D}([0, T], \mathbb{R}^{d+K})$  the set of all  $\mathbb{R}^{d+K}$ -valued measurable functions on  $[0, T]$  and by  $\mathbb{D}([0, T], \mathbb{R}^{d+K})$  the subset of all right-continuous functions with left limits.

*Step 1.* Let  $\tau_k^{(N)}(S)$  and  $m^{(N)}(S)$  be defined as in Sect. 5.1.1. To simplify the notation, in this section, we often suppress their dependences on  $S$  and  $N$  and simply write

$$m = m^{(N)}(S), \quad \tau_k = \tau_k^{(N)}(S).$$

Our first “naive” approximation  $F^{(N)} : \Omega \rightarrow \mathbb{D}([0, T], \mathbb{R}^{d+K})$  is defined as

$$F_t^{(N)}(S) = \sum_{k=0}^{m-1} S_{\tau_k} \mathbb{1}_{[\tau_k, \tau_{k+1})}(t) + S_T \mathbb{1}_{\{T\}}(t) \quad \text{for } t \in [0, T], S \in \Omega. \tag{5.2}$$

Note that  $F^{(N)}(\mathbb{S})$  is piecewise constant and  $\|F^{(N)}(\mathbb{S}) - \mathbb{S}\| \leq 1/2^N$ .

*Step 2.* Define a map

$$\begin{aligned} \zeta^{(N)} : \mathbb{R}_+^d &\rightarrow A^{(N)} := \{2^{-N}k : k = (k_1, \dots, k_{d+K}) \in \mathbb{N}^{d+K}\}, \\ \zeta^{(N)}(x)_i &:= 2^{-N} \lceil 2^N x_i \rceil, \quad i = 1, \dots, d + K. \end{aligned}$$

We then define our second approximation  $\check{F}^{(N)} : \Omega \rightarrow \mathbb{D}([0, T], \mathbb{R}^{d+K})$  by

$$\begin{aligned} \check{F}_t^{(N)}(S) &= (S_0 - \zeta^{(N+1)}(S_{\tau_1})) + \sum_{k=0}^{m-2} \zeta^{(N+k+1)}(S_{\tau_{k+1}}) \mathbb{1}_{[\tau_k, \tau_{k+1})}(t) \\ &\quad + \zeta^{(N+m)}(S_{\tau_m}) \mathbb{1}_{[\tau_{m-1}, T]}(t), \quad t \in [0, T]. \end{aligned}$$

*Step 3.* We now construct the shifted jump times  $\hat{\tau}_k^{(N)} : \Omega \rightarrow \mathbb{Q}_+ \cup \{T\}$ . Firstly, set  $\hat{\tau}_0^{(N)} = 0$ . Then for any  $S \in \Omega$  and  $k = 1, \dots, m^{(N)}(S)$ , define  $\Delta \tau_k^{(N)} := \tau_k^{(N)} - \tau_{k-1}^{(N)}$

and let  $\Delta \hat{\tau}_k^{(N)} = p_k/q_k$  with

$$(p_k, q_k) = \operatorname{argmin} \left\{ p + q : (p, q) \in \mathbb{N}^2, \right. \\ \left. \tau_{k-1}^{(N)} - \hat{\tau}_{k-1}^{(N)} < \frac{p}{q} \leq \Delta \tau_k^{(N)} + \tau_{k-1}^{(N)} - \hat{\tau}_{k-1}^{(N)} \right\}$$

if  $k < m^{(N)}(S)$  and  $\Delta \hat{\tau}_k^{(N)} = T - \hat{\tau}_{m^{(N)}-1}^{(N)}$  otherwise. Finally, set  $\hat{\tau}_k^{(N)} := \sum_{i=1}^k \Delta \hat{\tau}_i^{(N)}$ . Here we also suppress the dependences of these shifted jump times on  $S$  and  $N$  and write  $\hat{\tau}_k = \hat{\tau}_k^{(N)}(S)$ . Clearly  $0 = \hat{\tau}_0 < \hat{\tau}_1 < \hat{\tau}_2 < \dots < \hat{\tau}_m = T$ ,  $\tau_{k-1} < \hat{\tau}_k \leq \tau_k$  for all  $k < m$  and  $\hat{\tau}_m = \tau_m = T$ . These  $\hat{\tau}$  are the shifted versions of the  $\tau$  and are uniquely defined for any  $S$ . We are going to use the  $\hat{\tau}$  to define a class of approximating schemes.

We can define an approximation  $\hat{F}^{(N)} : \Omega \rightarrow \mathbb{D}([0, T], \mathbb{R}^{d+K})$  by

$$\hat{F}_t^{(N)}(S) = (S_0 - \zeta^{(N+1)}(S_{\tau_1})) + \sum_{k=0}^{m-2} \zeta^{(N+k+1)}(S_{\tau_{k+1}}) \mathbb{1}_{[\hat{\tau}_k, \hat{\tau}_{k+1})}(t) \\ + \zeta^{(N+m)}(S_{\tau_m}) \mathbb{1}_{[\hat{\tau}_{m-1}, T]}(t), \quad t \in [0, T]. \tag{5.3}$$

Notice that  $\hat{F}^{(N)}(S)$  is piecewise constant and

$$\|\hat{F}^{(N)}(S) - S\| \leq \|\hat{F}^{(N)}(S) - \check{F}^{(N)}(S)\| + \|\check{F}^{(N)}(S) - F^{(N)}(S)\| + \|F^{(N)}(S) - S\| \\ \leq \frac{2}{2^{N-1}} + \frac{2}{2^N} + \frac{1}{2^N} < \frac{1}{2^{N-3}}. \tag{5.4}$$

**Definition 5.4** Let  $\hat{\mathbb{D}}^{(N)} \subseteq \mathbb{D}([0, T], \mathbb{R}^{d+K})$  be the set of functions  $f = (f^{(i)})_{i=1}^{d+K}$  which satisfy the following:

1. For any  $i = 1, \dots, d + K$ ,  $f^{(i)}(0) = 1$ ;
2.  $f$  is piecewise constant with jumps at times  $t_1, \dots, t_{\ell-1} \in \mathbb{Q}_+$  for some  $\ell < \infty$ , where  $t_0 = t_{\ell_0} = 0 < t_1 < t_2 < \dots < t_{\ell-1} < T$ ;
3. For any  $k = 1, \dots, \ell - 1$  and  $i = 1, \dots, d + K$ ,  $f^{(i)}(t_k) - f^{(i)}(t_{k-1}) = j/2^{N+k}$  for  $j \in \mathbb{Z}$  with  $|j| \leq 2^k$ ;
4.  $\inf_{t \in [0, T], 1 \leq i \leq d+K} f^{(i)}(t) \geq -2^{-N+3}$ ;
5.  $\|f^{(i)}\| \leq \kappa + 1$  for  $i = d + 1, \dots, d + K$ , where  $\kappa = \max_{1 \leq j \leq K} \frac{\|X_j^{(c)}\|_\infty}{\mathcal{P}(X_j^{(c)})}$ ;
6. If  $f^{(i)}(t_k) = -2^{-N+3}$  for some  $i \leq d + K$  and  $k \leq \ell - 1$ , then  $f(t_j) = f(t_k)$  for all  $k < j < \ell$ ;
7. If  $f^{(i)}(t_k) = \kappa + 1$  for some  $i > d$  and  $k \leq \ell - 1$ , then  $f(t_j) = f(t_k)$  for all  $k < j < \ell$ .

It is clear that  $\hat{\mathbb{D}}^{(N)}$  is countable.

### 5.1.3 A countable probabilistic structure

Let  $\hat{\Omega} := \mathbb{D}([0, T], \mathbb{R}^{d+K})$  and denote by  $\hat{S} = (\hat{S}_t)_{0 \leq t \leq T}$  the canonical process on the space  $\hat{\Omega}$ .

The set  $\hat{\mathbb{D}}^{(N)}$  is a countable subset of  $\hat{\Omega}$ . There exists a local martingale measure  $\hat{\mathbb{P}}^{(N)}$  on  $\hat{\Omega}$  which satisfies  $\hat{\mathbb{P}}^{(N)}[\hat{\mathbb{D}}^{(N)}] = 1$  and  $\hat{\mathbb{P}}^{(N)}[\{f\}] > 0$  for all  $f \in \hat{\mathbb{D}}^{(N)}$ . In fact, such a local martingale measure  $\hat{\mathbb{P}}^{(N)}$  on  $\hat{\mathbb{D}}^{(N)}$  can be constructed “by hand” as a continuous-time Markov chain with jump times decided independently of the jump positions. Let  $\hat{\mathbb{F}}^{(N)} := (\hat{\mathcal{F}}_t^{(N)})_{0 \leq t \leq T}$  be the filtration generated by the process  $\hat{S}$  and satisfying the usual assumptions (right-continuous and  $\hat{\mathbb{P}}^{(N)}$ -complete).

In the last section, we saw definitions of  $\hat{\tau}_k^{(N)}$  on  $\Omega$ . Here we extend their definitions to  $\bigcup_{N \in \mathbb{N}} \hat{\mathbb{D}}^{(N)}$ . Define the jump times by setting  $\hat{\tau}_0(\hat{S}) = 0$  and for  $k > 0$ ,

$$\hat{\tau}_k(\hat{S}) = \inf\{t > \hat{\tau}_{k-1}(\hat{S}) : \hat{S}_t \neq \hat{S}_{t-}\} \wedge T. \tag{5.5}$$

Next we introduce the random time before  $T$ ,

$$m(\hat{S}) := \min\{k : \hat{\tau}_k(\hat{S}) = T\}.$$

Observe that for  $S \in \Omega$ , we have  $\hat{F}^{(N)}(S) \in \hat{\mathbb{D}}^{(N)}$ ,  $\hat{\tau}_k(\hat{F}^{(N)}(S)) = \hat{\tau}_k(S)$  for all  $k$  and  $m(\hat{F}^{(N)}(S)) = m^{(N)}(S)$ . It follows that the definitions are consistent. In this context, a trading strategy  $(\hat{\gamma}_t)_{t=0}^T$  on the filtered probability space  $(\hat{\Omega}, \hat{\mathbb{F}}^{(N)}, \hat{\mathbb{P}}^{(N)})$  is a predictable stochastic process. So  $\hat{\gamma}$  is a map from  $\mathbb{D}([0, T], \mathbb{R}^{d+K})$  to  $\mathcal{D}([0, T], \mathbb{R}^{d+K})$ . Now choose  $a \in \mathcal{D}([0, T], \mathbb{R}^{d+K})$  such that  $a \notin \hat{\gamma}(\hat{\mathbb{D}}^{(N)})$  and then define a mapping  $\phi : \mathbb{D}([0, T], \mathbb{R}^{d+K}) \rightarrow \mathcal{D}([0, T], \mathbb{R}^{d+K})$  by  $\phi(\hat{S}) = \hat{\gamma}(\hat{S})$  if  $\hat{S} \in \hat{\mathbb{D}}^{(N)}$ , and equal to  $a$  otherwise. Since  $\hat{\mathbb{P}}^{(N)}$  has full support on  $\hat{\mathbb{D}}^{(N)}$ , we get  $\hat{\gamma} = \phi(\hat{S})$   $\hat{\mathbb{P}}^{(N)}$ -a.s. In particular, for any  $A$  that is a Borel-measurable subset of  $\mathbb{R}^{d+K}$ , the symmetric difference of  $\{\hat{\gamma}_t \in A\}$  and  $\{\phi(\hat{S})_t \in A\}$  is a nullset for  $\hat{\mathbb{P}}^{(N)}$ . Thus  $\phi$  is a predictable map. Furthermore, since  $\hat{\mathbb{P}}^{(N)}$  charges all elements in  $\hat{\mathbb{D}}^{(N)}$ , for any  $v, \tilde{v} \in \hat{\mathbb{D}}^{(N)}$  and  $t \in [0, T]$ ,

$$v_u = \tilde{v}_u, \forall u \in [0, t] \implies \phi(v)_t = \phi(\tilde{v})_t. \tag{5.6}$$

In the sequel, we always consider the above version  $\phi(\hat{S})$  of a predictable process  $\hat{\gamma}$ .

We now formally define the probabilistic superreplication problem and later build a connection between the probabilistic superreplication problem on the discretised space and the pathwise discretised robust hedging problem. For the rest of the section, we write  $\int_{t_1}^{t_2}$  to mean  $\int_{(t_1, t_2]}$ .

