Martingale optimal transport and robust hedging in continuous time

Journal Article

Author(s):
Dolinsky, Yan; Soner, Mete (ID)

Martingale optimal transport and robust hedging in continuous time

Yan Dolinsky · H. Mete Soner

Abstract The duality between the robust (or equivalently, model independent) hedging of path dependent European options and a martingale optimal transport problem is proved. The financial market is modeled through a risky asset whose price is only assumed to be a continuous function of time. The hedging problem is to construct a minimal super-hedging portfolio that consists of dynamically trading the underlying risky asset and a static position of vanilla options which can be exercised at the given, fixed maturity. The dual is a Monge–Kantorovich type martingale transport problem of maximizing the expected value of the option over all martingale measures that have a given marginal at maturity. In addition to duality, a family of simple, piecewise constant super-replication portfolios that asymptotically achieve the minimal super-replication cost is constructed.

Keywords European options · Robust hedging · Min–max theorems · Prokhorov metric · Optimal transport

Mathematics Subject Classification (1991) 91G10 · 60G44

Y. Dolinsky
Department of Statistics, Hebrew University of Jerusalem, Jerusalem, Israel
e-mail: yan.dolinsky@mail.huji.ac.il

H. M. Soner (✉)
Department of Mathematics, ETH Zurich and Swiss Finance Institute, Zurich, Switzerland
e-mail: hmsoner@ethz.ch

1 Introduction

The original transport problem proposed by Monge [33] is to optimally move a pile
of soil to an excavation. Mathematically, given two measures ν and μ of equal mass,
we look for an optimal bijection of $\mathbb{R}^d$ which moves ν to μ, i.e., look for a map S so
that

$$\int_{\mathbb{R}^d} \varphi(S(x)) d\nu(x) = \int_{\mathbb{R}^d} \varphi(x) d\mu(x),$$

for all continuous functions φ. Then, with a given cost function c, the objective is to
minimize

$$\int_{\mathbb{R}^d} c(x, S(x)) \, d\nu(x)$$

over all bijections S.

In his classical papers [29,30], Kantorovich relaxed this problem by considering
a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$, whose marginals agree with ν and μ, instead of
a bijection. This generalization linearizes the problem. Hence it allows for an easy
existence result and enables one to identify its convex dual. Indeed, the dual elements
are real-valued continuous maps (g, h) of $\mathbb{R}^d$ satisfying the constraint

$$g(x) + h(y) \leq c(x, y). \tag{1.1}$$

The dual objective is to maximize

$$\int_{\mathbb{R}^d} g(x) \, d\nu(x) + \int_{\mathbb{R}^d} h(y) \, d\mu(y)$$

overall (g, h) satisfying the constraint 1.1. In the last decades an impressive theory
has been developed and we refer the reader to [1,40,41] and to the references therein.

In robust hedging problems, we are also given two measures. Namely, the initial and
the final distributions of a stock process. We then construct an optimal connection. In
general, however, the cost functional depends on the whole path of this connection and
not simply on the final value. Hence, one needs to consider processes instead of simply
the maps S. The probability distribution of this process has prescribed marginals at
final and initial times. Thus, it is in direct analogy with the Kantorovich measure.
But, financial considerations restrict the process to be a martingale (see Definition
2.4). Interestingly, the dual also has a financial interpretation as a robust hedging
(super-replication) problem. Indeed, the replication constraint is similar to (1.1). The
formal connection between the original Monge–Kantorovich problem and the financial
problem is further discussed in Remark 2.9 and also in the papers [6,21].

We continue by describing the robust hedging problem. Consider a financial market
consisting of one risky asset with a continuous price process. As in the classical paper

of Hobson [22], all call options are liquid assets and can be traded for a "reasonable" price that is known initially. Hence, the portfolio of an investor consists of static positions in the call options in addition to the usual dynamically updated risky asset. This leads us to a similar structure to that in [22] and in other papers [7,9,11,12,15, 16,21,24–28,32] which consider model-independent pricing. This approach is very closely related to path-wise proofs of well-known probabilistic inequalities [2,10]. Apart from the continuity of the price process no other model assumptions are placed on the dynamics of the price process.

In this market, we prove the Kantorovich duality, Theorem 2.7, and an approximation result, Theorem 2.10, for a general class of path-dependent options. The classical duality theorem, for a market with a risky asset whose price process is a semi–martingale, states that the minimal super-replication cost of a contingent claim is equal to the supremum of its expected value over all martingale measures that are equivalent to a given measure. We refer the reader to Delbaen and Schachermayer [17] (Theorem 5.7) for the case of general semi-martingale processes and to El-Karoui and Quenez [20] for its dynamic version in the diffusion case. Theorem 2.7 below, also provides a dual representation of the minimal super-replication cost but for model independent markets. The dual is given as the supremum of the expectations of the contingent claim over all martingale measures with *a given marginal at the maturity but with no dominating measure*. Since no probabilistic model is pre-assumed for the price process, the class of all martingale measures is quite large. Moreover, martingale measures are typically orthogonal to each other. These facts render the problem difficult.

In the literature, there are two earlier results in this direction. In a purely discrete setup, a similar result was recently proved by Beiglböck, Henry-Labordère and Penkner [6]. In their model, the investor is allowed to buy all call options at finitely many given maturities and the stock is traded only at these possible maturities. In this paper, however, the stock is traded in continuous time together with a static position in the calls with one maturity. In [6] the dual is recognized as a Monge–Kantorovich type optimal transport for martingale measures and the main tool in [6] is a duality result from optimal transport (see Theorem 2.14 in [31]).

In continuous time, Galichon, Henry-Labordère and Touzi [21] prove a different duality and then use the dual to convert the problem to an optimal control problem. There are two main differences between our result and the one proved in [21]. The duality result, Proposition 2.1 in [21], states that the minimal super-replication cost is given as the infimum over Lagrange multipliers and supremum over martingale measures *without* the final time constraint and the Lagrange multipliers are related to the constraint. Also the problem formulation is different. The model in [21] assumes a large class of possible martingale measures for the price process. The duality is then proved by extending an earlier unconstrained result proved in [38]. As in the unconstrained model of [14,38,39], the super replication is defined not path-wise but rather probabilistically through quasi-sure inequalities. Namely, the super-replication cost is the minimal initial wealth from which one can super-replicate the option almost surely with respect to all measures in a given class. In general, these measures are not dominated by one measure. As already mentioned this is the main difficulty and sets the current problem apart from the classical duality discussed earlier. However, our

duality result together with the results of [21] implies that these two approaches—namely, robust hedging through the path-wise definition of this paper and the quasi-sure definition of [14,38,39] yield the same value. This is proved in Sect. 3 below.

Our second result provides a class of portfolios which are managed on a finite number of random times and asymptotically achieve the minimal super-replication cost. This result may have practical implications allowing us to numerically investigate the corresponding discrete hedges, but we relegate this to a future study.

Robust hedging has been an active research area over the past decade. The initial paper of Hobson [22] studies the case of the lookback option. The connection to the Skorokhod embedding is also made in this paper and an explicit solution for the minimal super-replication cost is obtained. This approach is further developed by Brown, Hobson and Rogers [7], Cox and Obłój [11,13] and in several other papers, [23–28]. We refer the reader to the excellent survey of Hobson [23] for robust hedging and to Obłój [34] for the Skorokhod embedding problem. In particular, the recent paper by Cox and Wang [13] provides a discussion of various constructions of Root's solution of the Skorokhod embedding.

A similar modeling approach is applied to volatility options by Carr and Lee [9]. In a recent paper, Davis, Obłój and Raval [16] considers the variance swaps in a market with finitely many put options. In particular, in [16] the class of admissible portfolios is enlarged and numerical evidence is obtained by analyzing the S&P500 index options data. Furthermore, [16] contains a duality result in a simpler setting, using the classical Karlin–Isii duality in semi-infinite linear programming.

As already mentioned above, the dual approach is used by Galichon, Henry-Labordère and Touzi [21] and Henry-Labordère and Touzi [32] as well. In these papers, the duality provides a connection to stochastic optimal control which can then be used to compute the solution in a more systematic manner.

The proof of the main results is done in four steps. The first step is to reduce the problem to bounded claims. The second step is to represent the original robust hedging problem as a limit of robust hedging problems which live on a sequence of countable spaces. For these type of problems, robust hedging is the same as classical hedging, under the right choice of a probability measure. Thus we can apply the classical duality results for super-hedging of European options on a given probability space. The third step is to use the discrete structure and apply a standard min–max theorem (similar to the one used in [6]). The last step is to analyze the limit of the obtained prices in the discrete time markets. We combine methods from arbitrage-free pricing and limit theorems for stochastic processes.

The paper is organized as follows. The main results are formulated in the next section. In Sect. 3, the connection between the quasi sure approach and ours is proved. The two sections that follow are devoted to the proof of one inequality which implies the main results. The final section discusses a possible extension.

2 Preliminaries and main results

The financial market consists of a savings account which is normalized to unity $B_t \equiv 1$ by discounting and of a risky asset S_t, $t \in [0, T]$, where $T < \infty$ is the maturity date. Let $s := S_0 > 0$ be the initial stock price and without loss of generality, we set $s = 1$.

Denote by $\mathcal{C}^+[0, T]$ the set of all strictly positive functions $f : [0, T] \to \mathbb{R}_+$ which satisfy $f_0 = 1$. We assume that S_t is a continuous process. Then, any element of $\mathcal{C}^+[0, T]$ can be a possible path for the stock price process S. Let us emphasize that this the only assumption that we make on our financial market.

Denote by $\mathcal{D}[0, T]$ the space of all measurable functions $\upsilon : [0, T] \to \mathbb{R}$ with the norm $||\upsilon|| = \sup_{0 \le t \le T} |\upsilon_t|$. Let $G : \mathcal{D}[0, T] \to \mathbb{R}$ be a given deterministic map. We then consider a path dependent European option with the payoff

$$X = G(S), \tag{2.1}$$

where S is viewed as an element in $\mathcal{D}[0, T]$.

2.1 An assumption on the claim

Since our proof is through an approximation argument, we need the regularity of the pay-off functional G. Indeed, we first approximate the stock price process by piece-wise constant functions taking values in a finite set. We also discretize the jump times to obtain a countable set of possible price processes. This process necessitates a continuity assumption with respect to a Skorokhod type topology. Further discussion of this assumption is given in Remark 2.2. In particular, Asian and lookback type options satisfy the below condition. A possible generalization of our result to more general class of pay-offs is discussed in the final Sect. 6.

Let $\mathcal{D}_N[0, T]$ be the subset of $\mathcal{D}[0, T]$ that are piece-wise constant functions with N possible jumps i.e., $\upsilon \in \mathcal{D}_N[0, T]$ if and only if there exists a partition $t_0 = 0 < t_1 < t_2 < \cdots < t_N < T$ such that

$$\upsilon_t = \sum_{i=1}^{N} \upsilon_i \, \chi_{[t_{i-1}, t_i)}(t) + \upsilon_{N+1} \chi_{[t_N, T]}(t), \quad \text{where} \quad \upsilon_i := \upsilon_{t_{i-1}},$$

and we set χ_A be the characteristic function of the set A. We make the following standing assumption on G.

Assumption 2.1 There exists a constant $L > 0$ so that

$$|G(\omega) - G(\tilde{\omega})| \le L\|\omega - \tilde{\omega}\|, \quad \omega, \tilde{\omega} \in \mathcal{D}[0, T],$$

where as before, $|| \cdot ||$ is the sup norm.

Moreover, let $\upsilon, \tilde{\upsilon} \in \mathcal{D}_N[0, T]$ be such that $\upsilon_i = \tilde{\upsilon}_i$ for all $i = 1, \ldots, N$. Then,

$$|G(\upsilon) - G(\tilde{\upsilon})| \le L\|\upsilon\| \sum_{k=1}^{N} |\Delta t_k - \Delta \tilde{t}_k|,$$

where as usual $\Delta t_k := t_k - t_{k-1}$ and $\Delta \tilde{t}_k := \tilde{t}_k - \tilde{t}_{k-1}$.

Remark 2.2 In our setup, the process S represents the discounted stock price and $G(S)$ represents the discounted reward. Let $r > 0$ be the constant interest rate. Then, the payoff

$$G(S) := e^{-rT} H \left(e^{rT} S_T, \min_{0 \leq t \leq T} e^{rt} S_t, \max_{0 \leq t \leq T} e^{rt} S_t, \int_0^T e^{rt} S_t dt \right),$$

with a Lipschitz continuous function $H : \mathbb{R}^4 \to \mathbb{R}$ satisfies the above assumption.

The above condition on G is, in fact, a Lipschitz assumption with respect to a metric very similar to the Skorokhod one. However, it is *weaker* than to assume Lipschitz continuity with respect to the Skorokhod metric. Recall that this classical metric is given by

$$d(f, g) := \inf_{\lambda} \sup_{0 \leq t \leq T} \max \left(|f(t) - g(\lambda(t))|, |\lambda(t) - t| \right),$$

where the infimum is taken over all time changes. A time change is a strictly increasing continuous function which satisfy $\lambda(0) = 0$ and $\lambda(T) = T$. We refer the reader to Chapter 3 in [5] for more information. In particular, while $\int_0^T S_t dt$ is continuous with respect to the Skorokhod metric in $\mathcal{D}[0, T]$, it is not Lipschitz continuous with respect to this metric. Thus the above assumption is needed in order to include Asian options.

Moreover, from our proof of the main results it can be shown that Theorems 2.7 and 2.10 can be extended to payoffs of the form

$$e^{-rT} H \left(e^{rt_1} S_{t_1}, \ldots, e^{rt_k} S_{t_k}, \min_{0 \leq t \leq T} e^{rt} S_t, \max_{0 \leq t \leq T} e^{rt} S_t, \int_0^T e^{rt} S_t dt \right)$$

where H is Lipschitz and $0 < t_1 < \cdots < t_k \leq T$.   □

2.2 European calls

We assume that, at time zero, the investor is able to buy any call option with strike $K \geq 0$, for the price

$$C(K) := \int (x - K)^+ d\mu(x), \tag{2.2}$$

where μ is a given probability measure on $\mathbb{R}_+$. The measure μ is assumed to be derived from observed call prices that are liquidly traded in the market. One may also think of μ as describing the probabilistic belief (in the market) about the stock price distribution at time T. Then, an approximation argument implies that the price of a derivative security with the payoff $g(S_T)$ with a bounded, measurable g must be given by $\int g d\mu$. We then assume that this formula also holds for all $g \in \mathbb{L}^1(\mathbb{R}_+, \mu)$.

In particular, $C(0) = \int x d\mu(x)$. On the other hand the pay-off $C(0)$ is one stock. Hence, the value of $C(0)$ must be equal to the initial stock price S_0 which is normalized

to one. Therefore, although the probability measure μ is quite general, in view of our assumption (2.2) and arbitrage considerations, it should satisfy

$$C(0) = \int x d\mu(x) = S_0 = 1. \tag{2.3}$$

For technical reasons, we also assume that there exists $p > 1$ such that

$$\int x^p d\mu(x) < \infty. \tag{2.4}$$

2.3 Admissible portfolios

We continue by describing the continuous time trading in the underlying asset S. Since we do not assume any semi-martingale structure of the risky asset, this question is nontrivial. We adopt the path-wise approach and require that the trading strategy (in the risky asset) is of finite variation. Then, for any function $h : [0, T] \to \mathbb{R}$ of finite variation and continuous function $S \in \mathcal{C}[0, T]$, we use integration by parts to define

$$\int_0^t h_u dS_u := h_t S_t - h_0 S_0 - \int_0^t S_u dh_u,$$

where the last term in the above right hand side is the standard Stieltjes integral.

We are now ready to give the definition of semi-static portfolios and super-hedging. Recall the exponent p in (2.4).

Definition 2.3 1. We say that a map

$$\phi : A \subset \mathcal{D}[0, T] \to \mathcal{D}[0, T]$$

is *progressively measurable*, if for any $v, \tilde{v} \in A$,

$$v_u = \tilde{v}_u, \quad \forall u \in [0, t] \quad \Rightarrow \quad \phi(v)_t = \phi(\tilde{v})_t. \tag{2.5}$$

2. A *semi-static portfolio* is a pair $\pi := (g, \gamma)$, where $g \in \mathbb{L}^1(\mathbb{R}_+, \mu)$ and

$$\gamma : \mathcal{C}^+[0, T] \to \mathcal{D}[0, T]$$

is a progressively measurable map of bounded variation.

3. The corresponding *discounted portfolio value* is given by,

$$Z_t^\pi(S) = g(S_T)\chi_{\{t=T\}} + \int_0^t \gamma_u(S) dS_u, \quad t \in [0, T],$$

where χ_A is the indicator of the set A. A semi-static portfolio is *admissible*, if there exists $M > 0$ such that

$$Z_t^\pi(S) \geq -M\left(1 + \sup_{0 \leq u \leq t} S_u^p\right), \quad \forall t \in [0, T], \ S \in \mathcal{C}^+[0, T]. \quad (2.6)$$

4. An admissible semi-static portfolio is called *super-replicating*, if

$$Z_T^\pi(S) \geq G(S), \quad \forall S \in \mathcal{C}^+[0, T].$$

Namely, we require that for any possible value of the stock process, the portfolio value at maturity will be no less that the reward of the European claim.
5. The (minimal) *super-hedging cost* of G is defined by,

$$V(G) := \inf\left\{\int g d\mu : \exists \gamma \text{ such that } \pi := (g, \gamma) \text{ is super-replicating}\right\}.$$

Notice that the set of admissible portfolios depends on the exponent p which appears in the assumption (2.4). We suppress this possible dependence to simplify the exposition.

2.4 Martingale optimal transport

Since the dual formula refers to a probabilistic structure, we need to introduce that structure as well. Set $\Omega := \mathcal{C}^+[0, T]$ and let $\mathbb{S} = (\mathbb{S}_t)_{0 \leq t \leq T}$ be the canonical process given by $\mathbb{S}_t(\omega) := \omega_t$, for all $\omega \in \Omega$. Let $\mathcal{F}_t := \sigma(\mathbb{S}_s, 0 \leq s \leq t)$ be the canonical filtration (which is not right continuous).

The following class of probability measures are central to our results. Recall that we have normalized the stock prices to have initial value one. Therefore, the probability measures introduced below need to satisfy this condition as well.

Definition 2.4 A probability measure $\mathbb{Q}$ on the space $(\Omega, \mathcal{F})$ is a *martingale measure*, if the canonical process $(\mathbb{S}_t)_{t=0}^T$ is a local martingale with respect to $\mathbb{Q}$ and $\mathbb{S}_0 = 1$ $\mathbb{Q}$-a.s.

For a probability measure μ on $\mathbb{R}_+$, $\mathbb{M}_\mu$ is the set of all martingale measures $\mathbb{Q}$ such that the probability distribution of $\mathbb{S}_T$ under $\mathbb{Q}$ is equal to μ.

Note that if μ satisfies (2.3), then the canonical process $(\mathbb{S}_t)_{t=0}^T$ is a *martingale* (not only a local martingale) under any measure $\mathbb{Q} \in \mathbb{M}_\mu$. Indeed, a strict local martingale satisfies

$$1 = \mathbb{S}_0 > \mathbb{E}_\mathbb{Q}[\mathbb{S}_T] = \int x d\mu(x),$$

and it would be in contradiction with (2.3). We use $\mathbb{E}_\mathbb{Q}$ to denote the expectation with respect to $\mathbb{Q}$.

Remark 2.5 Observe that (2.3) yields that the set $\mathbb{M}_\mu$ is not empty. Indeed, consider a complete probability space $(\Omega^W, \mathcal{F}^W, P^W)$ together with a standard one-dimensional Brownian motion $(W_t)_{t=0}^\infty$, and the natural filtration $\mathcal{F}_t^W$ which is the completion of $\sigma\{W_s | s \leq t\}$. Then, there exists a function $f : \mathbb{R} \to \mathbb{R}_+$ such that the probability distribution of $f(W_T)$ is equal to μ. Define the martingale $M_t := E^W(f(W_T)|\mathcal{F}_t^W)$, $t \in [0, T]$. In view of (2.3), $M_0 = 1$. Since M is a Brownian martingale, it is continuous. Moreover, since μ has support on the positive real line, $f \geq 0$ and consequently, $M \geq 0$. Then, the distribution of M on the space Ω is an element in $\mathbb{M}_\mu$. This construction underpins Bass solution to the Skorokhod embedding problem (see [4]).

Remark 2.6 Clearly the duality is very closely related to fundamental theorem of asset pricing, which states the existence of a measure $\mathbb{Q} \in \mathbb{M}_\mu$. Since, as shown in the above remark such measures exist under our set of assumptions, the market considered in this paper is arbitrage-free. Then, a natural question that arises is whether our assumptions on the option prices and the measure μ can be replaced by the assumption of no-arbitrage. We do not address this very interesting question in this paper. However, several recent papers [3,8] study this question in discrete time.

The following is the main result of the paper. An outline of its proof is given in Sect. 2.6, below.

Theorem 2.7 *Assume that the European claim G satisfies the Assumption 2.1 and the probability measure μ satisfies (2.3) and (2.4). Then, the minimal super-hedging cost is given by*

$$V(G) = \sup_{\mathbb{Q}\in\mathbb{M}_\mu} \mathbb{E}_\mathbb{Q}[G(\mathbb{S})].$$

Remark 2.8 The above theorem provides a duality result for the robust semi-static hedging of a general pay-off. Many specific examples have been considered in the literature. Indeed, the initial paper of Hobson [22] explicitly provides the hedge for a lookback option. Similarly, using the random time change and Skorokhod embedding method [7,9,11], and several other papers analyze barrier options, lookback options and volatility options. Also, the path-wise proof of the Doob's maximal inequality given in [2] constructs an explicit portfolio which robustly hedges the power of the running maximum. We use this hedge in the proof of Lemma 4.1 as well.

Remark 2.9 One may consider the maximizer, if exists, of the expression

$$\sup_{\mathbb{Q}\in\mathbb{M}_\mu} \mathbb{E}_\mathbb{Q}[G(\mathbb{S})],$$

as the *optimal transport* of the initial probability measure $\nu = \delta_{\{1\}}$ to the final distribution μ. However, an additional constraint that the connection is a martingale is imposed. This in turn places a restriction on the measures, namely (2.3). The penalty function c is replaced by a more general functional G. In this context, one may also consider general initial distributions ν rather than Dirac measures. Then, the martingale measures with given marginals corresponds to the Kantorovich generalization of the mass transport problem.

The super-replication problem is also analogous to the *Kantorovich dual*. However, the dual elements reflect the fact that the cost functional depends on the whole path of the connection.

The reader may also consult [6] for a very clear discussion of the connection between robust hedging and optimal transport.

2.5 A discrete time approximation

Next we construct a special class of simple strategies which achieve asymptotically the super-hedging cost V.

For a positive integer N and any $S \in \mathcal{C}^+[0, T]$, set $\tau_0^{(N)}(S) = 0$. Then, recursively define

$$\tau_k^{(N)}(S) = \inf \left\{ t > \tau_{k-1}^{(N)}(S) : |S_t - S_{\tau_{k-1}^{(N)}(S)}| = \frac{1}{N} \right\} \wedge T, \qquad (2.7)$$

where we set $\tau_k^{(N)}(S) = T$, when the above set is empty. Also, define

$$H^{(N)}(S) = \min\{k \in \mathbb{N} : \tau_k^{(N)}(S) = T\}. \qquad (2.8)$$

Observe that for any $S \in \mathcal{C}^+[0, T]$, $H^{(N)}(S) < \infty$.

Denote by $\mathcal{A}_N$ the set of all portfolios for which the trading in the stock occurs only at the moments $0 = \tau_0^{(N)}(S) < \tau_1^{(N)}(S) < \cdots < \tau_{H^{(N)}(S)}^{(N)}(S) = T$. Formally, $\pi := (g, \gamma) \in \mathcal{A}_N$, if it is progressively measurable in the sense of (2.5) and it is of the form

$$\gamma_t(S) = \sum_{k=0}^{H^{(N)}(S)-1} \gamma_k(S) \chi_{(\tau_k^{(N)}(S), \tau_{k+1}^{(N)}(S)]}(t),$$

for some $\gamma_k(S)$'s. Note that, $\gamma_k(S)$ can depend on S only through its values up to time $\tau_k^{(N)}(S)$, so that γ_t is progressively measurable. Set

$$V_N(G) := \inf \left\{ \int g d\mu : \exists \gamma \text{ such that } \pi := (g, \gamma) \in \mathcal{A}_N \text{ is super-replicating} \right\}.$$

It is clear that for any integer $N \geq 1$, $V_N(G) \geq V(G)$. The following result proves the convergence to $V(G)$. This approximation result is the second main result of this paper. Also, it is the key analytical step in the proof of duality.

Theorem 2.10 *Under the assumptions of Theorem 2.7,*

$$\lim_{N \to \infty} V_N(G) = V(G).$$

2.6 Proofs of Theorems 2.7 and 2.10

Since $V_N \geq V$, Theorems 2.7 and 2.10 would follow from the following two inequalities,

$$\limsup_{N \to \infty} V_N(G) \leq \sup_{\mathbb{Q} \in \mathbb{M}_\mu} \mathbb{E}_{\mathbb{Q}}[G(\mathbb{S})] \tag{2.9}$$

and

$$V(G) \geq \sup_{\mathbb{Q} \in \mathbb{M}_\mu} \mathbb{E}_{\mathbb{Q}}[G(\mathbb{S})]. \tag{2.10}$$

The first inequality is the difficult one and it will be proved in Sects. 4 and 5. The second inequality is simpler and we provide its proof here.