As  $G$  is defined only on  $\Omega$ , to consider paths in  $\hat{\Omega}$ , we need to extend the domain of  $G$  to  $\hat{\Omega}$ . For most financial contracts, the extension is natural. However, we pursue a general approach here. We first define a projection  $\natural : \hat{\Omega} \rightarrow \mathcal{C}([0, T], \mathbb{R}^{d+K})$  by

$$\natural(\hat{S}) = \begin{cases} \hat{S}, & \text{if } \hat{S} \text{ is continuous,} \\ \sum_{k=0}^{m(\hat{S})-1} \left( \frac{\hat{S}_{\hat{\tau}_{k+1}} - \hat{S}_{\hat{\tau}_k}}{\hat{\tau}_{k+1} - \hat{\tau}_k} (t - \hat{\tau}_k) + \hat{S}_{\hat{\tau}_k} \right) \mathbb{1}_{[\hat{\tau}_k, \hat{\tau}_{k+1})}(t), & \text{if } \hat{S} \in \bigcup_{N \in \mathbb{N}} \hat{\mathbb{D}}^{(N)}, \\ \omega^1, & \text{otherwise,} \end{cases}$$

where  $\omega^1$  is the constant path equal to 1. Put differently, when  $\hat{S} \in \bigcup_{N \in \mathbb{N}} \hat{\mathbb{D}}^{(N)}$ ,  $\lrcorner(\hat{S})$  is the linear interpolation of the points

$$\left( (\hat{\tau}_0(\hat{S}), \hat{S}_{\hat{\tau}_0(\hat{S})}), \dots, (\hat{\tau}_m(\hat{S}), \hat{S}_{\hat{\tau}_m(\hat{S})}) \right).$$

We then can define  $\hat{G} : \hat{\Omega} \rightarrow \Omega$  using the projection  $\lrcorner$  by  $\hat{G}(\hat{S}) = G(\lrcorner(\hat{S}) \vee 0)$ , where  $\hat{S} \vee 0 := ((\hat{S}_t^{(1)} \vee 0, \dots, \hat{S}_t^{(d+K)} \vee 0))_{0 \leq t \leq T}$  for any  $S \in \hat{\Omega}$ . Note that  $G$  and  $\hat{G}$  are equal on  $\Omega$ . In addition, for every  $N \in \mathbb{N}$  and  $\hat{S} \in \hat{\mathbb{D}}^{(N)}$ , we have

$$\|\lrcorner(\hat{S}) - \hat{S}\| \leq 2^{-N+1}. \tag{5.7}$$

Therefore, we can deduce that

$$\begin{aligned} \|\lrcorner(\hat{F}(S)) \vee 0 - S\| &\leq \|\lrcorner(\hat{F}(S)) \vee 0 - \hat{F}(S) \vee 0\| + \|\hat{F}(S) \vee 0 - S\| \\ &\leq 2^{-N+1} + 2^{-N+3}, \quad \forall S \in \Omega, \end{aligned} \tag{5.8}$$

where the last inequality follows from (5.4) and (5.7). Similarly, for each  $D \in \mathbb{N}$ , we define  $\hat{m}^{(D)} : \hat{\Omega} \rightarrow \mathbb{N}$  by  $\hat{m}^{(D)}(\hat{S}) = m^{(D)}(\lrcorner(\hat{S}) \vee 0)$ . Then by Remark 4.2 and (5.8), when  $N$  is sufficiently large,

$$\hat{m}^{(D-2)}(\hat{F}^{(N)}(S)) \leq m^{(D)}(S), \quad \forall S \in \Omega. \tag{5.9}$$

**Definition 5.5**  $\hat{\gamma} : \hat{\Omega} \rightarrow \mathcal{D}([0, T], \mathbb{R}^{d+K})$  is  $\hat{\mathbb{P}}^{(N)}$ -admissible if  $\hat{\gamma}$  is predictable and bounded by  $N$ , and the stochastic integral  $(\int_0^t \hat{\gamma}_u(\hat{S}) d\hat{S}_u)_{0 \leq t \leq T}$ , which is well defined under  $\hat{\mathbb{P}}^{(N)}$ , satisfies that there is some  $M > 0$  such that

$$\int_0^t \hat{\gamma}_u(\hat{S}) d\hat{S}_u \geq -M \quad \hat{\mathbb{P}}^{(N)}\text{-a.s.}, t \in [0, T].$$

An admissible strategy  $\hat{\gamma}$  is said to  $\hat{\mathbb{P}}^{(N)}$ -superreplicate  $\hat{G}$  if

$$\int_0^T \hat{\gamma}_u(\hat{S}) d\hat{S}_u \geq \hat{G}(\hat{S}) \quad \hat{\mathbb{P}}^{(N)}\text{-a.s.}$$

The *superreplication cost* of  $\hat{G}$  is defined as

$$\hat{V}^{(N)} := \inf\{x : \exists \hat{\gamma} \text{ which is } \hat{\mathbb{P}}^{(N)}\text{-admissible and } \hat{\mathbb{P}}^{(N)}\text{-superreplicates } \hat{G} - x\}.$$

Similarly to [25], we can now connect the probabilistic superhedging problem and the discretised robust hedging problem.

**Definition 5.6** Given a predictable stochastic process  $(\hat{\gamma}_t)_{0 \leq t \leq T}$  on  $(\hat{\Omega}, \hat{\mathbb{F}}^{(N)}, \hat{\mathbb{P}}^{(N)})$ , we define  $\gamma^{(N)} : \Omega \rightarrow \mathbb{D}([0, T], \mathbb{R}^{d+K})$  by

$$\gamma_t^{(N)}(S) := \sum_{k=0}^{m-1} \hat{\gamma}_{\hat{\tau}_k}(\hat{F}^{(N)}(S)) \mathbb{1}_{(\tau_k, \tau_{k+1}]}(t), \tag{5.10}$$

where  $\tau_k = \tau_k^{(N)}(S)$ ,  $m = m^{(N)}(S)$  are given in Definition 4.1,  $\hat{F}^{(N)}$  in (5.3),  $\hat{\tau}_k = \hat{\tau}_k(\hat{F}^{(N)}(S))$  in (5.5), and we recall that  $m^{(N)}(S) = m(\hat{F}^{(N)}(S))$ .

**Lemma 5.7** *For any admissible strategy  $\hat{\gamma}$  in the sense of Definition 5.5,  $\gamma^{(N)}$  defined in (5.10) is  $\mathbb{F}$ -predictable.*

*Proof* We first show that if we equip  $\Omega$  and  $\mathbb{D}([0, T], \mathbb{R}^{d+K})$  with the respective  $\sigma$ -algebras  $\mathcal{F}_T$  and  $\hat{\mathcal{F}}_T^{(N)}$ , the function  $\hat{F}^{(N)} : \Omega \rightarrow \mathbb{D}([0, T], \mathbb{R}^{d+K})$  is measurable. Since  $\hat{\mathbb{P}}^{(N)}$  has full support on  $\hat{\mathbb{D}}^{(N)}$ , for any  $A \in \hat{\mathcal{F}}_T^{(N)}$ ,

$$\{\hat{F}^{(N)} \in A\} = \bigcup_{\hat{S} \in \hat{\mathbb{D}}^{(N)} \cap A} \{S \in \Omega : \hat{F}^{(N)}(S) = \hat{S}\}.$$

It is clear from the construction that  $\hat{\tau}_k^{(N)}$ ,  $m^{(N)}$  and  $\zeta^{(N+k)}$  are all  $\mathcal{F}_T$ -measurable, and we note that for any  $\hat{S} \in \hat{\mathbb{D}}^{(N)}$ ,

$$\{S \in \Omega : \hat{F}^{(N)}(S) = \hat{S}\} = \{S \in \Omega : m^{(N)} = m, \hat{\tau}_k^{(N)} = t_k, \zeta^{(N+i)} = s_k, \forall k < m\}$$

for some  $m, t_k, s_k$ . Therefore, we can conclude that  $\hat{F}^{(N)}$  has the desired measurability.

To prove that  $\gamma^{(N)}$  is  $\mathbb{F}$ -predictable, we need to show that  $\hat{\gamma}_{\hat{\tau}_k} \circ \hat{F}^{(N)}$  is  $\mathcal{F}_{\tau_k}$ -measurable. Galmarino’s test, see Dellacherie and Meyer [22, Theorem IV.100], states that given any  $\mathcal{F}_T$ -measurable random variable  $\phi : \Omega \rightarrow \mathbb{R}^{d+K}$  and any  $\mathbb{F}$ -stopping time  $\tau$ ,  $\phi$  is  $\mathcal{F}_\tau$ -measurable if and only if

$$\forall v, \omega \in \Omega : \quad v_u = \omega_u, \forall u \in [0, \tau(v)] \implies \phi(v) = \phi(\omega).$$

It follows from the definition of  $\hat{F}^{(N)}$  that for such  $v, \omega$  and  $\tau = \tau_k$ , we have  $\hat{F}_u^{(N)}(\omega) = \hat{F}_u^{(N)}(v), \forall u \in [0, \hat{\tau}_k)$ . Hence by (5.6),  $\hat{\gamma}_{\hat{\tau}_k}(\hat{F}^{(N)}(\omega)) = \hat{\gamma}_{\hat{\tau}_k}(\hat{F}^{(N)}(v))$ . Therefore Galmarino’s test implies that  $\gamma^{(N)}$  defined in (5.10) is  $\mathbb{F}$ -predictable.  $\square$

The following result is crucial. It states that the probabilistic superreplication value is asymptotically larger than the value of the discretised robust hedging problem. Recall that  $\lambda_{\mathcal{I}}(\omega) := \inf_{v \in \mathcal{I}} \|\omega - v\| \wedge 1$ .

**Proposition 5.8** *For uniformly continuous and bounded  $G, \alpha, \beta \geq 0$  and  $D \in \mathbb{N}$ , we have*

$$\begin{aligned} & \liminf_{N \rightarrow \infty} \mathbf{V}_{\mathcal{I}}^{(N)}(G(\mathbb{S}) - \beta \sqrt{m^{(D)}(\mathbb{S})} \wedge \alpha) \\ & \leq \liminf_{N \rightarrow \infty} \hat{\mathbf{V}}^{(N)}(\hat{G}(\hat{\mathbb{S}}) - \beta \sqrt{\hat{m}^{(D-2)}(\hat{\mathbb{S}})} \wedge \alpha - N \lambda_{\mathcal{I}}(\hat{\mathbb{S}})). \end{aligned} \tag{5.11}$$

*Proof* Fix  $N \geq 6$ . Let  $f_e : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be a modulus of continuity for  $G$  so that  $\lim_{x \rightarrow 0} f_e(x) = 0$ . Define  $G^{(N)} : \Omega \rightarrow \mathbb{R}$  as

$$G^{(N)}(S) := \hat{G}(S) - f_e(2^{-N+4}) - \frac{14(d+K)N}{2^N}.$$

Note that

$$\begin{aligned} \mathbf{V}_{\mathcal{I}}^{(N)}(G - \beta \sqrt{m^{(D)}} \wedge \alpha) &= \mathbf{V}_{\mathcal{I}}^{(N)}(G^{(N)} - \beta \sqrt{m^{(D)}} \wedge \alpha) \\ & \quad + f_e(2^{-N+4}) + \frac{14(d+K)N}{2^N}. \end{aligned}$$

Hence, to show (5.11), it suffices to show that

$$\mathbf{V}_{\mathcal{I}}^{(N)}(G^{(N)} - \beta\sqrt{m^{(D)}} \wedge \alpha) \leq \hat{\mathbf{V}}^{(N)}(\hat{G} - \beta\sqrt{\hat{m}^{(D-2)}} \wedge \alpha - N\lambda_{\mathcal{I}}). \tag{5.12}$$

The rest of the proof is structured to establish (5.12). Given a probabilistic semi-static portfolio  $\hat{\gamma}$  which superreplicates  $\hat{G} - \beta\sqrt{\hat{m}^{(D-2)}} \wedge \alpha - N\lambda_{\mathcal{I}} - x$ , we argue that the lifted trading strategy  $\gamma^{(N)}$  superreplicates  $G^{(N)} - \beta\sqrt{m^{(D)}} \wedge \alpha - x$  on  $\mathcal{I}$ . To simplify notations, **throughout the rest of the proof**, we fix  $S \in \mathcal{I}$  and write  $\hat{F} := \hat{F}^{(N)}(S)$ .

*Superreplication.* We first notice that for any  $j < m - 1$ ,

$$\begin{aligned} |(S_{\tau_{j+1}} - S_{\tau_j}) - (\hat{F}_{\hat{\tau}_j} - \hat{F}_{\hat{\tau}_{j-1}})| &\leq |S_{\tau_{j+1}} - \hat{F}_{\hat{\tau}_j}| + |S_{\tau_j} - \hat{F}_{\hat{\tau}_{j-1}}| \\ &\leq \frac{1}{2^{N+j+1}} + \frac{1}{2^{N+j}} = \frac{3}{2^{N+j+1}}. \end{aligned}$$

It follows that for any  $k < m$ ,

$$\begin{aligned} &\left| \int_0^{\tau_k} \gamma_u^{(N)}(S) dS_u - \int_0^{\hat{\tau}_k} \hat{\gamma}_u(\hat{F}) d\hat{F}_u \right| \\ &\leq \left| \sum_{j=0}^{k-1} \hat{\gamma}_{\hat{\tau}_j}(\hat{F})(S_{\tau_{j+1}} - S_{\tau_j}) - \sum_{j=0}^{k-1} \hat{\gamma}_{\hat{\tau}_{j+1}}(\hat{F})(\hat{F}_{\hat{\tau}_{j+1}} - \hat{F}_{\hat{\tau}_j}) \right| \\ &\leq \sum_{j=0}^{k-2} |\hat{\gamma}_{\hat{\tau}_{j+1}}(\hat{F})(S_{\tau_{j+2}} - S_{\tau_{j+1}}) - (\hat{F}_{\hat{\tau}_{j+1}} - \hat{F}_{\hat{\tau}_j})| + \frac{2(d+K)N}{2^{N-1}} \\ &\leq \sum_{j=0}^{\infty} \frac{N(d+K)}{2^{N+j+2}} + \frac{2(d+K)N}{2^{N-1}} \leq \frac{5(d+K)N}{2^N}. \end{aligned} \tag{5.13}$$

In addition,

$$\begin{aligned} &\left| \int_{\tau_{m-1}}^T \gamma_u^{(N)}(S) dS_u - \int_{\hat{\tau}_{m-1}}^T \hat{\gamma}_u(\hat{F}) d\hat{F}_u \right| \\ &= |\hat{\gamma}_{\hat{\tau}_{m-1}}(\hat{F})(S_T - S_{\tau_{m-1}}) - \hat{\gamma}_{\hat{\tau}_m}(\hat{F})(\hat{F}_{\hat{\tau}_m} - \hat{F}_{\hat{\tau}_{m-1}})| \leq \frac{N(d+K)}{2^N}. \end{aligned} \tag{5.14}$$

Hence,

$$\begin{aligned} x + \int_0^T \gamma_u^{(N)}(S) dS_u &\geq x + \int_0^T \hat{\gamma}_u(\hat{F}) d\hat{F}_u - \frac{5(d+K)N}{2^N} - \frac{(d+K)N}{2^N} \\ &\geq \hat{G}(\hat{F}) - \beta\sqrt{\hat{m}^{(D-2)}(\hat{F})} \wedge \alpha - N\lambda_{\mathcal{I}}(\hat{F}) - \frac{6(d+K)N}{2^N} \\ &\geq \hat{G}(\hat{F}) - \beta\sqrt{\hat{m}^{(D-2)}(\hat{F})} \wedge \alpha - N/2^{N-3} - \frac{6(d+K)N}{2^N} \\ &\geq G(S) - \beta\sqrt{m^{(D)}(S)} \wedge \alpha - f_e(2^{-N+4}) - \frac{14(d+K)N}{2^N} \\ &= G^{(N)}(S), \end{aligned}$$



where the second inequality follows from the superreplicating property of  $\hat{\gamma}$  and the fact that  $\hat{\mathbb{P}}^{(N)}[\{f\}] > 0, \forall f \in \hat{\mathbb{D}}^{(N)}$ , the third inequality is justified by (5.4), and the last inequality is due to (5.8) and (5.9).

*Admissibility.* Now, for a given  $t < T$ , let  $k < m$  be the largest integer so that  $\tau_k(S) \leq t$ . It follows from (5.13) and (5.14) that

$$\begin{aligned} \int_0^t \gamma_u^{(N)}(S) dS_u &= \int_0^{\tau_k} \gamma_u^{(N)}(S) dS_u + \int_{\tau_k}^t \gamma_u^{(N)}(S) dS_u \\ &\geq \int_0^{\hat{\tau}_k} \hat{\gamma}_u(\hat{F}) d\hat{F}_u - \frac{5(d+K)N}{2^N} - N(d+K) \max_i |S_t^{(i)} - S_{\tau_k}^{(i)}| \\ &\geq -M - \frac{6(d+K)N}{2^N}, \end{aligned}$$

where the last inequality follows from the admissibility of  $\hat{\gamma}$  and again the fact that  $\hat{\mathbb{P}}^{(N)}[\{f\}] > 0, \forall f \in \hat{\mathbb{D}}^{(N)}$ . Hence  $\gamma^{(N)}$  is admissible. □

### 5.2 Duality for the discretised problems

**Definition 5.9** Let  $\hat{\Pi}^{(N)}$  be the set of all probability measures  $\hat{\mathbb{Q}}$  which are equivalent to  $\hat{\mathbb{P}}^{(N)}$ . For any  $\kappa \geq 0$ , denote by  $\hat{\mathbb{M}}_{\mathcal{I}}^{(N)}(\kappa)$  the set of all probability measures  $\hat{\mathbb{Q}} \in \hat{\Pi}^{(N)}$  such that

$$\hat{\mathbb{Q}}\left[\left\{\omega \in \hat{\Omega} : \inf_{\nu \in \mathcal{I}} \|\hat{\mathbb{S}}(\omega) - \nu\| \geq 1/N\right\}\right] \leq \frac{\kappa}{N}$$

and

$$\mathbb{E}_{\hat{\mathbb{Q}}}\left[\sum_{k=1}^{m(\hat{\mathbb{S}})} \sum_{i=1}^{d+K} \left|\mathbb{E}_{\hat{\mathbb{Q}}}\left[\hat{\mathbb{S}}_{\hat{\tau}_k}^{(i)} \mid \hat{\mathcal{F}}_{\hat{\tau}_{k-1}}^{(N)}\right] - \hat{\mathbb{S}}_{\hat{\tau}_{k-1}}^{(i)}\right|\right] \leq \frac{\kappa}{N},$$

where  $\hat{\tau}_k = \hat{\tau}_k(\hat{\mathbb{S}})$  and  $m = m(\hat{\mathbb{S}})$  are defined in (5.5).

**Lemma 5.10** Let  $\kappa > 1$  and suppose  $\hat{G}$  is bounded by  $\kappa - 1$  and  $\mathcal{M}_{\mathcal{I}} \neq \emptyset$ . Then there are at most finitely many  $N \in \mathbb{N}$  such that  $\hat{\mathbb{M}}_{\mathcal{I}}^{(N)}(2\kappa) = \emptyset$ , and we have

$$\liminf_{N \rightarrow \infty} \hat{\mathbb{V}}^{(N)}(\hat{G}(\hat{\mathbb{S}}) - N\lambda_{\mathcal{I}}(\hat{\mathbb{S}})) \leq \liminf_{N \rightarrow \infty} \sup_{\hat{\mathbb{Q}} \in \hat{\mathbb{M}}_{\mathcal{I}}^{(N)}(2\kappa)} \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}})].$$

*Proof* For any  $\hat{\mathbb{Q}} \in \hat{\Pi}^{(N)}$ , the support of  $\hat{\mathbb{Q}}$  is  $\hat{\mathbb{D}}^{(N)}$  whose elements are piecewise constant. Therefore, the canonical process  $\hat{\mathbb{S}}$  is a semimartingale under  $\hat{\mathbb{Q}}$ . Moreover, it has the decomposition  $\hat{\mathbb{S}} = \hat{M}^{\hat{\mathbb{Q}}} + \hat{A}^{\hat{\mathbb{Q}}}$ , where

$$\begin{aligned} \hat{A}_t^{\hat{\mathbb{Q}}} &= \sum_{k=1}^{m(\hat{\mathbb{S}})} \left(\mathbb{E}_{\hat{\mathbb{Q}}}\left[\hat{\mathbb{S}}_{\hat{\tau}_k} \mid \hat{\mathcal{F}}_{\hat{\tau}_{k-1}}^{(N)}\right] - \hat{\mathbb{S}}_{\hat{\tau}_{k-1}}\right) \mathbb{1}_{[\hat{\tau}_k, \hat{\tau}_{k+1})}(t), \quad t < T, \\ \hat{A}_T^{\hat{\mathbb{Q}}} &:= \lim_{t \uparrow T} \hat{A}_t^{\hat{\mathbb{Q}}}, \end{aligned}$$

is a predictable process of bounded variation and  $\hat{M}^{\hat{Q}}$  is a martingale under  $\hat{Q}$ . Then similarly to Dolinsky and Soner [26], it follows from Example 2.3 and Proposition 4.1 in Föllmer and Kramkov [29] that

$$\begin{aligned} & \hat{V}^{(N)}(\hat{G}(\hat{S}) - N\lambda_{\mathcal{I}}(\hat{S})) \\ &= \sup_{\hat{Q} \in \hat{\Pi}^{(N)}} \mathbb{E}_{\hat{Q}} \left[ \hat{G}(\hat{S}) - N\lambda_{\mathcal{I}}(\hat{S}) - N \sum_{k=1}^{m(\hat{S})} \sum_{i=1}^{d+K} |\mathbb{E}_{\hat{Q}}[\hat{S}_{\hat{\tau}_k}^{(i)} | \hat{\mathcal{F}}_{\hat{\tau}_{k-1}}^{(N)}] - \hat{S}_{\hat{\tau}_{k-1}}^{(i)}| \right]. \end{aligned} \tag{5.15}$$

By Proposition 5.8,

$$\liminf_{N \rightarrow \infty} \hat{V}^{(N)}(\hat{G}(\hat{S}) - N\lambda_{\mathcal{I}}(\hat{S})) \geq \liminf_{N \rightarrow \infty} \mathbf{V}_{\mathcal{I}}^{(N)}(G) \geq \mathbf{P}_{\mathcal{I}}(G) > -\kappa.$$

Then, in (5.15), it suffices to consider the supremum over  $\hat{\mathbb{M}}_{\mathcal{I}}^{(N)}(2\kappa)$ . In particular,  $\hat{\mathbb{M}}_{\mathcal{I}}^{(N)}(2\kappa) \neq \emptyset$  for  $N$  large enough. □

### 5.3 Discretisation of the primal

Next, we show that we can lift any measure in  $\hat{\mathbb{M}}_{\mathcal{I}}^{(N)}(c)$  to a continuous martingale measure in  $\mathcal{M}_{\mathcal{I}}$  such that the difference of the expected value of  $G$  under this continuous martingale measure and the expected value of  $\hat{G}$  under the original measure is within a bounded error, which goes to zero as  $N \rightarrow \infty$ . Through this, we asymptotically connect the primal problems on the discretised space to the approximation of the primal problems on the space of continuous functions.

**Proposition 5.11** *Under the assumptions of Theorem 5.1, if  $G$  and all  $X_i^{(c)}/\mathcal{P}(X_i^{(c)})$  are bounded by  $\kappa - 1$  for some  $\kappa \geq 1$ , then for any  $\alpha, \beta \geq 0, D \in \mathbb{N}$ ,*

$$\begin{aligned} & \limsup_{N \rightarrow \infty} \sup_{\hat{Q} \in \hat{\mathbb{M}}_{\mathcal{I}}^{(N)}(2\kappa + 2\alpha)} \mathbb{E}_{\hat{Q}}[\hat{G}(\hat{S}) - \beta \sqrt{\hat{m}^{(D)}(\hat{S})} \wedge \alpha] \\ & \leq \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S}) - \beta \sqrt{\hat{m}^{(D-2)}(\hat{S})} \wedge \alpha]. \end{aligned} \tag{5.16}$$

*Proof* Let  $f_e : \mathbb{R}_+^{d+K} \rightarrow \mathbb{R}_+$  be a modulus of continuity of  $G$ , i.e.,

$$|G(\omega) - G(v)| \leq f_e(|\omega - v|) \quad \text{for any } \omega, v \in \Omega$$

and  $\lim_{x \searrow 0} f_e(x) = 0$ . Recall from Lemma 5.10 that  $\hat{\mathbb{M}}_{\mathcal{I}}^{(N)}(2\kappa + 2\alpha) \neq \emptyset$  for  $N$  large enough. Hence to show (5.16), it suffices to prove that for any  $\hat{Q} \in \hat{\mathbb{M}}_{\mathcal{I}}^{(N)}(2\kappa + 2\alpha)$ ,

$$\mathbb{E}_{\hat{Q}}[\hat{G}(\hat{S}) - \beta \sqrt{\hat{m}^{(D)}(\hat{S})} \wedge \alpha] \leq \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S}) - \beta \sqrt{m^{(D-2)}(\mathbb{S})} \wedge \alpha] + g(1/N),$$

for some  $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that  $\lim_{x \searrow 0} g(x) = 0$ . We fix  $N$  and  $\hat{Q} \in \hat{\mathbb{M}}_{\mathcal{I}}^{(N)}(2\kappa + 2\alpha)$  and prove the above inequality in four steps.

*Step 1.* We first construct a semimartingale  $\hat{Z} = \hat{M} + \hat{A}$  on a Wiener space  $(\Omega^W, \mathcal{F}^W, P^W)$  such that

$$|\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{S})] - E^W[\hat{G}(\hat{Z})]| \leq \kappa 2^{-N+1} \tag{5.17}$$

and

$$P^W \left[ \left\{ \omega \in \Omega^W : \inf_{\nu \in \mathcal{I}} \|\hat{M}(\omega) + \hat{A}(\omega) - \nu\| \geq 1/N \right\} \right] \leq \frac{2\kappa + 2\alpha}{N} + 2^{-N}, \tag{5.18}$$

where  $\hat{M}$  is constructed from a martingale and both have piecewise constant paths.