Let $\mathbb{Q} \in \mathbb{M}_\mu$ and let $\pi = (g, \gamma)$ be super-replicating. Since γ is progressively measurable in the sense of (2.5), the stochastic integral

$$\int_0^t \gamma_u(\mathbb{S}) d\mathbb{S}_u$$

is defined with respect to $\mathbb{Q}$. Also $\mathbb{Q}$ is a martingale measure. Hence, the above stochastic integral is a $\mathbb{Q}$ local- martingale. Moreover, from (2.6) we have,

$$\int_0^t \gamma_u(\mathbb{S}) d\mathbb{S}_u \geq -M\left(1 + \sup_{0 \leq u \leq t} |\mathbb{S}_t|^p\right), \quad t \in [0, T].$$

Also in view of (2.4) and the Doob–Kolmogorov inequality for the martingale $\mathbb{S}_t$,

$$\mathbb{E}_{\mathbb{Q}} \sup_{0 \leq t \leq T} |\mathbb{S}_t|^p \leq C_p \mathbb{E}_{\mathbb{Q}} |\mathbb{S}_T|^p = C_p \int |x|^p d\mu < \infty.$$

Therefore, $\mathbb{E}_{\mathbb{Q}} \int_0^T \gamma_u(\mathbb{S}) d\mathbb{S}_u \leq 0$. Since π is super-replicating, we conclude that

$$\mathbb{E}_{\mathbb{Q}}[G(\mathbb{S})] \leq \mathbb{E}_{\mathbb{Q}}\left(\int_0^T \gamma_u(\mathbb{S}) d\mathbb{S}_u + g(\mathbb{S}_T)\right) \leq \mathbb{E}_{\mathbb{Q}}[g(\mathbb{S}_T)] = \int g d\mu,$$

where in the last equality we again used the fact that the distribution of $\mathbb{S}_T$ under $\mathbb{Q}$ is equal to μ. This completes the proof of the lower bound. Together with (2.9), which will be proved later, it also completes the proofs of the theorems. □

3 Quasi sure approach and full duality

An alternate approach to define robust hedging is to use the notion of quasi sure super-hedging as was done in [21,38]. Let us briefly recall this notion. Let $\mathcal{Q}$ be the set of all martingale measures $\mathbb{Q}$ on the canonical space $\mathcal{C}^+[0, T]$ under which the canonical process $\mathbb{S}$ satisfies $\mathbb{S}_0 = 1$, $\mathbb{Q}$-a.s., has quadratic variation and satisfies $\mathbb{E}_\mathbb{Q} \sup_{0 \le t \le T} \mathbb{S}_t < \infty$. In this market, an admissible hedging strategy (or a portfolio) is defined as a pair $\pi = (g, \gamma)$, where $g \in \mathbb{L}_1(\mathbb{R}_+, \mu)$ and γ is a progressively measurable process such that the stochastic integral

$$\int\limits_0^t \gamma_u d\mathbb{S}_u, \quad t \in [0, T]$$

exists for any probability measure $\mathbb{Q} \in \mathcal{Q}$ and satisfies (2.6) $\mathbb{Q}$-a.s. We refer the reader to [38] for a complete characterization of this class. In particular, one does not restrict the trading strategies to be of bounded variation. A portfolio $\pi = (g, \gamma)$ is called an (admissible) *quasi-sure super-hedge*, provided that

$$g(\mathbb{S}_T) + \int\limits_0^T \gamma_u d\mathbb{S}_u \ge G(\mathbb{S}), \quad \mathbb{Q} \text{ a.s.,}$$

for all $\mathbb{Q} \in \mathcal{Q}$. Then, the minimal super-hedging cost is given by

$$V_{qs}(G) := \inf \left\{ \int g d\mu : \exists \gamma \text{ such that } \pi := (g, \gamma) \text{ is a quasi-sure super-hedge} \right\}.$$

Clearly,

$$V(G) \ge V_{qs}(G).$$

From simple arbitrage arguments it follows that

$$V_{qs}(G) \ge \inf_{\lambda \in \mathbb{L}^1(\mathbb{R}_+, \mu)} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_\mathbb{Q} \left(G(\mathbb{S}) - \lambda(\mathbb{S}_T) + \int \lambda d\mu \right)$$

where we set $\mathbb{E}_\mathbb{Q} \xi \equiv -\infty$, if $\mathbb{E}_\mathbb{Q} \xi^- = \infty$. Since inf sup $\ge$ sup inf, the above two inequalities yield,

$$V(G) \ge V_{qs}(G)$$

$$\ge \inf_{\lambda \in \mathbb{L}^1(\mathbb{R}_+, \mu)} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_\mathbb{Q} \left(G(\mathbb{S}) - \lambda(\mathbb{S}_T) + \int \lambda d\mu \right)$$

$$\ge \sup_{\mathbb{Q} \in \mathcal{Q}} \inf_{\lambda \in \mathbb{L}^1(\mathbb{R}_+, \mu)} \mathbb{E}_\mathbb{Q} \left(G(\mathbb{S}) - \lambda(\mathbb{S}_T) + \int \lambda d\mu \right).$$

Now if $\mathbb{Q} \in \mathbb{M}_\mu$, then the two terms involving λ are equal. So we first restrict the measures to the set $\mathbb{M}_\mu$ and then use Theorem 2.7. The result is

$$
V(G) \geq V_{qs}(G) \geq \inf_{\lambda \in \mathbb{L}^1(\mathbb{R}_+, \mu)} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_\mathbb{Q} \left(G(\mathbb{S}) - \lambda(\mathbb{S}_T) + \int \lambda \, d\mu \right)
$$

$$
\geq \sup_{\mathbb{Q} \in \mathcal{Q}} \inf_{\lambda \in \mathbb{L}^1(\mathbb{R}_+, \mu)} \mathbb{E}_\mathbb{Q} \left(G(\mathbb{S}) - \lambda(\mathbb{S}_T) + \int \lambda \, d\mu \right)
$$

$$
\geq \sup_{\mathbb{Q} \in \mathbb{M}_\mu} \mathbb{E}_\mathbb{Q} \left[G(\mathbb{S}) \right] = V(G).
$$

Hence, all terms in the above are equal. We summarize this in the following which can be seen as the full duality.

Proposition 3.1 *Assume that the European claim G satisfies Assumption* 2.1 *and the probability measure μ satisfies* (2.3)*,* (2.4)*. Then,*

$$
V(G) = V_{qs}(G) = \sup_{\mathbb{Q} \in \mathbb{M}_\mu} \mathbb{E}_\mathbb{Q} \left[G(\mathbb{S}) \right]
$$

$$
= \inf_{\lambda \in \mathbb{L}^1(\mathbb{R}_+, \mu)} \sup_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_\mathbb{Q} \left(G(\mathbb{S}) - \lambda(\mathbb{S}_T) + \int \lambda \, d\mu \right)
$$

$$
= \sup_{\mathbb{Q} \in \mathcal{Q}} \inf_{\lambda \in \mathbb{L}^1(\mathbb{R}_+, \mu)} \mathbb{E}_\mathbb{Q} \left(G(\mathbb{S}) - \lambda(\mathbb{S}_T) + \int \lambda \, d\mu \right).
$$

4 Proof of the main results

The rest of the paper is devoted to the proof of (2.9).

4.1 Reduction to bounded claims

The following result will be used in two places in the paper. The first place is Lemma 4.2 where we reduce the problem to claims that are bounded from above. The other place is Lemma 4.8.

Consider a claim with pay-off

$$
\alpha_K(S) := \|S\| \, \chi_{\{\|S\| \geq K\}} + \frac{\|S\|}{K}.
$$

Recall that $V_N(\alpha_K)$ is defined in Sect. 2.5.

Lemma 4.1

$$
\limsup_{K \to \infty} \ \limsup_{N \to \infty} V_N(\alpha_K) = 0.
$$

Proof In this proof, we always assume that $N > K > 1$. Let $\tau_k = \tau_k^{(N)}(S)$ and $n = H^{(N)}(S)$ be as in (2.7), (2.8), respectively, and set

$$\theta := \theta_N^{(K)}(S) = \min\{k : S_{\tau_k} \geq K - 1\} \wedge n.$$

Set $c_p := p/(p-1)$ where p as in (2.4). We define a portfolio $(g^{(N,K)}, \gamma^{(N,K)}) \in \mathcal{A}_N$ as follows. For $t \in (\tau_k, \tau_{k+1}]$ and $k = 0, 1, \ldots, n-1$, let

$$\gamma_t^{(N,K)}(S) = \gamma_{\tau_k}^{(N,K)}(S) = -\frac{p^2}{K(p-1)}\left(\max_{0 \leq i \leq k} S_{\tau_i}^{p-1}\right) - \frac{p^2}{(p-1)}\chi_{\{k \geq \theta\}}\left(\max_{\theta \leq i \leq k} S_{\tau_i}^{p-1}\right),$$

$$g^{(N,K)}(x) = \frac{1}{K}(1 + ((c_p x)^p - c_p)^+) + ((c_p x)^p - (c_p(K-1))^p)^+ + \frac{2}{N}.$$

We use Proposition 2.1 in [2] and the inequality $x < 1 + x^p$, $x \in \mathbb{R}_+$, to conclude that for any $t \in [0, T]$

$$g^{(N,K)}(S_t) + \int_0^t \gamma_u^{(N,K)} dS_u \geq \frac{\bar{S}_t}{K} + \bar{S}_t \, \chi_{\{\bar{S}_t \geq K\}},$$

where

$$\bar{S}_t := \max_{0 \leq u \leq t} S_u.$$

Therefore, $\pi^{(N,K)} := (g^{N,K}, \gamma^{(N,K)})$ satisfies (2.6) and super-replicates α_K. Hence,

$$V_N(\alpha_K) \leq \int g^{(N,K)} d\mu.$$

Also, in view of (2.4),

$$\limsup_{K \to \infty} \limsup_{N \to \infty} \int g^{(N,K)} d\mu = 0.$$

These two inequalities complete the proof of the lemma.   □

A corollary of the above estimate is the following reduction to claims that are bounded from above.

Lemma 4.2 *If suffices to prove* (2.9) *for claims G that are non-negative, bounded from above and satisfying Assumption* 2.1.

Proof We proceed in two steps. First suppose that (2.9) holds for nonnegative claims that are bounded from above. Then, the conclusions of Theorems 2.7 and 2.10 also hold for such claims.

Now let G be a non-negative claim satisfying Assumption 2.1. For $K > 0$, set

$$G_K := G \wedge K.$$

Then, G_K is bounded and (2.9) holds for G_K. Therefore,

$$\limsup_{N \to \infty} V_N(G_K) \leq \sup_{\mathbb{Q} \in \mathbb{M}_\mu} \mathbb{E}_{\mathbb{Q}}[G_K(\mathbb{S})] \leq \sup_{\mathbb{Q} \in \mathbb{M}_\mu} \mathbb{E}_{\mathbb{Q}}[G(\mathbb{S})].$$

In view of Assumption 2.1,

$$G(S) \leq G(0) + L\|S\|.$$

Hence, the set $\{G(S) \geq K\}$ is included in the set $\{L\|S\| + G(0) \geq K\}$ and

$$G \leq G_K + (L\|S\| + G(0) - K)\,\chi_{\{L\|S\|+G(0)\geq K\}}.$$

By the linearity of the market, this inequality implies that

$$V_N(G) \leq V_N(G_K) + V_N((L\|S\| + G(0) - K)\chi_{\{L\|S\|+G(0)\geq K\}}).$$

Moreover, in view of the previous lemma,

$$\limsup_{K \to \infty} \limsup_{N \to \infty} V_N((L\|S\| + G(0) - K)\chi_{\{L\|S\|+G(0)\geq K\}}) = 0.$$

Using these, we conclude that

$$\limsup_{N \to \infty} V_N(G) \leq \sup_{\mathbb{Q} \in \mathbb{M}_\mu} \mathbb{E}_{\mathbb{Q}}[G(\mathbb{S})].$$

Hence, (2.9) holds for all functions that are non-negative and satisfy Assumption 2.1. By adding an appropriate constant this results extends to all claims that are bounded from below and satisfying Assumption 2.1.

Now suppose that G is a general function that satisfies Assumption 2.1. For $c > 0$, set

$$\check{G}_c := G \vee (-c).$$

Then, $\check{G}$ is bounded from below and (2.9), Theorems 2.7 and 2.10 holds, i.e.,

$$\limsup_{N \to \infty} V_N(G) \leq \limsup_{N \to \infty} V_N(\check{G}_c) = \sup_{\mathbb{Q} \in \mathbb{M}_\mu} \mathbb{E}_{\mathbb{Q}}[\check{G}_c(\mathbb{S})].$$

By Assumption 2.1, $\check{G}_c(S) \leq G(S) + \check{e}_c(S)$ where the error function is

$$\check{e}_c(S) := (L\|S\| - G(0) - c)\chi_{\{L\|S\|-G(0)-c\geq 0\}}.$$

Since $\check{e}_c \geq 0$ and it satisfies the Assumption 2.1,

$$\sup_{\mathbb{Q}\in\mathbb{M}_\mu} \mathbb{E}_\mathbb{Q}[\check{e}_c(\mathbb{S})] = V(\check{e}_c) = \lim_{N\to\infty} V_N(\hat{e}_c).$$

In view of Lemma 4.1,

$$\limsup_{c\to\infty} \sup_{\mathbb{Q}\in\mathbb{M}_\mu} \mathbb{E}_\mathbb{Q}[\check{e}_c(\mathbb{S})] = \limsup_{c\to\infty} \limsup_{N\to\infty} V_N(\check{e}_c) = 0.$$

We combine the above inequalities to conclude that

$$\begin{aligned}
\limsup_{N\to\infty} V_N(G) &\leq \limsup_{c\to\infty} \sup_{\mathbb{Q}\in\mathbb{M}_\mu} \mathbb{E}_\mathbb{Q}[\check{G}_c(\mathbb{S})] \\
&\leq \sup_{\mathbb{Q}\in\mathbb{M}_\mu} \mathbb{E}_\mathbb{Q}[G(\mathbb{S})] + \limsup_{c\to\infty} \sup_{\mathbb{Q}\in\mathbb{M}_\mu} \mathbb{E}_\mathbb{Q}[\check{e}_c(\mathbb{S})] \\
&= \sup_{\mathbb{Q}\in\mathbb{M}_\mu} \mathbb{E}_\mathbb{Q}[G(\mathbb{S})].
\end{aligned}$$

This exactly (2.9).   □

4.2 A countable class of piecewise constant functions

In this section, we provide a piece-wise constant approximation of any continuous function S. Fix a positive integer N. For any $S \in \mathcal{C}^+[0, T]$, let $\tau_k^{(N)}(S)$ and $H^{(N)}(S)$ be the times defined in (2.7) and (2.8), respectively. To simplify the notation, we suppress their dependence on S and N and also set

$$n = H^{(N)}(S). \tag{4.1}$$

We first define the obvious piecewise constant approximation $\hat{S} = \hat{S}^{(N)}(S)$ using these times. Indeed, set

$$\hat{S}_t := \sum_{k=0}^{n-1} S_{\tau_k}\chi_{[\tau_k,\tau_{k+1})}(t) + \left[S_{\tau_{n-1}} + \frac{1}{N}sign(S_T - S_{\tau_{n-1}})\right]\chi_{\{T\}}(t). \tag{4.2}$$

The function, that takes S to $\hat{S}$ is a map of $\mathcal{C}^+[0, T]$ into the set of all functions with values in the target set

$$A^{(N)} = \{i/N : i = 0, 1, 2, \ldots, \}.$$

Indeed, $\hat{S}$ is behind the definition of the approximating costs V_N. However, this set of functions is not countable as the jump times are not restricted to a countable set. So, we provide yet another approximation by restricting the jump times as well.