Since the measure  $\hat{\mathbb{Q}}$  is supported on  $\hat{\mathbb{D}}^{(N)}$ , the canonical process  $\hat{S}$  is a pure jump process under  $\hat{\mathbb{Q}}$ , with a finite number of jumps  $\hat{\mathbb{Q}}$ -a.s. Consequently, there exists a deterministic positive integer  $m_0$  (depending on  $N$ ) such that

$$\hat{\mathbb{Q}}[m(\hat{S}) > m_0] < 2^{-N}. \tag{5.19}$$

It follows that

$$|\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{S})] - \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{S}^{\hat{\tau}_{m_0}})]| \leq \kappa 2^{-N+1}. \tag{5.20}$$

Notice that by the definition of  $\hat{\mathbb{D}}^{(N)}$ , the law of  $\hat{S}^{\hat{\tau}_{m_0}}$  under  $\hat{\mathbb{Q}}$  is also supported on  $\hat{\mathbb{D}}^{(N)}$ .

Let  $(\Omega^W, \mathcal{F}^W, P^W)$  be a complete probability space together with a standard  $(m_0 + 2)$ -dimensional Brownian motion  $\{W_t = (W_t^{(1)}, \dots, W_t^{(m_0+2)})\}_{t \geq 0}$  and let  $(\mathcal{F}_t^W)_{t \geq 0}$  be the  $P^W$ -completion of the natural filtration of  $W$ . With a small modification of Lemma 5.1 in Dolinsky and Soner [25], we can construct a sequence of stopping times (with respect to the Brownian filtration)  $\sigma_1 < \sigma_2 < \dots < \sigma_{m_0}$  together with  $\mathcal{F}_{\sigma_i}^W$ -measurable random variables  $Y_i, i = 1, \dots, m_0$  such that

$$\begin{aligned} & \mathcal{L}_{P^W}((\sigma_1, \dots, \sigma_{m_0}, Y_1, \dots, Y_{m_0})) \\ &= \mathcal{L}_{\hat{\mathbb{Q}}}((\hat{\tau}_1, \dots, \hat{\tau}_{m_0}, \hat{S}_{\hat{\tau}_1} - \hat{S}_{\hat{\tau}_0}, \dots, \hat{S}_{\hat{\tau}_{m_0}} - \hat{S}_{\hat{\tau}_{m_0-1}})); \end{aligned}$$

see Sect. A.3 for details. Define  $X_i$  as

$$X_i = E^W[Y_i | \mathcal{F}_{\sigma_{i-1}}^W \vee \sigma(\sigma_i)], \quad i = 1, \dots, m_0.$$

Note that by the definition of  $\hat{\mathbb{D}}^{(N)}$ , we have  $|Y_i| \leq 2^{-N}$  and hence also  $|X_i| \leq 2^{-N}$ . Also by the construction of the  $\sigma_i$  and  $Y_i$ , we have

$$E^W[Y_i | \mathcal{F}_{\sigma_{i-1}}^W \vee \sigma(\sigma_i)] = E^W[Y_i | \sigma_i, \mathbf{Y}_{i-1}],$$

where  $\sigma_i := (\sigma_1, \dots, \sigma_i), \mathbf{Y}_i := (Y_1, \dots, Y_i)$  and  $E^W$  is the expectation with respect to  $P^W$ . From these, we can construct a jump process  $(\hat{A}_t)_{0 \leq t \leq T}$  by

$$\hat{A}_t = \sum_{j=1}^{m_0} X_j \mathbb{1}_{[\sigma_j, T]}(t).$$

In particular, for  $k \leq m_0$ ,  $\hat{A}_{\sigma_k} = \sum_{j=1}^k X_j$ . Define a martingale  $(M_t)_{0 \leq t \leq T}$  via

$$M_t = 1 + E^W \left[ \sum_{j=1}^{m_0} (Y_j - X_j) \middle| \mathcal{F}_t^W \right], \quad t \in [0, T].$$

Since all Brownian martingales are continuous, so is  $M$ . Moreover, Brownian motion increments are independent and therefore

$$M_{\sigma_k} = 1 + \sum_{j=1}^k (Y_j - X_j) \quad P^W\text{-a.s., } k \leq m.$$

We now introduce a stochastic process  $(\hat{M}_t)_{0 \leq t \leq T}$  on the Brownian probability space by setting  $\hat{M}_t = M_{\sigma_k}$  for  $t \in [\sigma_k, \sigma_{k+1})$ ,  $k < m_0$ , and  $\hat{M}_t = \hat{M}_{\sigma_{m_0}}$  for  $t \in [\sigma_{m_0}, T]$ . Note that as  $|Y_i - X_i| \leq 2^{-N+1}$ , for any  $k \leq m_0$  and  $t \leq T$ , we have

$$\begin{aligned} |\hat{M}_{t \wedge \sigma_{k+1} \vee \sigma_k} - M_{t \wedge \sigma_{k+1} \vee \sigma_k}| &= \left| \sum_{j=k+1}^{m_0} E^W [(Y_j - X_j) | \mathcal{F}_{t \wedge \sigma_{k+1} \vee \sigma_k}^W] \right| \\ &= \left| \sum_{j=k+2}^{m_0} E^W [E^W [(Y_j - X_j) | \mathcal{F}_{\sigma_{j-1}}^W \vee \sigma(\sigma_j)] | \mathcal{F}_{t \wedge \sigma_{k+1} \vee \sigma_k}^W] \right. \\ &\quad \left. + E^W [Y_{k+1} - X_{k+1} | \mathcal{F}_{t \wedge \sigma_{k+1} \vee \sigma_k}^W] \right| \\ &= |E^W [Y_{k+1} - X_{k+1} | \mathcal{F}_{t \wedge \sigma_{k+1} \vee \sigma_k}^W]| \\ &\leq E^W [|Y_{k+1} - X_{k+1}| | \mathcal{F}_{t \wedge \sigma_{k+1} \vee \sigma_k}^W] \leq 2^{-N+1} \end{aligned}$$

and hence

$$\|\hat{M} - M\| < 2^{-N+2}. \tag{5.21}$$

We also notice that  $\hat{Z} = \hat{M} + \hat{A}$  satisfies  $\hat{Z}_0 = \hat{S}_0$  and

$$\begin{aligned} \mathcal{L}_{PW}((\sigma_1, \dots, \sigma_{m_0}, Z_{\sigma_1} - Z_0, \dots, Z_{\sigma_{m_0}} - Z_{\sigma_{m_0-1}})) \\ = \mathcal{L}_{\hat{Q}}((\hat{\tau}_1, \dots, \hat{\tau}_{m_0}, \hat{S}_{\hat{\tau}_1} - \hat{S}_{\hat{\tau}_0}, \dots, \hat{S}_{\hat{\tau}_{m_0}} - \hat{S}_{\hat{\tau}_{m_0-1}})). \end{aligned}$$

It follows that

$$E^W[\hat{G}(\hat{Z})] = \mathbb{E}_{\hat{Q}}[\hat{G}(\hat{S}^{\tau_{m_0}})].$$

In particular, by (5.20), we see that (5.17) holds, and also by (5.19) and the definition of  $\hat{M}$  and  $\hat{A}$ , (5.18) holds.

*Step 2.* We shall shortly construct a continuous martingale  $M^{\theta_0}$  from  $M$  such that  $M^{\theta_0}$  is bounded below by  $-2^{-N+2} - N^{-\frac{1}{2}}$  and

$$|E^W[\hat{G}(M^{\theta_0})] - \mathbb{E}_{\hat{\mathbb{Q}}}[G(\hat{S})]| \leq c^2 N^{-\frac{1}{2}} + 2f_e(N^{-\frac{1}{2}} + 2^{-N+2}) + 2^{-N}.$$

As the law of  $\hat{Z}$  under  $P^W$  is the same as that of  $\hat{S}^{m_0}$  under  $\hat{\mathbb{Q}}$ , it follows from the fact that  $\hat{\mathbb{Q}}$  is supported on  $\hat{\mathbb{D}}^{(N)}$  and any  $f \in \hat{\mathbb{D}}^{(N)}$  is above  $-2^{-N+3}$  that

$$\hat{Z} \geq -2^{-N+3} \quad P^W\text{-a.s.} \tag{5.22}$$

By combining this with (5.7) and (5.21), we can deduce that

$$\begin{aligned} \|\hat{\Delta}(\hat{Z}) - M\| &\leq \|\hat{\Delta}(\hat{Z}) - \hat{Z} \vee 0\| + \|\hat{Z} \vee 0 - \hat{Z}\| + \|\hat{Z} - M\| \\ &\leq 2^{-N+1} + 2^{N+3} + \|\hat{M} - M\| + \|\hat{A}\| \\ &\leq 2^{-N+4} + N^{-\frac{1}{2}}, \quad \text{whenever } \max_{1 \leq i \leq d+K} \sum_{k=1}^{m_0} |X_k^{(i)}| \leq N^{-\frac{1}{2}}. \end{aligned}$$

It follows that

$$\begin{aligned} |\hat{G}(M) - \hat{G}(\hat{Z})| &= |G(M \vee 0) - G(\hat{\Delta}(\hat{Z}) \vee 0)| \\ &\leq f_e(2^{-N+4} + N^{-\frac{1}{2}}), \quad \text{whenever } \max_{1 \leq i \leq d+K} \sum_{k=1}^{m_0} |X_k^{(i)}| \leq N^{-\frac{1}{2}}, \end{aligned}$$

where we use the fact that  $\|\hat{\Delta}(\hat{Z}) \vee 0 - M \vee 0\| \leq \|\hat{\Delta}(\hat{Z}) - M\|$ . Hence, since  $\hat{G}$  is bounded by  $\kappa$ ,

$$\begin{aligned} |E^W[\hat{G}(M)] - E^W[\hat{G}(\hat{Z})]| &\leq f_e(2^{-N+4} + N^{-\frac{1}{2}}) \\ &\quad + 2\kappa P^W \left[ \max_{1 \leq i \leq d+K} \sum_{k=1}^{m_0} |X_k^{(i)}| > N^{-\frac{1}{2}} \right]. \end{aligned}$$

Note that with the notation of the proof in Sect. A.3 in the Appendix,

$$\begin{aligned} X_k &= E^W[Y_k | \sigma_k, \mathbf{Y}_{k-1}] \stackrel{(d)}{=} \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{S}_{\hat{\tau}_k} - \hat{S}_{\hat{\tau}_{k-1}} | \hat{\mathbf{r}}_k, \Delta \hat{S}_{\hat{\tau}_{k-1}}] \\ &= \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{S}_{\hat{\tau}_k} - \hat{S}_{\hat{\tau}_{k-1}} | \hat{\mathcal{F}}_{\hat{\tau}_{k-1}}^{(N)}], \end{aligned}$$

where  $\Delta \hat{S}_k = \hat{S}_{\hat{\tau}_k} - \hat{S}_{\hat{\tau}_{k-1}} = \hat{S}_{\hat{\tau}_k} - \hat{S}_{\hat{\tau}_{k-1}}$  for  $k \leq m_0$  and hence

$$E^W \left[ \sum_{i=1}^{d+K} \sum_{k=1}^{m_0} |X_k^{(i)}| \right] = \mathbb{E}_{\hat{\mathbb{Q}}} \left[ \sum_{k=1}^{m_0} \sum_{i=1}^{d+K} |\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{S}_{\hat{\tau}_k}^{(i)} | \hat{\mathcal{F}}_{\hat{\tau}_{k-1}}^{(N)}] - \hat{S}_{\hat{\tau}_{k-1}}^{(i)}| \right].$$

By Markov’s inequality and the definition of  $\hat{\mathbb{M}}_{\mathcal{I}}^{(N)}(2\kappa + 2\alpha)$ , we have

$$\begin{aligned}
 P^W \left[ \sum_{i=1}^{d+K} \sum_{k=1}^{m_0} |X_k^{(i)}| > N^{-\frac{1}{2}} \right] &\leq \sqrt{N} E^W \left[ \sum_{i=1}^{d+K} \sum_{k=1}^{m_0} |X_k^{(i)}| \right] \\
 &\leq \sqrt{N} \mathbb{E}_{\hat{\mathbb{Q}}} \left[ \sum_{k=1}^{m(\hat{\mathbb{S}})} \sum_{i=1}^{d+K} \left| \mathbb{E}_{\hat{\mathbb{Q}}} \left[ \hat{\mathbb{S}}_{\hat{t}_k}^{(i)} \mid \hat{\mathcal{F}}_{\hat{t}_{k-1}}^{(N)} \right] - \hat{\mathbb{S}}_{\hat{t}_{k-1}}^{(i)} \right| \right] \\
 &\leq 2(\kappa + \alpha) N^{-\frac{1}{2}}. \tag{5.23}
 \end{aligned}$$

Therefore, we have

$$|E^W[\hat{G}(M)] - E^W[\hat{G}(\hat{Z})]| \leq f_e(2^{-N+4} + N^{-\frac{1}{2}}) + 4(\kappa + \alpha)^2 N^{-\frac{1}{2}}. \tag{5.24}$$

By (5.21)–(5.23),

$$\begin{aligned}
 P^W \left[ \inf_{0 \leq t \leq T} \min_{1 \leq i \leq d+K} M_t^{(i)} > -2^{-N+4} - N^{-\frac{1}{2}} \text{ and} \right. \\
 \left. \max_{d \leq i \leq d+K} \|M^{(i)}\| < \kappa + 1 + 2^{-N+2} + N^{-\frac{1}{2}} \right] \geq 1 - 2\kappa N^{-\frac{1}{2}}.
 \end{aligned}$$

Hence the stopped process  $M^{\theta_0}$ , with

$$\begin{aligned}
 \theta_0 := \inf \left\{ t \geq 0 : \min_{1 \leq i \leq d+K} M_t^{(i)} \leq -2^{-N+4} - N^{-\frac{1}{2}} \text{ or} \right. \\
 \left. \max_{d \leq i \leq d+K} \|M^{(i)}\| \geq \kappa + 1 + 2^{-N+2} + N^{-\frac{1}{2}} \right\},
 \end{aligned}$$

satisfies

$$|E^W[\hat{G}(M)] - E^W[\hat{G}(M^{\theta_0})]| \leq 4(\kappa + \alpha)^2 N^{-\frac{1}{2}}. \tag{5.25}$$

By (5.20), (5.24) and (5.25), it follows that

$$\begin{aligned}
 &|E^W[\hat{G}(M^{\theta_0})] - \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}})]| \\
 &\leq |E^W[\hat{G}(M^{\theta_0})] - E^W[\hat{G}(M)]| \\
 &\quad + |E^W[\hat{G}(M)] - E^W[\hat{G}(\hat{Z})]| + |\mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}}^{\hat{t}_{m_0}})] - \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}})]| \\
 &\leq 4(\kappa + \alpha)^2 N^{-\frac{1}{2}} + 4(\kappa + \alpha)^2 N^{-\frac{1}{2}} + f_e(2^{-N+4} + N^{-\frac{1}{2}}) + \kappa 2^{-N+1}. \tag{5.26}
 \end{aligned}$$

In addition, by (5.21) and (5.23), we can deduce from (5.18) that

$$\begin{aligned}
 P^W \left[ \left\{ \omega \in \Omega^W : \inf_{\nu \in \mathcal{I}} \|M^{\theta_0}(\omega) - \nu\| \geq 1/N + N^{-\frac{1}{2}} + 2^{-N+2} \right\} \right] \\
 \leq \frac{2\kappa + 2\alpha}{N} + 2^{-N} + 2(\kappa + \alpha) N^{-\frac{1}{2}},
 \end{aligned}$$

which for  $N$  large enough easily implies that

$$P^W \left[ \left\{ \omega \in \Omega^W : \inf_{\nu \in \mathcal{I}} \|M^{\theta_0}(\omega) - \nu\| \geq 4(\kappa + \alpha)N^{-\frac{1}{2}} \right\} \right] \leq 4(\kappa + \alpha)N^{-\frac{1}{2}}. \tag{5.27}$$

Similarly, by (5.21) and (5.23), we have

$$P^W [\|\hat{Z} - M^\theta\| \geq 2^{-N+2} + N^{-\frac{1}{2}}] \leq 2(\kappa + \alpha)N^{-\frac{1}{2}}. \tag{5.28}$$

*Step 3.* The next step is to modify the martingale  $M^{\theta_0}$  in such way that  $\Gamma$ , the new continuous martingale, is nonnegative. Write  $\epsilon_N = 2^{-N+4} + N^{-\frac{1}{2}}$  and define an  $\mathcal{F}_T^W$ -measurable random variable  $\Lambda \geq 0$  by  $\Lambda = (M_{T \wedge \theta_0} + \epsilon_N)/(1 + \epsilon_N)$ . Then

$$|\Lambda - M_{T_n \wedge \theta_0}| = \left| \epsilon_N \frac{1 - M_{T \wedge \theta_0}}{1 + \epsilon_N} \right| \leq \epsilon_N(1 + |M_{T \wedge \theta_0}|).$$

Note that for any  $i > d$ , we have  $\|\Lambda^{(i)}\| \leq \kappa + 1 + 2^{-N+2} + N^{-\frac{1}{2}} + \epsilon_N \leq \kappa + 2$  for  $N$  large enough. We now construct a continuous martingale from  $\Lambda$  by setting

$$\Gamma_t = E^W[\Lambda | \mathcal{F}_t^W], \quad t \in [0, T],$$

and  $\Lambda \geq 0$  implies that  $\Gamma$  is nonnegative, and  $\Gamma_0^{(i)} = 1$  for  $i \leq d + K$ . Hence  $\mathbb{P}^{(N)} := P^W \circ (\Gamma_t)^{-1} \in \underline{\mathcal{M}}$ .

We first note that for all  $i = 1, \dots, d + K$ ,

$$E^W[|M_{T \wedge \theta_0}^{(i)}|] = \mathbb{E}^W[M_{T \wedge \theta_0}^{(i)} + 2(M_{T \wedge \theta_0}^{(i)})^-] \leq \mathbb{E}^W[M_{T \wedge \theta_0}^{(i)} + 2] = 3.$$

Then by Doob’s martingale inequality,

$$\begin{aligned} P^W [\|\Gamma - M^{\theta_0}\| \geq \epsilon_N^{1/2}] &\leq \epsilon_N^{-1/2} \sum_{i=1}^{d+K} E^W[|\Lambda^{(i)} - M_{T \wedge \theta_0}^{(i)}|] \\ &\leq \epsilon_N^{-1/2} 4(d + K)\epsilon_N = 4(d + K)\epsilon_N^{1/2}. \end{aligned} \tag{5.29}$$

This together with (5.26), writing  $\kappa_1 = \kappa + \alpha$ , yields

$$\begin{aligned} &|E^W[G(\Gamma)] - \mathbb{E}_{\hat{\mathbb{Q}}}[G(\hat{\mathbb{S}})]| \\ &\leq E^W[|G(\Gamma) - \hat{G}(M^{\theta_0})|] + |E^W[\hat{G}(M^{\theta_0})] - \mathbb{E}_{\hat{\mathbb{Q}}}[\hat{G}(\hat{\mathbb{S}})]| \\ &\leq E^W \left[ |G(\Gamma) - G(M^{\theta_0} \vee 0)| \mathbb{1}_{\{\|\Gamma - M^{\theta_0}\| \leq \epsilon_N^{1/2}\}} \right] \\ &\quad + 8\kappa_1(d + K)\epsilon_N^{1/2} + 8\kappa_1^2 N^{-\frac{1}{2}} + f_e(2^{-N+4} + N^{-\frac{1}{2}}) + \kappa_1 2^{-N+1} \\ &\leq f_e(\epsilon_N^{1/2}) + 8\kappa_1(d + K)\epsilon_N^{1/2} + 9\kappa_1^2 \epsilon_N^{1/2} + f_e(\epsilon_N^{1/2}) \\ &\leq 2f_e(\epsilon_N^{1/2}) + 17\kappa_1^2(d + K)\epsilon_N^{1/2}. \end{aligned}$$

Finally, we can deduce from (5.27) and (5.29) that

$$\begin{aligned}
 P^W \left[ \left\{ \omega \in \Omega^W : \inf_{v \in \mathcal{I}} \|\Gamma(\omega) - v\| \geq 4\kappa_1 N^{-\frac{1}{2}} + \epsilon_N^{1/2} \right\} \right] \\
 \leq 4\kappa_1 N^{-\frac{1}{2}} + 4(d + K)\epsilon_N^{1/2},
 \end{aligned}
 \tag{5.30}$$

and from (5.28) and (5.29) that

$$P^W [\|\hat{Z} - \Gamma\| \geq \epsilon_N + \epsilon_N^{1/2}] \leq 2\kappa_1 N^{-\frac{1}{2}} + 4(d + K)\epsilon_N^{1/2}.
 \tag{5.31}$$

*Step 4.* The last step is to construct a new process  $\tilde{\Gamma}$  from  $\Gamma$  such that the law of  $\tilde{\Gamma}$  under  $P^W$  is an element of  $\mathcal{M}_{\mathcal{T}}$ . We write  $\eta_N = 4\kappa_1 N^{-\frac{1}{2}} + 4(d + K)\epsilon_N^{1/2}$  and

$$p_i^{(N)} := \mathbb{E}_{\mathbb{P}^{(N)}} [X_i^{(c)}(\mathbb{S}_T^{(1)}, \dots, \mathbb{S}_T^{(d)})]$$

for any  $i = 1, \dots, K$ , and define  $\tilde{p}_i^{(N)}$  by

$$\tilde{p}_i^{(N)} := \frac{\mathcal{P}(X_i^{(c)}) - (1 - \sqrt{\eta_N})p_i^{(N)}}{\sqrt{\eta_N}}.$$

We can deduce from (5.30) and (5.31) that

$$\begin{aligned}
 |\mathcal{P}(X_i^{(c)}) - p_i^{(N)}| &\leq E^W [ |X_i^{(c)}(\Gamma_T^{(1)}, \dots, \Gamma_T^{(d)}) - \mathcal{P}(X_i^{(c)})\Gamma_T^{(d+i)}| ] \\
 &\leq \mathcal{P}(X_i^{(c)})\eta_N + E^W \left[ |X_i^{(c)}(\Gamma_T^{(1)}, \dots, \Gamma_T^{(d)}) - \mathcal{P}(X_i^{(c)})\Gamma_T^{(d+i)}| \right. \\
 &\quad \left. \times \mathbb{1}_{\{|X_i^{(c)}(\Gamma_T^{(1)}, \dots, \Gamma_T^{(d)}) - \mathcal{P}(X_i^{(c)})\Gamma_T^{(d+i)}| > \eta_N\}} \right] \\
 &\leq \mathcal{P}(X_i^{(c)})\eta_N + 2(\kappa + 2)\mathcal{P}(X_i^{(c)})\eta_N, \quad \forall i = 1, \dots, K.
 \end{aligned}$$

It follows immediately that

$$\begin{aligned}
 |\tilde{p}_i^{(N)} - \mathcal{P}(X_i^{(c)})| &= \left( \frac{1}{\sqrt{\eta_N}} - 1 \right) |\mathcal{P}(X_i^{(c)}) - p_i^{(N)}| \\
 &\leq \frac{2(\kappa + 2)\mathcal{P}(X_i^{(c)})\eta_N}{\sqrt{\eta_N}} = 2(\kappa + 2)\mathcal{P}(X_i^{(c)})\sqrt{\eta_N}, \quad \forall i \leq K.
 \end{aligned}$$

Then it follows from Assumption 3.1 that when  $N$  is large enough, there exists a  $\tilde{\mathbb{P}}^{(N)} \in \underline{\mathcal{M}}_{\tilde{\mathcal{T}}}$  such that

$$\tilde{p}_i^{(N)} = \mathbb{E}_{\tilde{\mathbb{P}}^{(N)}} [X_i^{(c)}(\mathbb{S}_T^{(1)}, \dots, \mathbb{S}_T^{(d)})], \quad \forall i \leq K.$$

On the Wiener space  $(\Omega^W, \mathcal{F}^W, P^W)$ , or a suitable enlargement if necessary, there are continuous martingales  $\Gamma$  and  $\tilde{M}$  which have laws equal to  $\mathbb{P}^{(N)}$  and  $\tilde{\mathbb{P}}^{(N)}$  respectively, and an  $\mathcal{F}_T^W$ -measurable random variable  $\xi \in \{0, 1\}$  that is independent of  $\Gamma$  and  $\tilde{M}$  with

$$P^W [\xi = 1] = 1 - \sqrt{\eta_N}, \quad P^W [\xi = 0] = \sqrt{\eta_N}.$$



Define  $\mathcal{F}_T^W$ -measurable random variables  $\tilde{\Lambda}^{(i)}$  by

$$\begin{aligned} \tilde{\Lambda}^{(i)} &= \Gamma_T^{(i)} \mathbb{1}_{\{\xi=1\}} + \tilde{M}_T^{(i)} \mathbb{1}_{\{\xi=0\}}, \quad i = 1, \dots, d, \\ \tilde{\Lambda}^{(i)} &= X_{i-d}^{(c)}(\tilde{\Lambda}^{(1)}, \dots, \tilde{\Lambda}^{(d)})/\mathcal{P}(X_{i-d}^{(c)}), \quad d + K > i > d. \end{aligned}$$

We now construct a continuous martingale from  $\tilde{\Lambda}$  by setting

$$\tilde{\Gamma}_t = E^W[\tilde{\Lambda} | \mathcal{F}_t^W], \quad t \in [0, T].$$