Let $\hat{\Omega} := \mathbb{D}[0, T]$ be the space of all right continuous functions $f : [0, T] \to \mathbb{R}_+$ with left–hand limits (*càdlàg* functions). For integers N, k, let

$$U_k^{(N)} := \{i/(2^k N) : i = 1, 2, \ldots, \} \cup \{1/(i 2^k N) : i = 1, 2, \ldots, \},$$

be the sets of possible differences between two consecutive jump times. Next, we define subsets $\mathbb{D}^{(N)}$ of $\mathbb{D}[0, T]$.

Definition 4.3 A function $f \in \mathbb{D}[0, T]$ belongs to $\mathbb{D}^{(N)}$, if it satisfies the followings,

1. $f(0) \in \{1 - 1/N, 1 + 1/N\}$,
2. f is piecewise constant with jumps at times $t_1, \ldots, t_n$, where

$$t_0 = 0 < t_1 < t_2 < \cdots < t_n < T,$$

3. for any $k = 1, \ldots, n$, $|f(t_k) - f(t_{k-1})| = 1/N$,
4. for any $k = 1, \ldots, n$, $t_k - t_{k-1} \in U_k^{(N)}$.

We emphasize, in the fourth condition, the dependence of the set $U_k^{(N)}$ on k. So as k gets larger, jump times take values in a finer grid. Also, for technical reasons we will need that the functions value at 0 will be equal to $1 \pm 1/N$ but not 1.

We continue by defining an approximation of a generic stock price process S,

$$F^{(N)} : \mathcal{C}^+[0, T] \to \mathbb{D}^{(N)},$$

as follow. Recall $\tau_k = \tau_k^{(N)}(S)$, $n = H^{(N)}(S)$ from above and also from (2.7), (2.8). Set $\hat{\tau}_0 := 0$, $\hat{\tau}_n = T$ and for $k = 1, \ldots, n - 1$, define

$$\hat{\tau}_k := \sum_{i=1}^{k} \Delta \hat{\tau}_i,$$

$$\Delta \hat{\tau}_i = \max\{\Delta t \in U_i^{(N)} : \Delta t < \Delta \tau_i = \tau_i - \tau_{i-1}\}, \qquad i = 1, \ldots, n - 1.$$

Clearly, $0 = \hat{\tau}_0 < \hat{\tau}_1 < \cdots < \hat{\tau}_{n-1} < \hat{\tau}_n = T$ and $\hat{\tau}_k < \tau_k$ for all $k = 1, \ldots, n - 1$.

We are now ready to define $F^{(N)}(S)$. For $n = 1$, set

$$F^{(N)}(S) \equiv 1 + \frac{1}{N} sign(S_T - 1),$$

and for $n > 1$, define

$$F_t^{(N)}(S) = \sum_{k=1}^{n-1} S_{\tau_k} \chi_{[\hat{\tau}_{k-1}, \hat{\tau}_k)}(t)$$

$$+ \left(S_{\tau_{n-1}} + \frac{1}{N} sign(S_T - S_{\tau_{n-1}}) \right) \chi_{[\hat{\tau}_{n-1}, T]}(t). \qquad (4.3)$$

Observe that the value of the kth jump of the process $F^{(N)}(S)$ equals to the value of the $(k+1)$-th jump of the discretization $\hat{S}$ of the original process S. Indeed, for $n > 2$,

$$F_{\hat{\tau}_m}^{(N)} - F_{\hat{\tau}_{m-1}}^{(N)} = S_{\tau_{m+1}} - S_{\tau_m}, \qquad \forall m = 1, \ldots, n-2, \qquad (4.4)$$

and for $n \geq 2$,

$$F_{\hat{\tau}_{n-1}}^{(N)} - F_{\hat{\tau}_{n-2}}^{(N)} = \frac{1}{N} sign \left(S_T - S_{\tau_{n-1}} \right).$$

This shift is essential in order to deal with some delicate questions of adaptedness and predictability. We also recall that the jump times of $\hat{S}$ are the random times τ_k's while the jump times of $F^{(N)}(S)$ are $\hat{\tau}_k$'s and that all these times depend both on N and S. Moreover, by construction, $F^{(N)}(S) \in \mathbb{D}^{(N)}$. But, it may not be progressively measurable as defined in (2.5). However, we use $F^{(N)}$ only to lift progressively measurable maps defined on $\mathbb{D}^{(N)}$ to the initial space $\Omega = \mathcal{C}^+[0, T]$ and this yields progressively measurable maps on Ω. This procedure is defined and the measurability is proved in Lemma 4.7, below.

The following lemma shows that $F^{(N)}$ is close to S in the sense of Assumption 2.1. Let us emphasize that the following result is a consequence of the particular structure of $\mathbb{D}^{(N)}$ and in particular $U_k^{(N)}$'s.

Lemma 4.4 *Let $F^{(N)}$ be the map defined in* (4.3). *For any G satisfying the Assumption* 2.1 *with the constant L,*

$$|G(S) - G(F^{(N)}(S))| \leq \frac{4L\|S\|}{N}, \qquad \forall S \in \mathcal{C}^+[0, T].$$

Proof Set

$$\hat{F} := \hat{F}_t^{(N)}(S) := \sum_{k=0}^{n-1} S_{\tau_k} \chi_{[\hat{\tau}_k, \hat{\tau}_{k+1})}(t) + \left[S_{\tau_{n-1}} + \frac{1}{N} sign(S_T - S_{\tau_{n-1}}) \right] \chi_{\{T\}}(t).$$

Observe that $\hat{S}$ of (4.2) and $\hat{F}$ are like the functions υ and $\tilde{\upsilon}$ in that Assumption 2.1. Hence,

$$|G(\hat{S}) - G(\hat{F})| \leq L\|S\| \sum_{k=1}^{n} |\Delta \tau_k - \Delta \hat{\tau}_k|.$$

For $k < n$,

$$\Delta \hat{\tau}_k = \max\{\Delta t \in U_k^{(N)} \; : \; \Delta t < \Delta \tau_k \}.$$

The definition of $U_k^{(N)}$ implies that

$$0 \le \Delta \tau_k - \Delta \hat{\tau}_k \le \frac{1}{2^k N}, \quad k = 1, \ldots, n-1.$$

Therefore,

$$\sum_{k=1}^{n-1} |\Delta \tau_k - \Delta \hat{\tau}_k| \le \sum_{k=1}^{\infty} \frac{1}{2^k N} = \frac{1}{N}. \tag{4.5}$$

Combining the above inequalities, we arrive at

$$|G(\hat{S}) - G(\hat{F})| \le \frac{L\|S\|}{N}.$$

Set $F = F^{(N)}(S)$ and directly estimate that

$$\begin{aligned}
|G(S) - G(F)| &\le |G(S) - G(\hat{S})| + |G(\hat{S}) - G(\hat{F})| + |G(\hat{F}) - G(F)| \\
&\le L\|S - \hat{S}\| + \frac{L\|S\|}{N} + |G(\hat{F}) - G(F)| \\
&= \frac{3L\|S\|}{N} + |G(\hat{F}) - G(F)|.
\end{aligned}$$

Finally, we observe that by construction,

$$\|\hat{F} - F\| \le \frac{1}{N}, \quad \Rightarrow \quad |G(F) - G(\hat{F})| \le \frac{L}{N}.$$

The above inequalities completes the proof of the lemma. □

Remark 4.5 The proof of the above Lemma provides one of the reasons behind the particular structure of $U_k^{(N)}$. Indeed, (4.5) is a key estimate which provides a uniform upper bound for the sum of the differences over k. Since there is no upper bound on k, the approximating set $U_k^{(N)}$ for the k-th difference must depend on k. Moreover, it should have a summable structure over k. That explains the terms 2^k.

On the other hand, the reason for the part $\{1/(i2^k N) : i = 1, 2, \ldots\}$ in the definition of $U_k^{(N)}$ is to make sure that $\Delta \hat{\tau}_k > 0$. For probabilistic reasons (i.e. adaptability), we want $\hat{\tau}_k < \tau_k$. This forces us to approximate $\Delta \tau_k$ by $\Delta \hat{\tau}_k$ from below. This and $\Delta \hat{\tau}_k > 0$ would be possible only if $U_k^{(N)}$ has a subsequence converging to zero.

Hence, different sets of $U_k^{(N)}$'s are also possible provided that they have these two properties. □

4.3 A countable probabilistic structure

An essential step in the proof of (2.9) is a duality result for probabilistic problems. We first introduce this structure and then relate it to the problem V_N.

As before, let $\hat{\Omega} := \mathbb{D}[0, T]$ be the space of all right continuous functions $f :$ $[0, T] \rightarrow \mathbb{R}_+$ with left–hand limits (*càdlàg* functions). Denote by $\hat{\mathbb{S}} = (\hat{\mathbb{S}}_t)_{0 \leq t \leq T}$ the canonical process on the space $\hat{\Omega}$.

The set $\mathbb{D}^{(N)}$ defined in Definition 4.3 is a countable subset of $\hat{\Omega}$. We choose any probability measure $\hat{\mathbb{P}}^{(N)}$ on $\hat{\Omega}$ which satisfies $\hat{\mathbb{P}}^{(N)}(\mathbb{D}^{(N)}) = 1$ and $\hat{\mathbb{P}}^{(N)}(\{f\}) > 0$ for all $f \in \mathbb{D}^{(N)}$. Let $\hat{\mathcal{F}}_t^{(N)}$, $t \in [0, T]$ be the filtration generated by the process $\hat{\mathbb{S}}$ and contains $\hat{\mathbb{P}}^{(N)}$ null sets. Under the measure $\hat{\mathbb{P}}^{(N)}$, the canonical map $\hat{\mathbb{S}}$ has finitely many jumps. Let

$$0 = \hat{\tau}_0(\hat{\mathbb{S}}) < \hat{\tau}_1(\hat{\mathbb{S}}) < \cdots < \hat{\tau}_{\hat{H}(\hat{\mathbb{S}})}(\hat{\mathbb{S}}) < T,$$

be the jump times of $\hat{\mathbb{S}}$. Note that in Definition 4.3, the final jump time is always strictly less than T.

A trading strategy on the filtered probability space $(\hat{\Omega}, \hat{\mathcal{F}}^{(N)}, (\hat{\mathcal{F}}_t^{(N)})_{t=0}^T, \hat{\mathbb{P}}^{(N)})$ is a predictable stochastic process $(\hat{\gamma}_t)_{t=0}^T$. Thus, it is a function $\hat{\gamma} : \mathbb{D}[0, T] \rightarrow \mathcal{D}[0, T]$. Let $a \in \mathcal{D}[0, T]$ be such that $a \notin \hat{\gamma}(\mathbb{D}^{(N)})$. Define a map $\phi : \mathbb{D}[0, T] \rightarrow \mathcal{D}[0, T]$, by $\phi(\omega) = \hat{\gamma}(\omega)$ if $\omega \in \mathbb{D}^{(N)}$, and equal to a otherwise. Clearly, $\hat{\mathbb{P}}^{(N)}$ almost surely, $\hat{\gamma} = \phi(\hat{\mathbb{S}})$. Also, since $\hat{\mathbb{P}}^{(N)}$ is non-zero on every point in $\mathbb{D}^{(N)}$, the definition of the predictable sigma algebra implies that ϕ is a predictable map. Namely, for any $v, \tilde{v} \in \mathbb{D}[0, T]$ and $t \in [0, T]$

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

Indeed, arguing by contraposition, if there were $t \in [0, T]$ and $v, \tilde{v} \in \mathbb{D}^{(N)}$ such that $v_u = \tilde{v}_u$ for all $u \in [0, t)$ and $\phi(v)_t \neq \phi(\tilde{v})_t$. Then, we would conclude that the event $\{\hat{\gamma}_t = \phi(v)_t\} \notin \hat{\mathcal{F}}_{t-}^{(N)}$. However, this would be in contradiction with the predictability of the process $\hat{\gamma}$. (Recall that $\mathcal{F}_{t-}^{(N)}$ is the smallest σ–algebra which contains $\mathcal{F}_s^{(N)}$ for any $s < t$). Hence, any predictable process $\hat{\gamma}$ has a version ϕ that is progressively measurable in the sense of Definition 2.3. In what follows, we always use this progressively measurable version of any predictable process. In particular, the following can be seen as the probabilistic counterpart of the Definition 2.3.

Definition 4.6 1. A (probabilistic) *semi-static portfolio* is a pair $(h, \hat{\gamma})$ such that $\hat{\gamma} : \mathbb{D}[0, T] \rightarrow \mathcal{D}[0, T]$ is predictable and the stochastic integral $\int_0^{\cdot} \hat{\gamma}_u d\hat{\mathbb{S}}_u$ exists (with respect to the measure $\mathbb{P}^{(N)}$), and $h : A^{(N)} \rightarrow \mathbb{R}$.

2. A semi-static portfolio is $\hat{\mathbb{P}}^{(N)}$-*admissible*, if h is bounded and there exists $M > 0$ such that

$$\int_0^t \hat{\gamma}_u d\hat{\mathbb{S}}_u \geq -M, \quad \hat{\mathbb{P}}^{(N)} - a.s., \ t \in [0, T]. \tag{4.6}$$

3. An admissible semi-static portfolio is $\hat{\mathbb{P}}^{(N)}$-*super-replicating* G, if

$$h(\hat{\mathbb{S}}_T) + \int_0^T \hat{\gamma}_u d\hat{\mathbb{S}}_u \geq G(\hat{\mathbb{S}}), \quad \hat{\mathbb{P}}^{(N)} - a.s. \tag{4.7}$$

4.4 Approximating μ

Recall the set $\mathcal{A}_N$ of portfolios used in the definition of V_N in Sect. 2.5.

Next we provide a connection between the probabilistic super-replication and the discrete robust problem. However, the option h in the Definition 4.6 above is defined only on $A^{(N)}$ while the static part of the hedges in $\mathcal{A}_N$ are functions defined on $\mathbb{R}_+$. So for a given $h : A^{(N)} \to \mathbb{R}$, we define the following operator

$$g^{(N)} := \mathcal{L}^{(N)}(h) : \mathbb{R}_+ \to \mathbb{R}$$

by

$$g^{(N)}(x) := (1 + \lfloor Nx \rfloor - Nx)h(\lfloor Nx \rfloor/N) + (Nx - \lfloor Nx \rfloor)h((1 + \lfloor Nx \rfloor)/N),$$

where for a real number r, $\lfloor r \rfloor$ is the largest integer that is not larger than r.

Next, define a measure $\mu^{(N)}$ on the set $A^{(N)}$ by

$$\mu^{(N)}(\{0\}) := \int_{[0,1/N)} (1 - Nx) \, d\mu(x)$$

and for any positive integer k,

$$\mu^{(N)}(\{k/N\}) := \int_{[(k-1)/N,k/N)} (Nx + 1 - k) \, d\mu(x) + \int_{[k/N,(k+1)/N)} (1 + k - Nx) \, d\mu(x).$$

This construction has the following important property. For any bounded function $h : A^{(N)} \to \mathbb{R}$, let $g^{(N)} = \mathcal{L}^{(N)}(h)$ be as above. Then,

$$\int h \, d\mu^{(N)} = \int g^{(N)} d\mu. \tag{4.8}$$

In particular, by taking $h \equiv 1$, we conclude that $\mu^{(N)}$ is a probability measure. Also, since for continuous h, $g^{(N)}$ converges pointwise to h, one may directly show (by Lebesgue's dominated convergence theorem) that $\mu^{(N)}$ converges weakly to μ.

4.5 Probabilistic super-replication

Recall the probabilistic super-replication problem introduced in Definition 4.6. Let G be a European claim as before and N be a positive integer. Then, the probabilistic super-replication problem is given by,

$$\hat{V}_N(G) = \inf\left\{\int h\, d\mu^{(N)} : \exists\, \hat{\gamma} \text{ s.t. } (h, \hat{\gamma}) \text{ is a } \hat{\mathbb{P}}^{(N)} \text{ admissible super hedge of } G\right\}.$$

We continue by establishing a connection between the probabilistic super hedging $\hat{V}_N$ and the discrete robust problem V_N. Suppose that we are given a probabilistic semi-static portfolio $\hat{\pi} = (h, \hat{\gamma})$ in the sense of Definition 4.6. We lift this portfolio to a semi-static portfolio $\pi^{(N)} = (g^{(N)}, \gamma^{(N)}) \in \mathcal{A}_N$. Indeed, let $g^{(N)} = \mathcal{L}^{(N)}(h)$ be as in Sect. 4.4 and define $\gamma^{(N)} : \mathcal{C}^+[0, T] \to \mathcal{D}[0, T]$ by

$$\gamma_t^{(N)}(S) = \sum_{k=1}^{n-1} \hat{\gamma}_{\hat{\tau}_k}(F^{(N)}(S))\chi_{(\tau_k, \tau_{k+1}]}(t),$$

where $\tau_k = \tau_k^{(N)}(S)$ are as in (2.7), n is as in (4.1) and $F^{(N)}(S)$, $\hat{\tau}_k := \hat{\tau}_k(S)$ are as in (4.3). Note that the random integer n is the number of crossings of magnitude of no less than $1/N$. Moreover, by construction it is exactly one more than the number of jumps of $F^{(N)}$. Also notice that

$$\gamma_t^{(N)}(S) = 0, \quad \forall\, t \in [0, \tau_1].$$

Lemma 4.7 *For any probabilistic semi-static portfolio $(h, \hat{\gamma})$, $\gamma^{(N)}$ defined above is progressively measurable in the sense of* (2.5).

Proof Let $S, \tilde{S} \in \mathcal{C}^+[0, T]$ and $t \in [0, T]$ be such that $S_u = \tilde{S}_u$ for all $u \leq t$. We need to show that

$$\gamma_t^{(N)}(S) = \gamma_t^{(N)}(\tilde{S}).$$

Since the above clearly holds for $t = 0$ and $t = T$, we may assume that $t \in (0, T)$. Set

$$k_t(S) := k_t^{(N)}(S) := \min\{i \geq 1, : \tau_i^{(N)} \geq t\} - 1,$$

so that $0 \leq k_t(S) < n$ and

$$t \in (\tau_{k_t(S)}^{(N)}, \tau_{k_t(S)+1}^{(N)}].$$

It is clear that $k_t(S) = k_t(\tilde{S})$. If $k_t(S) = k_t(\tilde{S}) = 0$, then $\gamma_t^{(N)}(S) = \gamma_t^{(N)}(\tilde{S}) = 0$. So we assume that $k_t(S) > 0$ and use the definition of $\hat{\tau}_k$ to conclude that

$$\theta := \hat{\tau}_{k_t(S)} = \hat{\tau}_{k_t(\tilde{S})}(\tilde{S}).$$

Since $0 < k_t(S) < n$, we have $n > 1$ and $F_t^{(N)}$ is given by (4.3), i.e.,

$$F_t^{(N)}(S) = \sum_{k=1}^{n-1} S_{\tau_k} \chi_{[\hat{\tau}_{k-1}, \hat{\tau}_k)}(t) + \left(S_{\tau_{n-1}} + \frac{1}{N} sign(S_T - S_{\tau_{n-1}}) \right) \chi_{[\hat{\tau}_{n-1}, T]}(t).$$

Now, for any $u < \theta = \hat{\tau}_{k_t(S)} = \hat{\tau}_{k_t(\tilde{S})}(\tilde{S})$, the above definition implies that

$$F_u^{(N)}(S) = S_{\tau_k}, \quad F_u^{(N)}(\tilde{S}) = \tilde{S}_{\tau_k}, \quad \text{for some} \quad k \le k_t(S) = k_t(\tilde{S}).$$

Since by definition $\tau_{k_t(S)}(S) < t$, we conclude that

$$F_u^{(N)}(S) = F_u^{(N)}(\tilde{S}), \quad \forall\, u \in [0, \theta).$$

Therefore, by the predictability of $\hat{\gamma}$ we have $\gamma_t^{(N)}(S) = \gamma_t^{(N)}(\tilde{S})$. □

The following lemma provides a natural and a crucial connection between the probabilistic super-replication and the discrete robust problem.

Recall the set $\mathcal{A}_N$ of portfolios used in the definition of V_N in Sect. 2.5.

Lemma 4.8 *Suppose G is bounded from above and satisfies the Assumption* 2.1. *Then,*

$$\limsup_{N \to \infty} V_N(G) \le \limsup_{N \to \infty} \hat{V}_N(G).$$

Proof Set

$$G^{(N)}(S) := G(S) - \frac{5L\|S\|}{N}.$$

We first show that

$$V_N(G^{(N)}) \le \hat{V}_N(G).$$

To prove the above inequality, suppose that a portfolio $(h, \hat{\gamma})$ is a $\hat{\mathbb{P}}^{(N)}$-admissible super hedge of G. Then it suffices to construct a map $\gamma^{(N)} : \mathcal{C}^+[0, T] \to \mathcal{D}[0, T]$ and $g^{(N)} : \mathbb{R}_+ \to \mathbb{R}$ such that the semi-static portfolio $\pi^{(N)} := (g^{(N)}, \gamma^{(N)})$ is admissible, belongs to $\mathcal{A}_N$ and super-replicates $G^{(N)}$ in the sense of Definition 2.3.

Let $g^{(N)} = \mathcal{L}^{(N)}(h)$ be as in Sect. 4.4 and $\gamma^{(N)}$ be the probabilistic portfolio considered in Lemma 4.7. We claim that $\pi^{(N)}$ is the desired portfolio. In view of Lemma 4.7, we need to show that $\pi^{(N)}$ is in $\mathcal{A}_N$ and super-replicates the $G^{(N)}$ in the sense of Definition 2.3.

To simplify the notation, we set $F := F^{(N)}(S)$.

Admissibility of $\gamma^{(N)}$ By construction trading is only at the random times τ_k's. Therefore, $\pi^{(N)} \in \mathcal{A}_N$ provided that it satisfies the lower bound (2.6) for every $t \in [0, T]$. We first claim that for any $S \in \mathcal{C}^+[0, T]$ and for every $k \le n - 1$,

$$\int_0^{\tau_k} \gamma_u^{(N)}(S) dS_u = \int_{[0,\hat{\tau}_{k-1}]} \hat{\gamma}_u(F) dF_u.$$

Since $\gamma^{(N)} \equiv 0$ on $[0, \tau_1]$, the above trivially holds for $k = 1$. So we assume that $1 < k \leq n-1$. In particular, $n > 2$. Then, we use (4.4) and the definitions to compute that

$$\int_{[0,\hat{\tau}_{k-1}]} \hat{\gamma}_u(F) dF_u = \sum_{m=1}^{k-1} \hat{\gamma}_{\hat{\tau}_m}(F)(F_{\hat{\tau}_m} - F_{\hat{\tau}_{m-1}}) = \sum_{m=1}^{k-1} \hat{\gamma}_{\hat{\tau}_m}(F)(S_{\tau_{m+1}} - S_{\tau_m})$$

$$= \sum_{m=1}^{k} \gamma_{\tau_{m+1}}^{(N)}(F)(S_{\tau_{m+1}} - S_{\tau_m}) = \int_{\tau_1}^{\tau_k} \gamma_u^{(N)}(S) dS_u$$

$$= \int_0^{\tau_k} \gamma_u^{(N)}(S) dS_u.$$

The last identity follows from the fact that $\gamma^{(N)}$ is zero on the interval $[0, \tau_1]$.

Now, for a given $t \in [0, T)$ and $S \in C^+[0, T]$, let $k \leq n-1$ be the largest integer so that $\tau_k \leq t$. Construct a function $\tilde{F} \in \mathbb{D}^{(N)}$ by,

$$\tilde{F}_{[0,\hat{\tau}_k)} = F_{[0,\hat{\tau}_k)}, \ (\text{i.e.,} \ \tilde{F}_u = F_u, \ \forall \, u \in [0, \hat{\tau}_k),)$$

and

$$\tilde{F}_u = 2F_{\hat{\tau}_{k-1}} - F_{\hat{\tau}_k}, \ u \geq \hat{\tau}_k.$$

Note that the constructed function $\tilde{F}$ depends on S and N, since both F and the stopping times τ_k depend on them. But we suppress these dependences. Since

$$\tilde{F}_{\hat{\tau}_k} - \tilde{F}_{\hat{\tau}_{k-1}} = -[F_{\hat{\tau}_k} - F_{\hat{\tau}_{k-1}}] = \pm 1/N,$$

and since

$$|S_t - S_{\tau_k}| \leq 1/N,$$

there exists $\lambda \in [0, 1]$ (depending on t, N, S) such that

$$S_t - S_{\tau_k} = \lambda(F_{\hat{\tau}_k} - F_{\hat{\tau}_{k-1}}) + (1 - \lambda)(\tilde{F}_{\hat{\tau}_k} - \tilde{F}_{\hat{\tau}_{k-1}}).$$

Since F and $\tilde{F}$ agree on $[0, \hat{\tau}_k)$ and $\hat{\gamma}$ is predictable, $\hat{\gamma}_u(F) = \hat{\gamma}_u(\tilde{F})$ for all $u \leq \hat{\tau}_k$. Also, for $u \in (\tau_k, t) \subset (\tau_k, \tau_{k+1})$, $\gamma_u^{(N)}(S) = \hat{\gamma}_{\hat{\tau}_k}(F)$ and

$$\int_0^t \gamma_u^{(N)}(S)dS_u = \int_0^{\tau_k} \gamma_u^{(N)}(S)dS_u + \int_{\tau_k}^t \gamma_u^{(N)}(S)dS_u$$

$$= \int_{[0,\hat{\tau}_{k-1}]} \hat{\gamma}_u(F)dF_u + \hat{\gamma}_{\hat{\tau}_k}(F)[S_t - S_{\tau_k}].$$

Since F is piece-wise constant with jumps only at the stopping times $\hat{\tau}_i$'s,

$$\int_{[0,\hat{\tau}_k]} \hat{\gamma}_u(F)dF_u = \int_{[0,\hat{\tau}_{k-1}]} \hat{\gamma}_u(F)dF_u + \int_{(\hat{\tau}_{k-1},\hat{\tau}_k]} \hat{\gamma}_u(F)dF_u$$

$$= \int_{[0,\hat{\tau}_{k-1}]} \hat{\gamma}_u(F)dF_u + \hat{\gamma}_{\hat{\tau}_k}(F)[F_{\hat{\tau}_k} - F_{\hat{\tau}_{k-1}}].$$

We calculate the same integral for $\tilde{F}$ using the fact that $F = \tilde{F}$ on $[0, \tau_k)$. The result is

$$\int_{[0,\hat{\tau}_k]} \hat{\gamma}_u(\tilde{F})d\tilde{F}_u = \int_{[0,\hat{\tau}_{k-1}]} \hat{\gamma}_u(F)dF_u + \hat{\gamma}_{\hat{\tau}_k}(F)[\tilde{F}_{\hat{\tau}_k} - \tilde{F}_{\hat{\tau}_{k-1}}].$$

Therefore,

$$\int_0^t \gamma_u^{(N)}(S)dS_u = \lambda \int_{[0,\hat{\tau}_k]} \hat{\gamma}_u(F)dF_u + (1-\lambda) \int_{[0,\hat{\tau}_k]} \hat{\gamma}_u(\tilde{F})d\tilde{F}_u.$$

Since $F, \tilde{F} \in \mathbb{D}^{(N)}$ and $\hat{\mathbb{P}}^{(N)}(F), \hat{\mathbb{P}}^{(N)}(\tilde{F}) > 0$, (4.6) imply that

$$\int_{[0,\hat{\tau}_k]} \hat{\gamma}_u(F)dF_u \geq -M, \quad \text{and} \quad \int_{[0,\hat{\tau}_k]} \hat{\gamma}_u(\tilde{F})d\tilde{F}_u \geq -M.$$

Hence, $\gamma^{(N)}$ satisfies (2.6) and $\pi^{(N)} \in \mathcal{A}_N$.