It follows from the fact that  $\xi$  is independent of  $\Gamma$  and  $\tilde{M}$  that

$$\begin{aligned} \tilde{\Gamma}_0^{(i+d)} &= E^W[\tilde{\Gamma}_T^{(i+d)} | \mathcal{F}_0^W] \\ &= (1 - \sqrt{\eta_N})E^W[X_i^{(c)}(\Gamma_T^{(1)}, \dots, \Gamma_T^{(d)})/\mathcal{P}(X_i^{(c)})] \\ &\quad + \sqrt{\eta_N}E^W[X_i^{(c)}(\tilde{M}_T^{(1)}, \dots, \tilde{M}_T^{(d)})/\mathcal{P}(X_i^{(c)})] \\ &= \frac{(1 - \sqrt{\eta_N})p_i^{(N)} + \sqrt{\eta_N}\tilde{p}_i^{(N)}}{\mathcal{P}(X_i^{(c)})} = 1, \quad 1 \leq i \leq K, \end{aligned}$$

and

$$\begin{aligned} \tilde{\Gamma}_0^{(i)} &= E^W[\tilde{\Gamma}_T^{(i)} | \mathcal{F}_0^W] = E^W[\tilde{\Lambda}_T^{(i)} | \mathcal{F}_0^W] \\ &= (1 - \sqrt{\eta_N})E^W[\Gamma_T^{(i)}] + \sqrt{\eta_N}E^W[\tilde{M}_T^{(i)}] = 1, \quad i \leq d. \end{aligned}$$

Hence  $\tilde{\mathbb{P}} := P^W \circ (\tilde{\Gamma}_T)^{-1} \in \mathcal{M}_{\mathcal{I}}$ . Also by the independence between  $\xi$  and  $(\Gamma, \tilde{M})$ , we have

$$E^W[|\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)}|] = \sqrt{\eta_N}E^W[|\tilde{M}_T^{(i)} - \Gamma_T^{(i)}|] \leq 2\sqrt{\eta_N}, \quad i \leq d,$$

and by (5.30),

$$P^W[|\Gamma_T^{(i)} - \tilde{\Lambda}^{(i)}| > \eta_N] \leq \eta_N + \sqrt{\eta_N} \leq 2\sqrt{\eta_N}, \quad i > d,$$

which implies that

$$\begin{aligned} E^W[|\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)}|] &= 2E^W[(\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)})^+] - E^W[\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)}] \\ &= 2E^W[(\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)})^+] \\ &\leq 2\eta_N + 2E^W\left[\tilde{\Lambda}^{(i)} \mathbb{1}_{\{|\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)}| > \eta_N\}}\right] \\ &\leq 2\eta_N + 4(\kappa + 2)\sqrt{\eta_N} \leq 14\kappa\sqrt{\eta_N}, \quad i = d + 1, \dots, K. \end{aligned}$$

Then by Doob’s martingale inequality,

$$P^W[\|\tilde{\Gamma} - \Gamma\| \geq \kappa\eta_N^{1/4}] \leq \frac{1}{\kappa\eta_N^{1/4}} \sum_{i=1}^{d+K} E^W[|\tilde{\Lambda}^{(i)} - \Gamma_T^{(i)}|] \leq 14(d + K)\eta_N^{1/4}$$

and hence

$$\begin{aligned} |\mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] - \mathbb{E}_{\mathbb{P}^{(N)}}[G(\mathbb{S})]| &= |E^W[G(\tilde{\Gamma}) - G(\Gamma)]| \\ &\leq f_e(\kappa\eta_N^{1/4}) + E^W\left[|G(\Gamma) - G(\tilde{\Gamma})|\mathbb{1}_{\{\|\tilde{\Gamma}-\Gamma\|\geq\kappa\eta_N^{1/4}\}}\right] \\ &\leq f_e(\kappa\eta_N^{1/4}) + 28\kappa(d + K)\eta_N^{1/4}. \end{aligned}$$

In addition, we can deduce from (5.31) that

$$P^W[\|\hat{Z} - \tilde{\Gamma}\| \geq \kappa\eta_N^{1/4} + \epsilon_N + \epsilon_N^{1/2}] \leq 2\kappa_1 N^{-\frac{1}{2}} + 4(d + K)\epsilon_N^{1/2} + 14(d + K)\kappa\eta_N^{1/4}.$$

Notice that when  $N$  is sufficiently large such that  $\kappa\eta_N^{1/4} + \epsilon_N + \epsilon_N^{1/2} < 2^{-D-1}$ , we can deduce from (5.7) and (5.22) that on the event

$$\{\omega \in \Omega^W : \|\hat{Z}(\omega) - \tilde{\Gamma}(\omega)\| < \kappa\eta_N^{1/4} + \epsilon_N + \epsilon_N^{1/2} \text{ and } \hat{Z}(\omega) \in \hat{\mathbb{D}}^{(N)}\},$$

we have

$$\begin{aligned} |\downarrow(\hat{Z}) \vee 0 - \tilde{\Gamma}| &\leq |\downarrow(\hat{Z}) \vee 0 - \downarrow(\hat{Z})| + |\downarrow(\hat{Z}) - \hat{Z}| + |\hat{Z} - \tilde{\Gamma}| \\ &< 2^{-N+3} + 2^{-N+1} + 2^{-D-1} \leq 2^{-D}, \end{aligned}$$

and hence by Remark 4.2, the inequality  $\hat{m}^{(D)}(\hat{Z}) \geq m^{(D-2)}(\tilde{\Gamma})$  holds on

$$\{\omega \in \Omega^W : \|\hat{Z}(\omega) - \tilde{\Gamma}(\omega)\| < \kappa\eta_N^{1/4} + \epsilon_N + \epsilon_N^{1/2} \text{ and } \hat{Z}(\omega) \in \hat{\mathbb{D}}^{(N)}\}.$$

It follows that

$$\begin{aligned} \mathbb{E}_{\hat{\mathbb{Q}}}[&\beta\sqrt{\hat{m}^{(D)}(\hat{\mathbb{S}}) \wedge \alpha}] \geq \mathbb{E}_{\hat{\mathbb{Q}}}[&\beta\sqrt{\hat{m}^{(D)}(\hat{\mathbb{S}}^{m_0}) \wedge \alpha}] = E^W[\beta\sqrt{\hat{m}^{(D)}(\hat{Z}) \wedge \alpha}] \\ &\geq E^W[\beta\sqrt{m^{(D-2)}(\tilde{\Gamma}) \wedge \alpha}] - \alpha(2\kappa_1 N^{-\frac{1}{2}} + 4(d + K)\epsilon_N^{1/2} + 14(d + K)\kappa\eta_N^{1/4}). \end{aligned}$$

□

**Acknowledgements** Zhaoxu Hou gratefully acknowledges a PhD studentship from the Oxford-Man Institute of Quantitative Finance and support from Balliol College in Oxford. Jan Oblój gratefully acknowledges funding received from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement no. 335421 and is also thankful to the Oxford-Man Institute of Quantitative Finance and St John’s College in Oxford for their financial support. We are grateful for helpful discussions we have had with Mathias Beiglböck, Bruno Bouchard, Yan Dolinsky, Kostas Kardaras, Marcel Nutz, Mete Soner, Peter Spoida, Nizar Touzi and Johannes Wiesel.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## Appendix

### A.1 Proof of Lemma 4.3

Let  $f_e : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be a modulus of continuity for  $G$  and  $X_1^{(c)}, \dots, X_K^{(c)}$  so that  $\lim_{x \searrow 0} f_e(x) = 0$ . Now fix  $N$  and  $\mathbb{P}^{(N)} \in \mathcal{M}_{\pi^{(N)}, \mathcal{I}} \cap \underline{\mathcal{M}}$ . By the definition of  $\underline{\mathcal{M}}$ , there exist a complete probability space  $(\Omega^W, \mathcal{F}^W, P^W)$ , where  $\mathbb{F}^W = (\mathcal{F}_t^W)_{0 \leq t \leq T}$  is the  $P^W$ -completion of the natural filtration of a  $\tilde{d}$ -dimensional Brownian motion  $W$  on  $[0, T]$ , and a continuous martingale  $M$  defined on  $(\Omega^W, \mathcal{F}_T^W, \mathbb{F}^W, P^W)$  such that  $\mathbb{P}^{(N)} = P^W \circ M^{-1}$ .

Write  $\epsilon_N := d_p(\pi^{(N)}, \pi)$ . Saying that  $(\pi^{(N)})$  converges to  $\pi$  weakly is equivalent to saying that  $\epsilon_N \rightarrow 0$  as  $N \rightarrow \infty$ . Fix  $N$ . If  $\epsilon_N = 0$ , it is trivially true that

$$\sup_{\mathbb{P} \in \underline{\mathcal{M}} \cap \mathcal{M}_{\pi^{(N)}, \mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S}) - \beta \sqrt{m^{(D)}(\mathbb{S})} \wedge \alpha] = \sup_{\mathbb{P} \in \underline{\mathcal{M}} \cap \mathcal{M}_{\pi, \mathcal{I}}} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S}) - \beta \sqrt{m^{(D)}(\mathbb{S})} \wedge \alpha].$$

Therefore we only consider the case that  $\epsilon_N > 0$ . It follows from Strassen [49, Corollary of Theorem 11] that, possibly increasing  $d$ , we can find an  $\mathcal{F}_T^W$ -measurable random variable  $\Lambda$  such that  $\Lambda^{(d+i)} = X_i^{(c)}(\Lambda^{(1)}, \dots, \Lambda^{(d)})$  for every  $i \leq K$ ,

$$\mathcal{L}_{P^W}(\Lambda^{(1)}, \dots, \Lambda^{(d)}) = \pi \quad \text{and} \quad P^W[|\Lambda^{(i)} - M_T^{(i)}| > 2\epsilon_N] < 2\epsilon_N, \quad \forall i \leq d. \tag{A.1}$$

We now construct a continuous martingale from  $\Lambda$  by setting

$$\Gamma_t = E^W[\Lambda | \mathcal{F}_t^W], \quad t \in [0, T],$$

where  $E^W$  is the expectation with respect to  $P^W$ . Note that by the uniform continuity of  $X_i^{(c)}$ , we have

$$|\Lambda^{(d+i)} - M_T^{(d+i)}| \leq f_e(2\epsilon_N), \quad \forall i \leq K, \quad \text{when } |\Lambda^{(j)} - M_T^{(j)}| \leq 2\epsilon_N, \quad \forall j \leq d.$$

Hence, for every  $i \leq K$ ,

$$P^W[|\Lambda^{(i+d)} - M_T^{(i+d)}| > f_e(2\epsilon_N)] \leq 2d\epsilon_N.$$

Observe that  $E^W[\Lambda^{(i)}] = E^W[M_T^{(i)}] = 1$  and  $\Lambda^{(i)} \geq 0$   $P^W$ -a.s. for  $i \leq d + K$ . Then using (A.1), we get for  $i = 1, \dots, d$  that

$$\begin{aligned} E^W[|\Lambda^{(i)} - M_T^{(i)}|] &= 2E^W[(\Lambda^{(i)} - M_T^{(i)})^+] - E^W[\Lambda^{(i)} - M_T^{(i)}] \\ &= 2E^W[(\Lambda^{(i)} - M_T^{(i)})^+] \\ &\leq 4\epsilon_N + 2E^W\left[\Lambda^{(i)} \mathbb{1}_{\{|\Lambda^{(i)} - M_T^{(i)}| > 2\epsilon_N\}}\right] \\ &\leq 4\epsilon_N + 2E^W\left[\Lambda^{(i)} \mathbb{1}_{\{|\Lambda^{(i)} - M_T^{(i)}| > 2\epsilon_N\}} \mathbb{1}_{\{\Lambda^{(i)} > 1/\sqrt{\epsilon_N}\}}\right] + 4\sqrt{\epsilon_N} \\ &\leq 4\epsilon_N + 2 \int_{\{x_i \geq \frac{1}{\sqrt{\epsilon_N}}\} \cap \mathbb{R}_+^d} x_i \pi(dx_1, \dots, dx_d) + 4\sqrt{\epsilon_N}. \end{aligned}$$

Similarly, for every  $i \leq K$ ,

$$\begin{aligned} E^W [|\Lambda^{(d+i)} - M_T^{(d+i)}|] &= 2E^W [(\Lambda^{(d+i)} - M_T^{(d+i)})^+] \\ &\leq 2f_e(2\epsilon_N) + 2E^W \left[ \Lambda^{(i)} \mathbb{1}_{\{|\Lambda^{(d+i)} - M_T^{(d+i)}| > f_e(2\epsilon_N)\}} \right] \\ &\leq 2f_e(2\epsilon_N) + 4d \frac{\|X_i^{(c)}\|_\infty}{\mathcal{P}(X_i^{(c)})} \epsilon_N. \end{aligned}$$

We now define  $\eta_N$  by

$$\begin{aligned} \eta_N &:= 2f_e(2\epsilon_N) + 4\epsilon_N + 4d \sum_{i=1}^K \frac{\|X_i^{(c)}\|_\infty}{\mathcal{P}(X_i^{(c)})} \epsilon_N \\ &\quad + 2 \sum_{i=1}^d \int_{\{x_i \geq \frac{1}{\sqrt{\epsilon_N}}\} \cap \mathbb{R}_+^d} x_i \pi(dx_1, \dots, dx_d) + 4\sqrt{\epsilon_N}. \end{aligned}$$