Super-replication We need to show that

$$g^{(N)}(S_T) + \int_0^T \gamma_u^{(N)}(S)dS_u \geq G^{(N)}(S).$$

We proceed almost exactly as in the proof of admissibility. Again we define a modification $\bar{F} \in \mathbb{D}^{(N)}$ by $\bar{F}_{[0,\hat{\tau}_{n-2})} = F_{[0,\hat{\tau}_{n-2})}$ and $\bar{F}_u = \bar{F}_{\hat{\tau}_{n-2}}$ for $u \geq \hat{\tau}_{n-2}$. Set

$$\hat{\lambda} := N|S_T - S_{\tau_{n-1}}|.$$

Then $\hat{\lambda} \in [0, 1]$ and by the construction of $g^{(N)}$,

$$g^{(N)}(S_T) = \hat{\lambda}h(F_T) + (1 - \hat{\lambda})h(\bar{F}_T).$$

Hence,

$$g^{(N)}(S_T) + \int_0^T \gamma_u^{(N)}(S)dS_u$$

$$= \hat{\lambda}\left[h(F_T) + \int_0^T \hat{\gamma}_u(F)dF_u\right] + (1 - \hat{\lambda})\left[h(\bar{F}_T) + \int_0^T \hat{\gamma}_u(\bar{F})d\bar{F}_u\right]$$

$$\geq \hat{\lambda}G(F) + (1 - \hat{\lambda})G(\bar{F}).$$

Since $\|F - \bar{F}\| \leq 1/N$, Assumption 2.1 and Lemma 4.4 imply that

$$\left|G(S) - G(\bar{F})\right| \geq |G(S) - G(F)| + \left|G(F) - G(\bar{F})\right| \leq \frac{5L\|S\|}{N}.$$

Consequently,

$$\hat{\lambda}G(F) + (1 - \hat{\lambda})G(\bar{F}) \geq G^{(N)}(S)$$

and we conclude that $\pi^{(N)}$ is super-replication $G^{(N)}$.

Completion of the proof We have shown that

$$V_N(G - 5L\|S\|/N) \leq \hat{V}_N(G).$$

Moreover, the linearity of the market yields that the super-replication cost is subadditive. Hence,

$$V_N(G) \leq V_N(5L\|S\|/N) + V_N(G - 5L\|S\|/N).$$

Therefore,

$$V_N(G) \leq V_N(5L\|S\|/N) + \hat{V}_N(G).$$

Finally, by Lemma 4.1,

$$\limsup_{N \to \infty} V_N(5L\|S\|/N) = 0.$$

We use the above inequalities to complete the proof of the lemma.   □

4.6 First duality

Recall the countable set $\mathbb{D}^{(N)} \subset \hat{\Omega}$ and its probabilistic structure that were introduced in Sect. 4.3. We consider two classes of measures on this set.

Definition 4.9 1. We say that a probability measure $\mathbb{Q}$ on the space $(\hat{\Omega}, \hat{\mathcal{F}})$ is a *martingale measure* if the canonical process $(\hat{\mathbb{S}}_t)_{t=0}^T$ is a local martingale with respect to $\mathbb{Q}$.
2. $\mathbb{M}_N$ is the set of all martingale measures that are supported on $\mathbb{D}^{(N)}$.
3. For a given $K > 0$, $\mathbb{M}_N^{(K)}$ is the set of all measures $\mathbb{Q} \in \mathbb{M}_N$ that satisfy

$$\sum_{k=0}^{\infty} |\mathbb{Q}(\hat{\mathbb{S}}_T = k/N) - \mu^{(N)}(\{k/N\})| < \frac{K}{N}. \qquad (4.9)$$

$\square$

The following follows from known duality results. We will combine it with Lemma 4.8 and Proposition 5.1, which will be proved in the next section to complete the proof of the inequality (2.9).

Lemma 4.10 *Suppose that $G \geq 0$ is bounded from above by K and satisfies the Assumption 2.1. Then, for any positive integer N,*

$$\hat{V}_N(G) \leq \left(\sup_{\mathbb{Q} \in \mathbb{M}_N^{(K)}} \mathbb{E}_{\mathbb{Q}}[G(\hat{\mathbb{S}})] \right)^+,$$

where the right hand side of the above inequality is zero when the set $\mathbb{M}_N^{(K)} = \emptyset$ is empty.

Proof Fix N and define the set

$$\mathcal{Z} = \mathcal{Z}^{(N)} := \{h : A^{(N)} \to \mathbb{R} : |h(z)| \leq N, \ \forall z\}.$$

Set

$$\mathbb{V} := \inf_{h \in \mathcal{Z}} \sup_{\mathbb{Q} \in \mathbb{M}_N} \left(\mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{S}}) - h(\hat{\mathbb{S}}_T)) + \int h d\mu^{(N)} \right).$$

Clearly, for any $\epsilon > 0$, there exists $h_\epsilon \in \mathcal{Z}$ such that

$$\sup_{\mathbb{Q} \in \mathbb{M}_N} \mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{S}}) - h_\epsilon(\hat{\mathbb{S}}_T)) + \int h_\epsilon d\mu^{(N)} \leq \mathbb{V} + \epsilon.$$

By construction, the support of the measure $\hat{\mathbb{P}}^{(N)}$ is $\mathbb{D}^{(N)}$. Also all elements of $\mathbb{D}^{(N)}$ are piece-wise constant. Therefore, under $\hat{\mathbb{P}}^{(N)}$ the canonical process $\hat{\mathbb{S}}$ is trivially a

semi-martingale and we may use the results of the seminal paper [17]. In particular, by Theorem 5.7 in [17], for

$$x_\epsilon = \sup_{\mathbb{Q} \in \mathbb{M}_N} \mathbb{E}_\mathbb{Q}(G(\hat{\mathbb{S}}) - h_\epsilon(\hat{\mathbb{S}}_T)) + \epsilon,$$

there exists an *admissible* portfolio strategy $\hat{\gamma}$ such that

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

Therefore, $(h_\epsilon + x_\epsilon, \hat{\gamma})$ satisfies (4.6)–(4.7), consequently

$$\hat{V}_N(G) \leq x_\epsilon + \int h_\epsilon d\mu^{(N)} \leq \sup_{\mathbb{Q} \in \mathbb{M}_N} \mathbb{E}_\mathbb{Q}(G(\hat{\mathbb{S}}) - h_\epsilon(\hat{\mathbb{S}}_T)) + \int h_\epsilon d\mu^{(N)} + \epsilon \leq \mathbb{V} + 2\epsilon.$$

We now let ϵ to zero to conclude that

$$\hat{V}_N(G) \leq \inf_{h \in Z} \sup_{\mathbb{Q} \in \mathbb{M}_N} \left(\mathbb{E}_\mathbb{Q}(G(\hat{\mathbb{S}}) - h(\hat{\mathbb{S}}_T)) + \int h d\mu^{(N)} \right). \qquad (4.10)$$

The next step is to interchange the order of the above infimum and supremum. Consider the vector space $\mathbb{R}^{A^{(N)}}$ of all functions $f : A^{(N)} \to \mathbb{R}$ equipped with the topology of point-wise convergence. Clearly, this space is locally convex. Also, since $A^{(N)}$ is countable, $\mathcal{Z}$ is a compact subset of $\mathbb{R}^{A^{(N)}}$. The set $\mathbb{M}_N$ can be naturally considered as a convex subspace of the vector space $\mathbb{R}^{\mathbb{D}^{(N)}}$.

Now, define the function $\mathcal{G} : \mathcal{Z} \times \mathbb{M}_N \to \mathbb{R}$, by

$$\mathcal{G}(h, \mathbb{Q}) = \mathbb{E}_\mathbb{Q}(G(\hat{\mathbb{S}}) - h(\hat{\mathbb{S}}_T)) + \int h d\mu^{(N)}.$$

Notice that $\mathcal{G}$ is affine in each of the variables. From the bounded convergence theorem, it follows that $\mathcal{G}$ is continuous in the first variable. Next, we apply the min-max theorem, Theorem 45.8 in [37] to $\mathcal{G}$. The result is,

$$\inf_{h \in \mathcal{Z}} \sup_{\mathbb{Q} \in \mathbb{M}_N} \mathcal{G}(h, \mathbb{Q}) = \sup_{\mathbb{Q} \in \mathbb{M}_N} \inf_{h \in \mathcal{Z}} \mathcal{G}(h, \mathbb{Q}).$$

This together with (4.10) yields,

$$\hat{V}_N(G) \leq \sup_{\mathbb{Q} \in \mathbb{M}_N} \inf_{h \in \mathcal{Z}} \left(\mathbb{E}_\mathbb{Q}(G(\hat{\mathbb{S}}) - h(\hat{\mathbb{S}}_T)) + \int h d\mu^{(N)} \right). \qquad (4.11)$$

Finally, for any measure $\mathbb{Q} \in \mathbb{M}_N$, define $h^\mathbb{Q} \in \mathcal{Z}$ by

$$h^\mathbb{Q}(k/N) = N sign(\mathbb{Q}(\hat{\mathbb{S}}_T = k/N) - \mu^{(N)}(\{k/N\})), \quad k = 0, 1, \ldots.$$

In view of (4.11),

$$\hat{V}_N(G) \leq \sup_{\mathbb{Q}\in\mathbb{M}_N} \left(\mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{S}})) + \int h^{\mathbb{Q}} d\mu^{(N)} - \mathbb{E}_{\mathbb{Q}} h^{\mathbb{Q}}(\hat{\mathbb{S}}_T) \right)$$

$$= \sup_{\mathbb{Q}\in\mathbb{M}_N} \left\{ \mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{S}})) - N \sum_{k=0}^{\infty} |\mathbb{Q}(\hat{\mathbb{S}}_T = k/N) - \mu^{(N)}(\{k/N\})| \right\}$$

Suppose that $\mathbb{Q} \notin \mathbb{M}_N^{(K)}$. Then,

$$N \sum_{k=0}^{\infty} |\mathbb{Q}(\hat{\mathbb{S}}_T = k/N) - \mu^{(N)}(\{k/N\})| \geq K.$$

Since G is bounded by K, this implies that

$$\mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{S}})) - N \sum_{k=0}^{\infty} |\mathbb{Q}(\hat{\mathbb{S}}_T = k/N) - \mu^{(N)}(\{k/N\})| \leq 0.$$

Hence,

$$\hat{V}_N(G) \leq \left(\sup_{\mathbb{Q}\in\mathbb{M}_N^{(K)}} \left\{ \mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{S}})) - N \sum_{k=0}^{\infty} |\mathbb{Q}(\hat{\mathbb{S}}_T = k/N) - \mu^{(N)}(\{k/N\})| \right\} \right)^+$$

$$\leq \left(\sup_{\mathbb{Q}\in\mathbb{M}_N^{(K)}} \mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{S}})) \right)^+.$$

$\square$

5 Approximation of Martingale measures

In this final section, we prove the asymptotic connection between the approximating martingale measures $\mathbb{M}_N^{(K)}$ defined in Definition 4.9 and the continuous martingale measures $\mathbb{M}_\mu$ satisfying the marginal constraint at the final time, defined in Definition 2.4.

The following proposition completes the proof of the inequality (2.9) and consequently the proofs of the main theorems when the claim $G \geq 0$ is bounded from above. The general case then follows from Lemma 4.2.

Proposition 5.1 *Suppose that $G \geq 0$ is bounded from above by K and satisfies the Assumption 2.1. Assume that μ satisfies (2.3)–(2.4). Then*

$$\limsup_{N\to\infty} \left(\sup_{\mathbb{Q}\in\mathbb{M}_N^{(K)}} \mathbb{E}_{\mathbb{Q}}[G(\hat{\mathbb{S}})] \right)^+ \leq \sup_{\mathbb{Q}\in\mathbb{M}_\mu} \mathbb{E}_{\mathbb{Q}}[G(\mathbb{S})].$$

We prove the above proposition not through a compactness argument as one may expect. Instead, we show that any given measure $\mathbb{Q} \in \mathbb{M}_N^{(K)}$ has a lifted version in $\mathbb{M}_\mu$ that is close to $\mathbb{Q}$ in some sense. The set $\mathbb{M}_\mu \neq \emptyset$ is not empty, thus $\sup_{\mathbb{Q} \in \mathbb{M}_\mu} \mathbb{E}_\mathbb{Q}[G(\mathbb{S})] \geq 0$. Therefore, without loss of generality, we can assume that for sufficiently large N the set $\mathbb{M}_N^{(K)}$ is not empty, otherwise the proposition is trivially satisfied. Hence, the above proposition is a direct consequence of the below lemma.

Recall the Lipschitz constant L in Assumption 2.1.

Lemma 5.2 *Under the hypothesis of Proposition 5.1, there exists a function $f_K(\epsilon, N)$ satisfying,*

$$\lim_{\epsilon \downarrow 0} \lim_{N \to \infty} f_K(\epsilon, N) = 0$$

so that for any $\hat{\mathbb{Q}} \in \mathbb{M}_N^{(K)}$ and $\epsilon > 0$,

$$\mathbb{E}_{\hat{\mathbb{Q}}}[G(\hat{\mathbb{S}})] \leq f_K(\epsilon, N) + \sup_{\mathbb{Q} \in \mathbb{M}_\mu} \mathbb{E}_\mathbb{Q}[G(\mathbb{S})].$$

Proof Fix $\epsilon \in (0, 1)$, a positive integer N and $\hat{\mathbb{Q}} \in \mathbb{M}_N^{(K)}$. Recall that G is bounded from above by K.

Shift of the initial value Denote by $\mathbb{D}_1^{(N)}$ the set of all functions $f \in \mathbb{D}[0, T]$ which satisfy $f(0) = 1$ and the conditions 2–4 in Definition 4.3. Define a map $H : \mathbb{D}^{(N)} \to \mathbb{D}_1^{(N)}$ by $H(f) = f + 1 - f(0)$. Consider the measure $\mathbb{Q}_1 = H \circ \hat{\mathbb{Q}}$, clearly $\mathbb{Q}_1$ is a martingale measure.

Jump times Since the probability measure $\hat{\mathbb{Q}}$ is supported on the set $\mathbb{D}^{(N)}$, the canonical process $\hat{\mathbb{S}}$ is a purely jump process under $\mathbb{Q}_1$, with a finite number of jumps. Introduce the jump times by setting $\tau_0 = 0$ and for $k > 0$,

$$\tau_k = \inf\{t > \tau_{k-1} : \hat{\mathbb{S}}_t \neq \hat{\mathbb{S}}_{t\text{-}}\} \wedge T.$$

Next we introduce the largest random time

$$\hat{N} := \min\{k : \tau_k = T\}.$$

Then, $\hat{N} < \infty$ almost surely and consequently, there exists a deterministic positive integer m (depending on ϵ) such that

$$\mathbb{Q}_1(\hat{N} > m) < \epsilon. \tag{5.1}$$

By the definition of the set $\mathbb{D}^{(N)}$, there is a decreasing sequence of strictly positive numbers $t_k \downarrow 0$, with $t_1 = T$, such that for $i = 1, \ldots, m$,

$$\tau_i - \tau_{i-1} \in \{t_k\}_{k=1}^\infty \cup \{0\}, \quad \mathbb{Q}_1 - a.s.$$

Wiener space Let $(\Omega^W, \mathcal{F}^W, P^W)$ be a complete probability space together with a standard $m+2$–dimensional Brownian motion $\left\{ W_t = \left(W_t^{(1)}, W_t^{(2)}, \ldots, W_t^{(m+2)} \right) \right\}_{t=0}^{\infty}$, and the natural filtration $\mathcal{F}_t^W = \sigma\{W_s | s \leq t\}$. The next step is to construct a martingale Z on the Brownian probability space $(\Omega^W, \mathcal{F}^W, P^W)$ together with a sequence of stopping times (with respect to the Brownian filtration) $\sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_m$ such that the distribution (under the Wiener measure P^W) of the random vector $(\sigma_1, \ldots, \sigma_m, Z_{\sigma_1}, \ldots, Z_{\sigma_m})$ is equal to the distribution of the random vector $(\tau_1, \ldots, \tau_m, \hat{\mathbb{S}}_{\tau_1}, \ldots, \hat{\mathbb{S}}_{\tau_m})$ under the measure $\mathbb{Q}_1$. Namely,

$$((\sigma_1, \ldots, \sigma_m, Z_{\sigma_1}, \ldots, Z_{\sigma_m}), P^W) = ((\tau_1, \ldots, \tau_m, \hat{\mathbb{S}}_{\tau_1}, \ldots, \hat{\mathbb{S}}_{\tau_m}), \mathbb{Q}_1). \quad (5.2)$$

The construction is done by induction, at each step k we construct the stopping time σ_k and Z_{σ_k} such that the conditional probability is the same as in the case of the canonical process $\hat{\mathbb{S}}$ under the measure $\mathbb{Q}_1$.

Construction of σ's and Z For an integer n and given $x_1, \ldots, x_n$, introduce the notation

$$\vec{x}_n := (x_1, \ldots, x_n).$$

Also set

$$\mathbb{T} := \{t_k\}_{k=1}^{\infty}.$$

For $k = 1, \ldots, m$, define the functions $\Psi_k, \Phi_k : \mathbb{T}^k \times \{-1, 1\}^{k-1} \to [0, 1]$ by

$$\Psi_k(\vec{\alpha}_k; \vec{\beta}_{k-1}) := \mathbb{Q}_1(\tau_k - \tau_{k-1} \geq \alpha_k \mid A), \quad (5.3)$$

where

$$A := \{\tau_i - \tau_{i-1} = \alpha_i, \ \hat{\mathbb{S}}_{\tau_i} - \hat{\mathbb{S}}_{\tau_{i-1}} = \beta_i/N, \ i \leq k - 1\},$$

and

$$\Phi_k(\vec{\alpha}_k; \vec{\beta}_{k-1}) = \mathbb{Q}_1(\hat{\mathbb{S}}_{\tau_k} - \hat{\mathbb{S}}_{\tau_{k-1}} = 1/N \mid B), \quad (5.4)$$

where

$$B = \{\tau_k < T, \tau_j - \tau_{j-1} = \alpha_j, \ \hat{\mathbb{S}}_{\tau_i} - \hat{\mathbb{S}}_{\tau_{i-1}} = \beta_i/N, \ j \leq k, i \leq k - 1\}.$$

As usual we set $\mathbb{Q}_1(\cdot | \emptyset) \equiv 0$. Next, for $k \leq m$, we define the maps $\Gamma_k, \Theta_k : \mathbb{T}^k \times \{-1, 1\}^{k-1} \to [-\infty, \infty]$, as the unique solutions of the following equations,

$$P^W(W_{\alpha_k}^{(1)} < \Gamma_k(\vec{\alpha}_k; \vec{\beta}_{k-1})) = \Phi_k(\vec{\alpha}_k; \vec{\beta}_{k-1}), \quad (5.5)$$

and

$$P^W(W_{t_l}^{(1)} - W_{t_{l+1}}^{(1)} < \Theta_k(\vec{\alpha}_k; \vec{\beta}_{k-1})) = \frac{\Psi_k(\vec{\alpha}_{k-1}, t_l; \vec{\beta}_{k-1})}{\Psi_k(\vec{\alpha}_{k-1}, t_{l+1}; \vec{\beta}_{k-1})}, \qquad (5.6)$$

where $l \in \mathbb{N}$ is given by $\alpha_k = t_l \in \mathbb{T}$. From the definitions it follows that $\Psi_k(\vec{\alpha}_{k-1}, t_l; \vec{\beta}_{k-1}) \le \Psi_k(\vec{\alpha}_{k-1}, t_{l+1}; \vec{\beta}_{k-1})$. Thus if $\Psi_k(\vec{\alpha}_{k-1}, t_{l+1}; \vec{\beta}_{k-1}) = 0$ for some l, then also $\Psi_k(\vec{\alpha}_{k-1}, t_l; \vec{\beta}_{k-1}) = 0$. We set $0/0 \equiv 0$.

Set $\sigma_0 \equiv 0$ and define the random variables $\sigma_1, \ldots, \sigma_m, Y_1, \ldots, Y_m$ by the following recursive relations

$$\sigma_1 = \sum_{k=1}^{\infty} t_k \chi_{\{W_{t_k}^{(1)} - W_{t_{k+1}}^{(1)} > \Theta_1(t_k)\}} \prod_{j=k+1}^{\infty} \chi_{\{W_{t_j}^{(1)} - W_{t_{j+1}}^{(1)} < \Theta_1(t_j)\}}, \qquad (5.7)$$

$$Y_1 = 2\chi_{\{W_{\{\sigma_1\}}^{(2)} > \Gamma_1(\sigma_1)\}} - 1,$$

and for $i > 1$

$$\sigma_i = \sigma_{i-1} + \Delta_i$$

$$Y_i = \chi_{\{\sigma_i < T\}} \left(2\chi_{\{W_{\sigma_i}^{(i+1)} - W_{\sigma_{i-1}}^{(i+1)} > \Gamma_i(\vec{\Delta\sigma}_i, \vec{Y}_{i-1})\}} - 1 \right),$$

where $\Delta_i = t_k$ on the set $A_i \cap B_{i,k} \cap C_{i,k}$ and zero otherwise. These sets are given by,

$$A_i := \{|Y_{i-1}| > 0\},$$

$$B_{i,k} := \{W_{t_k+\sigma_{i-1}}^{(1)} - W_{t_{k+1}+\sigma_{i-1}}^{(1)} > \Theta_i(\vec{\sigma}_{i-1}, t_k; \vec{Y}_{i-1})\},$$

$$C_{i,k} := \bigcap_{j=k+1}^{\infty} \{W_{t_j+\sigma_{i-1}}^{(1)} - W_{t_{j+1}+\sigma_{i-1}}^{(1)} < \Theta_i(\vec{\Delta\sigma}_{i-1}, t_j; \vec{Y}_{i-1})\}.$$

Since t_k is decreasing with $t_1 = T$, $\sigma_1 \le \sigma_2 \le \cdots \le \sigma_m$ and they are stopping times with respect to the Brownian filtration. Let $k \le m$ and $(\vec{\alpha}_k; \vec{\beta}_{k-1}) \in \mathbb{T}^k \times \{-1, 1\}^{k-1}$. There exists $m \in \mathbb{N}$ such that $\alpha_k = t_m \in \mathbb{T}$. From (5.7) to (5.8), the strong Markov property and the independency of the Brownian motion increments it follows that

$$P^W(\sigma_k - \sigma_{k-1} \ge \alpha_k | (\vec{\Delta\sigma}_{k-1}; \vec{Y}_{k-1}) = (\vec{\alpha}_{k-1}; \vec{\beta}_{k-1}))$$

$$= P^W \left(\bigcap_{j=m}^{\infty} (W_{t_j+\sigma_{k-1}}^{(1)} - W_{t_{j+1}+\sigma_{k-1}}^{(1)} < \Theta_k(\vec{\alpha}_{k-1}, t_j; \vec{\beta}_{k-1})) \right)$$

$$= \prod_{j=m}^{\infty} P^W(W_{t_j+\sigma_{k-1}}^{(1)} - W_{t_{j+1}+\sigma_{k-1}}^{(1)} < \Theta_k(\vec{\alpha}_{k-1}, t_j; \vec{Y}_{k-1}))$$

$$= \Psi_k(\vec{\alpha}_k, \vec{\beta}_{k-1}), \qquad (5.8)$$

where the last equality follows from (5.6) and the fact that

$$\lim_{l \to \infty} \Psi_k(\alpha_1, \ldots, \alpha_{k-1}, t_l, \beta_1, \ldots, \beta_{k-1}) = 1.$$

Similarly, from (5.5) and (5.8), we have

$$
\begin{aligned}
& P^W (Y_k = 1 | \sigma_k < T, \vec{\Delta \sigma}_k = \vec{\alpha}_k, \vec{Y}_{k-1} = \vec{\beta}_{k-1}) \\
& = P^W \left(W^{(k+1)}_{\sum_{i=1}^{k} \alpha_i} - W^{(k+1)}_{\sum_{i=1}^{k-1} \alpha_i} < \Gamma_k(\vec{\alpha}_k; \vec{\beta}_{k-1}) \right) \\
& = \Phi_k(\vec{\alpha}_k; \vec{\beta}_{k-1}).
\end{aligned}
\tag{5.9}
$$

Using (5.3)–(5.4) and (5.8)–(5.9), we conclude that

$$
\left(\left(\vec{\sigma}_m; \frac{1}{N} \vec{Y}_m \right), P^W \right) = ((\vec{\tau}_m; \vec{\Delta \hat{\mathbb{S}}}_m), \mathbb{Q}_1)
$$

where $\Delta \hat{\mathbb{S}}_k = \hat{\mathbb{S}}_{\tau_k} - \hat{\mathbb{S}}_{\tau_{k-1}}, k \leq m$.

Continuous martingale Set

$$
Z_t = 1 + \frac{1}{N} E^W \left(\sum_{i=1}^{m} Y_i | \mathcal{F}_t^W \right), \quad t \in [0, T].
\tag{5.10}
$$

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

$$
Z_{\sigma_k} = 1 + \frac{1}{N} \sum_{i=1}^{k} Y_i, \quad P^W \text{a.s.}, \quad k \leq m.
\tag{5.11}
$$

By the construction of Y and σ's, we conclude that (5.2) holds with the process Z.

Measure in $\mathbb{M}_\mu$ The next step in the proof is to modify the martingale Z in such way that the distribution of the modified martingale is an element of $\mathbb{M}_\mu$. For any two probability measures ν_1, ν_2 on $\mathbb{R}$, Prokhorov's metric is defined by

$$
\begin{aligned}
d(\nu_1, \nu_2) = \inf\{ \delta > 0 : \nu_1(A) \leq \nu_2(A^\delta) + \delta \quad \text{and} \\
\nu_2(A) \leq \nu_1(A^\delta) + \delta, \quad \forall A \in \mathcal{B}(\mathbb{R}) \},
\end{aligned}
$$

where $\mathcal{B}(\mathbb{R})$ is the set of all Borel sets $A \subset \mathbb{R}$ and $A^\delta := \bigcup_{x \in A} (x - \delta, x + \delta)$ is the δ–neighborhood of A. It is well known that convergence in the Prokhorov metric is equivalent to weak convergence, (for more details on Prokhorov's metric see [35], Chapter 3, Section 7).