Note that  $\eta_N \rightarrow 0$  as  $N \rightarrow \infty$ . Then by Doob’s martingale inequality,

$$P^W [\|\Gamma - M\| \geq \eta_N^{1/2}] \leq \eta_N^{-1/2} \sum_{i=1}^{d+K} E^W [|\Lambda^{(i)} - M_T^{(i)}|] \leq (d + K)\eta_N^{1/2}.$$

It follows that

$$\begin{aligned} |E^W [G(\Gamma) - G(M)]| &\leq 2(d + K)\|G\|_\infty \eta_N^{1/2} \\ &\quad + E^W \left[ |G(\Gamma) - G(M)| \mathbb{1}_{\{\|\Gamma - M\| < \eta_N^{1/2}\}} \right] \\ &\leq 2(d + K)\|G\|_\infty \eta_N^{1/2} + f_e(\eta_N^{1/2}). \end{aligned}$$

Note that by (4.1) in Remark 4.2, for  $N$  sufficiently large,

$$E^W [\beta\sqrt{m^{(D)}}(M) \wedge \alpha] \geq E^W [\beta\sqrt{m^{(D-2)}}(\Gamma) \wedge \alpha] - \alpha\eta_N^{1/2}.$$

As  $\mathbb{P}^{(N)} \in \mathcal{M}_{\pi^{(N)}, \mathcal{I}} \cap \underline{\mathcal{M}}$  is arbitrary,

$$\begin{aligned} &\sup_{\mathbb{P} \in \underline{\mathcal{M}} \cap \mathcal{M}_{\pi^{(N)}, \mathcal{I}}} \mathbb{E}_{\mathbb{P}} [G(\mathbb{S}) - \beta\sqrt{m^{(D)}}(\mathbb{S}) \wedge \alpha] \\ &\leq c_N \sup_{\mathbb{P} \in \underline{\mathcal{M}} \cap \mathcal{M}_{\pi, \mathcal{I}}} \mathbb{E}_{\mathbb{P}} [G(\mathbb{S}) - \beta\sqrt{m^{(D-2)}}(\mathbb{S}) \wedge \alpha], \end{aligned}$$

where  $c_N := 2(2(d + K)\|G\|_\infty \eta_N^{1/2} + f_e(\eta_N^{1/2}) + \alpha\eta_N^{1/2})$ . The inequality asserted in Lemma 4.3 now follows. □

### A.2 Proof of Lemma 4.4

Choose  $\kappa > 2 \vee (\|G\|_\infty + \alpha)$ . We first observe that

$$\begin{aligned} & \inf_{X \in \text{Lin}(\mathcal{X}), N \geq 0} \left( \sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D)}} \wedge \alpha - X - N\lambda_{\mathfrak{F}}] + \mathcal{P}(X) \right) \\ &= \lim_{N \rightarrow \infty} \inf_{X \in \text{Lin}_N(\mathcal{X})} \left( \sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D)}} \wedge \alpha - X - N\lambda_{\mathfrak{F}}] + \mathcal{P}(X) \right). \end{aligned}$$

Define the function  $\mathcal{G} : \text{Lin}_N(\mathcal{X}) \times \mathcal{M}_s \rightarrow \mathbb{R}$  by

$$\begin{aligned} \mathcal{G}(X, \mathbb{P}) &:= \lim_{\epsilon \searrow 0} \inf_{\tilde{\mathbb{P}} \in \mathcal{M}_s, d_p(\tilde{\mathbb{P}}, \mathbb{P}) < \epsilon} \mathbb{E}_{\tilde{\mathbb{P}}}[G - \beta\sqrt{m^{(D-2)}} \wedge \alpha - X - N\lambda_{\mathfrak{F}}] + \mathcal{P}(X) \\ &= \lim_{\epsilon \searrow 0} \inf_{\tilde{\mathbb{P}} \in \mathcal{M}_s, d_p(\tilde{\mathbb{P}}, \mathbb{P}) < \epsilon} \mathbb{E}_{\tilde{\mathbb{P}}}[-\beta\sqrt{m^{(D-2)}} \wedge \alpha] \\ &\quad + \mathbb{E}_{\mathbb{P}}[G - N\lambda_{\mathfrak{F}} - X] + \mathcal{P}(X). \end{aligned}$$

Then by (4.1) in Remark 4.2, for any sequence  $(\mathbb{P}^{(k)})_{k \geq 1}$  converging to  $\mathbb{P}$  weakly,

$$\mathbb{E}_{\mathbb{P}}[-\beta\sqrt{m^{(D)}}(\mathbb{S}) \wedge \alpha] \leq \liminf_{k \rightarrow \infty} \mathbb{E}_{\mathbb{P}^{(k)}}[-\beta\sqrt{m^{(D-2)}}(\mathbb{S}) \wedge \alpha]$$

and hence

$$\begin{aligned} & \lim_{N \rightarrow \infty} \inf_{X \in \text{Lin}_N(\mathcal{X})} \left( \sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D)}} \wedge \alpha - X - N\lambda_{\mathfrak{F}}] + \mathcal{P}(X) \right) \\ & \leq \lim_{N \rightarrow \infty} \inf_{X \in \text{Lin}_N(\mathcal{X})} \sup_{\mathbb{P} \in \mathcal{M}_s} \mathcal{G}(X, \mathbb{P}). \end{aligned} \tag{A.2}$$

The next step is to interchange the order of the infimum and supremum. Notice that when we fix  $\mathbb{P}$ ,  $\mathcal{G}$  is affine in the first variable and continuous due to the dominated convergence theorem. In addition, by definition,  $\mathcal{G}$  is lower semi-continuous in the second variable. Furthermore,  $\mathcal{G}$  is convex in the second variable. To justify this, we notice that  $\mathbb{P} \mapsto \mathbb{E}_{\mathbb{P}}[-\beta\sqrt{m^{(D-2)}}(\mathbb{S}) \wedge \alpha]$  is a linear functional, and it follows that for each  $\epsilon > 0$  and  $\lambda \in [0, 1]$ ,

$$\begin{aligned} & \inf_{\tilde{\mathbb{P}} \in \mathcal{M}_s, d_p(\tilde{\mathbb{P}}, \lambda\mathbb{P}^{(1)} + (1-\lambda)\mathbb{P}^{(2)}) < \epsilon} \mathbb{E}_{\tilde{\mathbb{P}}}[-\beta\sqrt{m^{(D-2)}}(\mathbb{S}) \wedge \alpha] \\ & \leq \lambda \inf_{\tilde{\mathbb{P}} \in \mathcal{M}_s, d_p(\tilde{\mathbb{P}}, \mathbb{P}^{(1)}) < \epsilon} \mathbb{E}_{\tilde{\mathbb{P}}}[-\beta\sqrt{m^{(D-2)}}(\mathbb{S}) \wedge \alpha] \\ & \quad + (1-\lambda) \inf_{\tilde{\mathbb{P}} \in \mathcal{M}_s, d_p(\tilde{\mathbb{P}}, \mathbb{P}^{(2)}) < \epsilon} \mathbb{E}_{\tilde{\mathbb{P}}}[-\beta\sqrt{m^{(D-2)}}(\mathbb{S}) \wedge \alpha]. \end{aligned}$$

Since  $\text{Lin}_N(\mathcal{X})$  is convex and compact, it follows that we can now apply a minimax theorem (see Terkelsen [50, Corollary 2]) to  $\mathcal{G}$  and derive

$$\lim_{N \rightarrow \infty} \inf_{X \in \text{Lin}_N(\mathcal{X})} \sup_{\mathbb{P} \in \mathcal{M}_s} \mathcal{G}(X, \mathbb{P}) = \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_s} \inf_{X \in \text{Lin}_N(\mathcal{X})} \mathcal{G}(X, \mathbb{P}).$$

Combining this with (A.2) yields

$$\begin{aligned} & \lim_{N \rightarrow \infty} \inf_{X \in \text{Lin}_N(\mathcal{X})} \left( \sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D)}} \wedge \alpha - X - N\lambda_{\mathfrak{F}}] + \mathcal{P}(X) \right) \\ & \leq \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_s} \inf_{X \in \text{Lin}_N(\mathcal{X})} \left( \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D-2)}} \wedge \alpha - X - N\lambda_{\mathfrak{F}}] + \mathcal{P}(X) \right). \end{aligned} \tag{A.3}$$

We begin to verify (4.9). Since  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}^\eta \cap \mathcal{M}_s \neq \emptyset$  for all  $\eta > 0$ ,

$$\sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[-X - N\lambda_{\mathfrak{F}}] + \mathcal{P}(X) \geq 0, \quad \forall X \in \mathcal{X}, N \in \mathbb{R}_+.$$

Hence for every  $N$  and  $X \in \text{Lin}_N(X)$ ,

$$\begin{aligned} & \sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D)}} \wedge \alpha - X - N\lambda_{\mathfrak{F}}] + \mathcal{P}(X) \\ & \geq -\|G\|_\infty - \alpha + \sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[-X - N\lambda_{\mathfrak{F}}] + \mathcal{P}(X) \geq -\kappa, \end{aligned}$$

and therefore

$$\lim_{N \rightarrow \infty} \inf_{X \in \text{Lin}_N(\mathcal{X})} \left( \sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D)}} \wedge \alpha - X - N\lambda_{\mathfrak{F}}] + \mathcal{P}(X) \right) \geq -\kappa.$$

By the arguments used to justify the inequality between (4.5) and (4.6) in Sect. 4.1, in the sup term of (A.3), it suffices to consider probability measures in  $\mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}^{\eta_N} \cap \mathcal{M}_s$ , where  $\eta_N = 2\kappa/\sqrt{N}$ . Hence we have

$$\begin{aligned} & \lim_{N \rightarrow \infty} \inf_{X \in \text{Lin}_N(\mathcal{X})} \left( \sup_{\mathbb{P} \in \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D)}} \wedge \alpha - X - N\lambda_{\mathfrak{F}}] + \mathcal{P}(X) \right) \\ & \leq \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}^{\eta_N} \cap \mathcal{M}_s} \inf_{X \in \text{Lin}_N(\mathcal{X})} \left( \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D-2)}} \wedge \alpha - X] + \mathcal{P}(X) \right) \\ & \leq \lim_{N \rightarrow \infty} \sup_{\mathbb{P} \in \mathcal{M}_{\mathcal{X}, \mathcal{P}, \mathfrak{F}}^{\eta_N} \cap \mathcal{M}_s} \mathbb{E}_{\mathbb{P}}[G - \beta\sqrt{m^{(D-2)}} \wedge \alpha], \end{aligned}$$

where the second inequality follows from the fact that  $-X \in \text{Lin}_N(\mathcal{X})$  for every  $X \in \text{Lin}_N(\mathcal{X})$ . This completes the verification of (4.9). □

### A.3 Construction of $\sigma$ and $Y$ in Proposition 5.11

Given a sequence  $a_1, a_2, \dots$ , we denote by  $\mathbf{a}_m := (a_1, \dots, a_m)$  the vector of its first  $m$  elements. We denote by  $\Pi(E)$  the set of probability measures on  $E$ . In addition, set

$$\begin{aligned} \mathbb{T} & := \mathbb{Q}_+ \cup \{T - b : b \in \mathbb{Q}_+ \cap [0, T]\}, \\ \mathcal{S}_k & := \left\{ \frac{1}{2^{N+k}}(a_1, \dots, a_{d+k}) : a_j \in \mathbb{Z}, |a_j| \leq 2^k, j = 1, \dots, d + K \right\}. \end{aligned}$$

For  $k = 1, \dots, m_0$ , define the function  $\Psi_k : \mathbb{T}^k \times \mathcal{S}_1 \times \dots \times \mathcal{S}_k \rightarrow \Pi(\mathbb{R})$  by

$$\Psi_k(\boldsymbol{\alpha}_k; \boldsymbol{\beta}_k) := \mathcal{L}_{\hat{\mathbb{Q}}}(\hat{\tau}_{k+1} - \hat{\tau}_k | \hat{\tau}_i - \hat{\tau}_{i-1} = \alpha_i, \hat{S}_{\hat{\tau}_i} - \hat{S}_{\hat{\tau}_{i-1}} = \beta_i, i \leq k) \tag{A.4}$$

and  $\Phi_k : \mathbb{T}^k \times \mathcal{S}_1 \times \dots \times \mathcal{S}_{k-1} \rightarrow \Pi(\mathbb{R}^{d+K})$  by

$$\Phi_k(\boldsymbol{\alpha}_k; \boldsymbol{\beta}_{k-1}) := \mathcal{L}_{\hat{\mathbb{Q}}}(\hat{S}_{\hat{\tau}_k} - \hat{S}_{\hat{\tau}_{k-1}} | \hat{\tau}_j - \hat{\tau}_{j-1} = \alpha_j, \hat{S}_{\hat{\tau}_i} - \hat{S}_{\hat{\tau}_{i-1}} = \beta_i, j \leq k, i \leq k-1), \tag{A.5}$$

where we set  $\mathcal{L}_{\hat{\mathbb{Q}}}(\cdot | \emptyset) \equiv \delta_0$ , the Dirac measure on the zero vector. Next, let  $\mathcal{B}$  be the set of barriers; see [27, Def. 2.1]. Then by Theorem 2.3 in [27], for any  $k \leq m_0$ , we can find  $\Upsilon_k : \mathbb{T}^k \times \mathcal{S}_1 \times \dots \times \mathcal{S}_k \rightarrow [-\infty, \infty]$  and  $B_k : \mathbb{T}^k \times \mathcal{S}_1 \times \dots \times \mathcal{S}_k \rightarrow \mathcal{B}$  such that

$$\mathcal{L}_{PW} \left( \sum_{\ell=0}^{\infty} s_{k,\ell+1} \mathbb{1}_{\{\Upsilon_k(\boldsymbol{\alpha}_k; \boldsymbol{\beta}_{k-1, s_{k,\ell}}) \leq W_{\alpha_k}^{(1)} < \Upsilon_k(\boldsymbol{\alpha}_k; \boldsymbol{\beta}_{k-1, s_{k,\ell+1}})\}} \right) = \Phi_k(\boldsymbol{\alpha}_k; \boldsymbol{\beta}_k), \tag{A.6}$$

where  $\{s_{k,1}, s_{k,2}, \dots, s_{k,\ell}, \dots\}$  is an enumeration of  $\mathcal{S}_k$  and

$$\mathcal{L}_{PW}(\tau_{B_k(\boldsymbol{\alpha}_k; \boldsymbol{\beta}_k)}(W^{(1)})) = \Psi_k(\boldsymbol{\alpha}_k; \boldsymbol{\beta}_k), \tag{A.7}$$

where  $\tau_{B_k(\boldsymbol{\alpha}_k; \boldsymbol{\beta}_k)}(W^{(1)})$  is the first hitting time of  $B_k(\boldsymbol{\alpha}_k; \boldsymbol{\beta}_k)$  by  $W^{(1)}$ .