Let ν_1 and ν_2, be the distributions of $\hat{\mathbb{S}}_{\tau_m}$ and $\hat{\mathbb{S}}_T$ respectively, under the measure $\mathbb{Q}_1$. Let ν_3 be the be the distributions of $\hat{\mathbb{S}}_T$ under the measure $\hat{\mathbb{Q}}$. In view of (5.1), $d(\nu_1, \nu_2) < \epsilon$. From the definition of the measure $\mathbb{Q}_1$ it follows that $d(\nu_2, \nu_3) < \frac{2}{N}$. Moreover, (4.9) implies that $d(\nu_3, \mu^{(N)}) < \frac{K}{N}$ and $\mu^{(N)}$ converges to μ weakly. Hence, the preceding inequalities, together with this convergence yield that for all sufficiently large N, $d(\nu_1, \mu) < 2\epsilon$. Finally, we observe that in view of (5.2), $(Z_T, P^W) = \nu_1$.

We now use Theorem 4 on page 358 in [35] and Theorem 1 in [36] to construct a measurable function $\psi : \mathbb{R}^2 \to \mathbb{R}$ such that the random variable $\Lambda := \psi(Z_T, W_T^{(m+2)})$ satisfies

$$(\Lambda, P^W) = \mu \quad \text{and} \quad P^W(|\Lambda - Z_T| > 2\epsilon|) < 2\epsilon. \tag{5.12}$$

We define a martingale by,

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

In view of (5.12), the distribution of the martingale Γ is an element in $\mathbb{M}_\mu$. Hence,

$$\sup_{\mathbb{Q} \in \mathbb{M}_\mu} \mathbb{E}_{\mathbb{Q}}[G(\mathbb{S})] \geq E^W(G(\Gamma)). \tag{5.13}$$

We continue with the estimate that connects the distribution of Γ to $\mathbb{Q} \in \mathbb{M}_N^{(K)}$. Observe that $E^W \Lambda = E^W Z_T = 1$. This together with (5.12), positivity of $Z + \frac{1}{N}$ and Λ, and the Holder inequality yields

$$\begin{aligned}
E^W |\Lambda - Z_T| &= 2E^W (\Lambda - Z_T)^+ - E^W (\Lambda - Z_T) \\
&= 2E^W (\Lambda - Z_T)^+ \\
&\leq 4\epsilon + \frac{2}{N} + 2E^W (\Lambda \chi_{\{|\Lambda - Z_T| > 2\epsilon\}}) \\
&\leq 4\epsilon + \frac{2}{N} + 2 \left(\int x^p d\mu(x) \right)^{1/p} (2\epsilon)^{1/q},
\end{aligned} \tag{5.14}$$

where $p > 1$ is as (2.4) and $q = p/(p-1)$. From (5.14) and the Doob inequality for the martingale $\Gamma_t - Z_t$, $t \in [0, 1]$ we obtain

$$E^W (\chi_{\{\|\Gamma - Z\| > \epsilon^{1/2q}\}}) \leq \frac{E^W |\Lambda - Z_T|}{\epsilon^{1/2q}} \leq \frac{4\epsilon + \frac{2}{N} + 2 \left(\int x^p d\mu(x) \right)^{1/p} (2\epsilon)^{1/q}}{\epsilon^{1/2q}}. \tag{5.15}$$

We now introduce a stochastic process $(\hat{Z}_t)_{t=0}^T$, on the Brownian probability space, by, $\hat{Z}_t = Z_{\sigma_k}$ for $t \in [\sigma_k, \sigma_{k+1})$, $k < m$ and for $t \in [\sigma_m, T]$, we set $\hat{Z}_t = Z_{\sigma_m}$. On the space $(\hat{\Omega}, \mathbb{Q}_1)$ let $\tilde{\mathbb{S}}_t = \hat{\mathbb{S}}_{t \wedge \tau_m}$, $t \in [0, T]$. Recall that G is bounded by K. We now use the Assumption (2.1) together with (5.1) and (5.11) to arrive at

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}_1}(G(\hat{\mathbb{S}})) - \mathbb{E}_{\mathbb{Q}_1}(G(\tilde{\mathbb{S}})) &\leq K\epsilon \\
|E^W (G(Z)) - E^W (G(\hat{Z}))| &\leq LE^W \|Z - \hat{Z}\| \leq \frac{L}{N}.
\end{aligned} \tag{5.16}$$

Recall that by (5.2), $(\hat{Z}, P^W) = (\tilde{\mathbb{S}}, \mathbb{Q}_1)$. Thus, $E^W(G(\hat{Z})) = \mathbb{E}_{\mathbb{Q}_1}(G(\tilde{\mathbb{S}}))$. This together with Assumption 2.1 and (5.16) yields

$$\mathbb{E}_{\hat{\mathbb{Q}}}(G(\hat{\mathbb{S}})) \leq \frac{L}{N} + \mathbb{E}_{\mathbb{Q}_1}(G(\hat{\mathbb{S}})) \leq \frac{2L}{N} + K\epsilon + E^W(G(Z)). \qquad (5.17)$$

From Assumption 2.1, (5.13)–(5.15) and (5.17) we obtain

$$\sup_{\mathbb{Q}\in\mathbb{M}_\mu} \mathbb{E}_{\mathbb{Q}}[G(\mathbb{S})] \geq E^W(G(\Gamma))$$

$$\geq E^W(G(Z)) - L\epsilon^{1/2q} - K E^W(\chi_{\{\|\Gamma - Z\| > \epsilon^{1/2q}\}})$$

$$\geq \mathbb{E}_{\hat{\mathbb{Q}}}(G(\hat{\mathbb{S}})) - f_K(\epsilon, N),$$

where

$$f_K(\epsilon, N) = \frac{2L}{N} + K\epsilon + L\epsilon^{1/2q} + K\frac{4\epsilon + \frac{2}{N} + \left(\int x^p d\mu(x)\right)^{1/p} (2\epsilon)^{1/q}}{\epsilon^{1/2q}}.$$

$\square$

6 Possible extensions

In this paper, we prove a Kantorovich type duality for a super-replication problem in financial market with no prior probability structure. The dual is a martingale optimal problem.

The main theorem holds for nonlinear path-dependent options satisfying Assumption 2.1. Although this condition is satisfied by most of the examples, it is an interesting question to characterize the class of functions for which the duality holds. A possible procedure for extending the proof is the following. Assumption 2.1 is used in the proofs of Lemmas 4.2, 4.4 and 4.8. In Lemma 4.2, only the linear growth implied by the assumption is used and one may replace this assumption by an appropriate growth condition on the function G. In particular, if G is bounded no assumption would be required.

Since the inequality (2.10) holds for any measurable function G, we need to extend the proof of the inequality (2.9). We may achieve this by modifying the right hand side of formula (4.7) in Definition 4.6 and use a sequence of functions $G_n(\hat{\mathbb{S}})$ satisfying the Assumption 2.1 and $G_n \downarrow G$ as n approaches to ∞. Under this structure, we skip Lemma 4.4, and prove Lemma 4.8 directly. The final step would be a modification of Proposition 5.1 to the following claim

$$\limsup_{N\to\infty} \sup_{\mathbb{Q}\in\mathbb{M}_N^{(K)}} \mathbb{E}_{\mathbb{Q}}[G_N(\hat{\mathbb{S}})] \leq \sup_{\mathbb{Q}\in\mathbb{M}_\mu} \mathbb{E}_{\mathbb{Q}}[G(\mathbb{S})].$$

This extension technique also applies to Barrier options. In this case, we use the approximating sequence as the payoffs G_n of Barrier options with a larger (than the original payoff G) corridor. The main concern here is to discretize the process in a way adapted to the barriers.

Two other important extensions are to the case of many stocks and the inclusion of the possibility of jumps into the stock price process. We believe that for the multi-dimensional case, a discretization based proof would be possible. The main difficulty here is to appropriately define the crossing times and use them to obtain a piece-wise constant approximation of a generic stock price process.

Finally, the discretization technique developed in this paper also applies to markets with frictions. Indeed, recently, the authors proved the duality in a finite time model with proportional transactions costs [19] using an earlier result of [18].

References

1. Ambrosio, L., Pratelli, A.: Existence and stability results in the L^1 theory of optimal transportation. In: Optimal Transportation and Applications (Martina Franca, 2001). Lecture Notes in Mathematics, Vol. 1813, pp. 123–160. Springer, Berlin (2003)
2. Acciaio, B., Beiglböck, M., Penkner, F., Schachermayer, W., Temme, J.: A trajectorial interpretation of Doob's Martingale inequalities. Ann. Appl. Prob. **23/4**, 1494–1505 (2013)
3. Acciaio, B., Beiglböck, M., Schachermayer, W.: A Model-free Version of the Fundamental Theorem of Asset Pricing and the Super-Replication Theorem. Math. Finance (2013, to appear)
4. Bass, R.F.: Skorokhod imbedding via stochastic integrals. In: Seminaire de Probabilites, XVII. Lecture Notes in Mathematics, Vol. 986, pp. 221–224. Springer, Berlin (1983)
5. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
6. Beiglböck, M., Henry-Labordère, P., Penkner, F.: Model-independent bounds for option prices: a mass transport approach. Finance Stoch. **17**, 477–501 (2013)
7. Brown, H., Hobson, D., Rogers, L.C.G.: Robust hedging of barrier options. Math. Finance **11**, 285–314 (2001)
8. Bouchard, B., Nutz, N.: Arbitrage and duality in non-dominated discrete-dime models. arxiv:1305.6008 (2013, preprint)
9. Carr, P., Lee, R.: Hedging variance options on continuous semimartingales. Finance Stoch. **14**, 179–207 (2010)
10. Cox, A.M.G., Hobson, D., Obłój, J.: Pathwise inequalities for local time: applications to Skorokhod embeddings and optimal stopping. Ann. Appl. Probab. **18/5**, 1870–1896 (2008)
11. Cox, A.M.G., Obłój, J.: Robust pricing and hedging of double no-touch options. Finance Stoch. **15**, 573–605 (2011)
12. Cox, A.M.G., Obłój, J.: Robust hedging of double touch barrier options. SIAM J. Financial Math. **2**, 141–182 (2011)
13. Cox, A.M.G., Wang, J.: Root's barrier: construction, optimality and applications to variance options. Ann. Appl. Probab. **23**, 859–894 (2013)
14. 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)
15. Davis, M.H.A., Hobson, D.: The range of traded option prices. Math. Finance **17/1**, 1–14 (2007)
16. Davis, M.H.A., Obłój, J., Raval, V.: Arbitrage bounds for prices of weighted variance swaps. Math. Finance. doi:10.1111/mafi.12021
17. Delbaen, F., Schachermayer, W.: A general version of the fundamental theorem of asset pricing. Math. Ann. **300**, 463–520 (1994)
18. Dolinsky, Y., Soner, H.M.: Duality and Convergence for Binomial Markets with Friction. Finance Stoch. **17/3**, 447–475 (2013)
19. Dolinsky, Y., Soner, H.M.: Robust hedging with proportional transaction costs. arxiv:1302.0590 (2013, preprint).

20. El Karoui, N., Quenez, M.-C.: Dynamic programming and pricing of contingent claims in an incomplete market. SIAM J. Control Opt. **33/1**, 29–66 (1995)
21. 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. (to appear)
22. Hobson, D.: Robust hedging of the lookback option. Finance Stoch. **2**, 329–347 (1998)
23. Hobson, D.: The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices. Princeton Lectures on Mathematical Finance. Springer, Paris (2010)
24. Hobson, D., Klimmek, M.: Model independent hedging strategies for variance swaps. Finance Stoch. **16**, 611–649 (2012)
25. Hobson, D., Laurence, P., Wang, T.H.: Static-arbitrage optimal sup-replicating strategies for basket options. Insur. Math. Econ. **37**, 553–572 (2005)
26. Hobson, D., Laurence, P., Wang, T.H.: Static-arbitrage upper bounds for the prices of basket options. Quant. Finance **5**, 329–342 (2005)
27. Hobson, D., Neuberger, A.: Robust bounds for forward start options. Math. Finance **22**, 31–56 (2012)
28. Hobson, D., Pedersen, J.L.: The minimum maximum of a continuous martingale with given initial and terminal laws. Ann. Probab. **30**, 978–999 (2002)
29. Kantorovich, L.V.: On the transfer of masses. Dokl. Akad. Nauk. SSSR (in Russian) **37**, 227–229 (1942)
30. Kantorovich, L.V.: On a problem of Monge. Uspekhi Mat. Nauk. **3**, 225–226 (1948)
31. Kellerer, H.G.: Duality theorems for marginal problems. Z. Wahrsch. Verw. Gebiete **67**, 399–432 (1984)
32. Obłój, J., Henry-Labordère, P., Spoida, P., Touzi, N.: Maximum of Martingales given marginals. arxiv 1203.6877 (2013)
33. Monge, G.: Memoire sur la Theorie des Déblais et des Remblais. Histoire de L'Acad. des Sciences de Paris, pp 666–704 (1781)
34. Obłój, J.: The Skorokhod embedding problem and its offspring. Probab. Surv. **1**, 321–390 (2004)
35. Shiryaev, A.N.: Probability. Springer, New York (1984)
36. Skorokhod, A.V.: On a representation of random variables. Theory Probab. Appl. **21**, 628–632 (1976)
37. Strasser, H.: Mathematical theory of statistics. In: de Gruyter Studies in Mathematics, Vol. 7. Springer, Berlin (1985)
38. Soner, H.M., Touzi, N., Zhang, J.: Dual formulation of second order target problems. Ann. Appl. Probab. **23/1**, 308–347 (2013)
39. Soner, H.M., Touzi, N., Zhang, J.: Wellposedness of second order backward SDEs. Probab. Theory Relat. Fields **153**, 149–190 (2012)
40. Villani, C.: Topics in optimal transportation. In: Graduate Studies in Mathematics, Vol. 58. AMS, Providence (2003)
41. Villani, C.: Optimal transport. old and new. Grundlehren der mathematischen Wissenschaften, Vol. 338. Springer, Berlin (2009)