Now we set  $\sigma_0 \equiv 0$  and define the random variables  $\sigma_1, \dots, \sigma_{m_0}, Y_1, \dots, Y_{m_0}$  recursively by

$$\begin{aligned} \Delta_i &= \tau_{B_{i-1}(\Delta_{i-1}; Y_{i-1})}(\{W_{t+\sigma_{i-1}}^{(1)} - W_{\sigma_{i-1}}^{(1)}\}_{t \geq 0}), \\ \sigma_i &= \sigma_{i-1} + \Delta_i, \\ Y_i &= \mathbb{1}_{\{\sigma_i < T\}} \sum_{j=1}^{\infty} s_{i,j} \mathbb{1}_{\{\Upsilon_i(\Delta_i; \mathbf{Y}_{i-1}; s_{i,j-1}) \leq W_{\sigma_i}^{(i+1)} - W_{\sigma_{i-1}}^{(i+1)} < \Upsilon_i(\Delta_i; \mathbf{Y}_{i-1}; s_{i,j})\}}. \end{aligned}$$

Note that the  $\sigma_i$  are stopping times with respect to the Brownian filtration. Fix  $k \leq m_0$  and  $(\boldsymbol{\alpha}_k; \boldsymbol{\beta}_{k-1}) \in \mathbb{T}^k \times \mathcal{S}_1 \times \dots \times \mathcal{S}_{k-1}$ . By the strong Markov property and the independence of the Brownian motion increments, it follows from (A.7) that

$$\mathcal{L}_{PW}(\Delta_k | (\boldsymbol{\Delta}_{k-1}, \mathbf{Y}_{k-1}) = (\boldsymbol{\alpha}_{k-1}, \boldsymbol{\beta}_{k-1})) = \Psi_{k-1}(\boldsymbol{\alpha}_{k-1}; \boldsymbol{\beta}_{k-1}). \tag{A.8}$$

Similarly, from (A.6) and (A.8), we have

$$\mathcal{L}_{PW}(Y_k | \boldsymbol{\Delta}_k = \boldsymbol{\sigma}_k, \mathbf{Y}_{k-1} = \boldsymbol{\beta}_{k-1}) = \Phi(\boldsymbol{\alpha}_k; \boldsymbol{\beta}_{k-1}). \tag{A.9}$$

Therefore, using (A.4), (A.5), (A.8) and (A.9), we conclude that

$$\mathcal{L}_{PW}((\boldsymbol{\sigma}_{m_0}; \mathbf{Y}_{m_0})) = \mathcal{L}_{\hat{\mathbb{Q}}}((\hat{\boldsymbol{\tau}}_{m_0}; \boldsymbol{\Delta} \hat{\mathbf{S}}_{m_0})),$$

where  $\boldsymbol{\Delta} \hat{\mathbf{S}}_k = \hat{S}_{\hat{\tau}_k} - \hat{S}_{\hat{\tau}_{k-1}}, k \leq m_0$ .

## References

1. Acciaio, B., Beiglböck, M., Penkner, F., Schachermayer, W.: A model-free version of the fundamental theorem of asset pricing and the superreplication theorem. *Math. Finance* **26**, 233–251 (2016)
2. Avellaneda, M., Levy, A., Parás, A.: Pricing and hedging derivative securities in markets with uncertain volatilities. *Appl. Math. Finance* **2**, 73–88 (1995)
3. Bayraktar, E., Zhou, Z.: On arbitrage and duality under model uncertainty and portfolio constraints. *Math. Finance* **27**, 988–1012 (2017)
4. Beiglböck, M., Cox, A.M.G., Huesmann, M., Perkowski, N., Prömel, D.: Pathwise superreplication via Vovk’s outer measure. *Finance Stoch.* **21**, 1141–1166 (2017)
5. Beiglböck, M., Henry-Labordère, P., Penkner, F.: Model-independent bounds for option prices: a mass transport approach. *Finance Stoch.* **17**, 477–501 (2013)
6. Biagini, S., Bouchard, B., Kardaras, C., Nutz, M.: Robust fundamental theorem for continuous processes. *Math. Finance* **27**, 963–987 (2017)
7. Bichteler, K.: Stochastic integration and  $L^p$ -theory of semimartingales. *Ann. Probab.* **9**, 49–89 (1981)
8. Black, F., Scholes, M.: The pricing of options and corporate liabilities. *J. Polit. Econ.* **81**, 637–654 (1973)
9. Bogachev, V.I.: *Measure Theory*, vols. I, II. Springer, Berlin (2007)
10. Bouchard, B., Nutz, M.: Arbitrage and duality in nondominated discrete-time models. *Ann. Appl. Probab.* **25**, 823–859 (2015)
11. Breeden, D.T., Litzenberger, R.H.: Prices of state-contingent claims implicit in option prices. *J. Bus.*, 621–651 (1978)
12. Brown, H., Hobson, D., Rogers, L.C.G.: Robust hedging of barrier options. *Math. Finance* **11**, 285–314 (2001)
13. Burzoni, M., Frittelli, M., Hou, Z., Maggis, M., Oblój, J.: Pointwise arbitrage pricing theory in discrete time. Preprint (2016). Available online at [arXiv:1612.07618v2](https://arxiv.org/abs/1612.07618v2)
14. Burzoni, M., Frittelli, M., Maggis, M.: Universal arbitrage aggregator in discrete-time markets under uncertainty. *Finance Stoch.* **20**, 1–50 (2016)
15. Cassese, G.: Asset pricing with no exogenous probability measure. *Math. Finance* **18**, 23–54 (2008)
16. Cox, A.M.G., Hou, Z., Oblój, J.: Robust pricing and hedging under trading restrictions and the emergence of local martingale models. *Finance Stoch.* **20**, 669–704 (2016)
17. Cox, A.M.G., Oblój, J.: Robust pricing and hedging of double no-touch options. *Finance Stoch.* **15**, 573–605 (2011)
18. Cox, A.M.G., Wang, J.: Root’s barrier: construction, optimality and applications to variance options. *Ann. Appl. Probab.* **23**, 859–894 (2013)
19. Davis, M.H.A., Oblój, J., Raval, V.: Arbitrage bounds for prices of weighted variance swaps. *Math. Finance* **24**, 821–854 (2014)
20. Delbaen, F., Schachermayer, W.: A general version of the fundamental theorem of asset pricing. *Math. Ann.* **300**, 463–520 (1994)
21. Delbaen, F., Schachermayer, W.: The fundamental theorem of asset pricing for unbounded stochastic processes. *Math. Ann.* **312**, 215–250 (1998)
22. Dellacherie, C., Meyer, P.: *Probabilities and Potential*. North-Holland Mathematics Studies (1978). Chapters I–IV
23. Denis, L., Martini, C.: A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. *Ann. Appl. Probab.* **16**, 827–852 (2006)
24. Dolinsky, Y., Soner, H.M.: Robust hedging with proportional transaction costs. *Finance Stoch.* **18**, 327–347 (2014)
25. Dolinsky, Y., Soner, H.M.: Martingale optimal transport and robust hedging in continuous time. *Probab. Theory Relat. Fields* **160**, 391–427 (2014)
26. Dolinsky, Y., Soner, H.M.: Martingale optimal transport in the Skorokhod space. *Stoch. Process. Appl.* **125**, 3893–3931 (2015)
27. Ekström, E., Janson, S.: The inverse first-passage problem and optimal stopping. *Ann. Appl. Probab.* **26**, 3154–3177 (2016)
28. Fahim, A., Huang, Y.: Model-independent superhedging under portfolio constraints. *Finance Stoch.* **20**, 51–81 (2016)
29. Föllmer, H., Kramkov, D.: Optional decompositions under constraints. *Probab. Theory Relat. Fields* **109**, 1–25 (1997)
30. Galichon, A., Henry-Labordère, P., Touzi, N.: A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. *Ann. Appl. Probab.* **24**, 312–336 (2014)



31. Guo, G., Tan, X., Touzi, N.: Tightness and duality of martingale transport on the Skorokhod space. *Stoch. Process. Appl.* **127**, 927–956 (2017)
32. Henry-Labordère, P., Oblój, J., Spoida, P., Touzi, N.: The maximum maximum of a martingale with given  $n$  marginals. *Ann. Appl. Probab.* **26**, 1–44 (2016)
33. Hobson, D.G.: Robust hedging of the lookback option. *Finance Stoch.* **2**, 329–347 (1998)
34. Hou, Z.: A robust approach to pricing–hedging duality and related problems in mathematical finance. PhD thesis, University of Oxford (2017). Available online at <https://ora.ox.ac.uk/objects/uuid:4a21584a-f898-43ac-bfa5-914fec17961e>
35. Jacod, J., Shiryaev, A.N.: *Limit Theorems for Stochastic Processes*, 2nd edn. Springer, Berlin (2002)
36. Knight, F.: *Risk, Uncertainty and Profit*. Houghton Mifflin, Boston (1921)
37. Lyons, T.J.: Uncertain volatility and the risk-free synthesis of derivatives. *Appl. Math. Finance* **2**, 117–133 (1995)
38. Merton, R.C.: Theory of rational option pricing. *Bell J. Econ.* **4**, 141–183 (1973)
39. Mykland, P.A.: Financial options and statistical prediction intervals. *Ann. Stat.* **31**, 1413–1438 (2003)
40. Nadochiy, S., Oblój, J.: Robust trading of implied skew. *Int. J. Theor. Appl. Finance* **20**, 1–41 (2017)
41. Neufeld, A., Nutz, M.: Superreplication under volatility uncertainty for measurable claims. *Electron. J. Probab.* **18**(48), 1–14 (2013)
42. Oblój, J.: Skorokhod Embedding. In: Cont, R. (ed.) *Encyclopedia of Quantitative Finance*, pp. 1653–1657. Wiley, Chichester (2010)
43. Oblój, J., Ulmer, F.: Performance of robust hedges for digital double barrier options. *Int. J. Theor. Appl. Finance* **15**, 1–34 (2012)
44. Possamaï, D., Royer, G., Touzi, N.: On the robust superhedging of measurable claims. *Electron. Commun. Probab.* **18**, 1–13 (2013)
45. Riedel, F.: Financial economics without probabilistic prior assumptions. *Decis. Econ. Finance* **38**, 75–91 (2015)
46. Rogers, L.C.G., Williams, D.: *Diffusions, Markov Processes and Martingales: Volume 2, Itô Calculus*. Cambridge University Press, Cambridge (2000)
47. Samuelson, P.: Rational theory of warrant pricing. *Ind. Manage. Rev.* **6**, 13–31 (1965)
48. Soner, H.M., Touzi, N., Zhang, J.: Quasi-sure stochastic analysis through aggregation. *Electron. J. Probab.* **16**(2), 1844–1879 (2011)
49. Strassen, V.: The existence of probability measures with given marginals. *Ann. Math. Stat.* **36**, 423–439 (1965)
50. Terkelsen, F.: Some minimax theorems. *Math. Scand.* **31**, 405–413 (1972)
51. Vovk, V.: Continuous-time trading and the emergence of probability. *Finance Stoch.* **16**, 561–609 (2012)
52. Whalley, A.E., Wilmott, P.: An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Math. Finance* **7**, 307–324 (1997)