MARTINGALE OPTIMAL TRANSPORT AND ROBUST HEDGING IN CONTINUOUS TIME

YAN DOLINSKY AND H.METE SONER

HEBREW UNIVERSITY OF JERUSALEM AND ETH ZURICH

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 has the 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.

1. Introduction

The original transport problem proposed by Monge [18] 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 [15, 16], 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, 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

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

Date: January 22, 2013.

2010 *Mathematics Subject Classification.* 91G10, 60G44.

Key words and phrases. European Options, Robust Hedging, Min–Max Theorems, Prokhorov Metric, Optimal transport.

Research partly supported by the European Research Council under the grant 228053-FiRM, by the ETH Foundation and by the Swiss Finance Institute.

The dual objective function 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, 23, 24] 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 2.7 and also in the papers [4] and [12].

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 [13], 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 as in [13] and in other papers that consider robust hedging of model-independent price bounds. 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 2.6, and an approximation result, Theorem 2.8, for a general class of path-dependent options. The classical duality theorem, for a market with a given 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 & Schachermayer [10] (Theorem 5.7) for the general semi-martingale processes and to El-Karoui & Quenez [11] for its dynamic version in the diffusion case. Theorem 2.6 below, also provides a dual representation of the minimal super-replication cost but for the general 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. Indeed, typically martingale measures are orthogonal to each other and this fact renders the problem difficult. However, the additional feature of the market that the call prices are known, introduces the new marginal constraint. This in turn, provides the connection to the problem of optimal transport.

In the literature, there are two earlier results in this direction. In a purely discrete setup, a similar result was recently proved by Beiglbock, Henry-Labordère and Penkner [4]. 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 [4] the dual is recognized as a Monge-Kantorovich type optimal transport for martingale measures and is essentially used.

In continuous time, Galichon, Henry-Labordère and Touzi [12] 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 [12]. The duality result, Proposition 2.1 in [12], 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 [12] 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 [21]. As in the unconstrained model of [21, 22], 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 with probability one with respect to all measures in a given class. In general, these measures are not dominated with one measure. As already mentioned this is the main difficulty and differs the current problem from the classical duality discussed earlier. However, our duality result together with the results of [12] implies that these two approaches – namely, robust hedging through the path-wise definition of this paper and the quasi-sure definition of [12, 21] – yield the same value. This is proved in Section 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.

The robust hedging has been an active research area over the past decade. The initial paper of Hobson [13] 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 was further developed by Brown, Hobson and Rogers [5] and Cox and Obloj [7] and in several other papers. We refer the reader to the excellent survey of Hobson [14] and the references therein. Also a similar modeling approach was applied to volatility options by Carr and Lee [6]. We refer to the recent paper by Cox and Wang [8] for more information and a discussion of various constructions of the Root's solution of the Skorokhod embedding. Also in a recent paper, Davis and Obloj [9] considered the variance swaps in a market with finitely many put options. In particular, in [9] the class of admissible portfolios is enlarged and numerical evidence is obtained by analyzing the S&P500 index options data.

As already mentioned above, the dual approach is used by Galichon, Henry-Labordère and Touzi [12] and Henry-Labordère and Touzi [17] as well. In these papers, the duality provides a connection to stochastic optimal control which can be then used to compute the solution in a more systematic manner.

Our approach 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 the classical hedging, under the right choice of the probability measure. Thus we can apply the classical duality results for super–hedging of European options on a given probability space. The last step is to analyze the limit of the obtained prices. We combine methods from arbitrage–free pricing and limit theorems for stochastic processes.

The paper is organized as follows. Main results are formulated in the next section. In Section 3, the connection between the quasi sure approach and ours is proved. The last two sections are devoted to the proof of one inequality which implies the main results.

Acknowledgements. The authors would like to thank Mathias Beiglbock, David Belius, Alex Cox, Jan Obloj, Walter Schachermayer and Nizar Touzi for insightful discussions and comments. In particular, Section 3 resulted from discussions with Obloj and Touzi and Shachermayer pointed out a mistake in an earlier version.

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 emphasis 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 approach is through approximation, we need the regularity of the pay-off functional G. A discussion of this assumption is given in Remark 2.2. In particular, it is related to the classical Skorokhod topology, and Asian and lookback type options satisfy the below condition. We need the following definition. Let $\mathcal{D}_N[0,T]$ be the subset of $\mathcal{D}[0,T]$ that are piecewise constant with N possible jumps $t_0 = 0 < t_1 < t_2 < \ldots < t_N \le T$, i.e., $\upsilon \in \mathcal{D}_N[0,T]$ if and only if

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

We make the following standing assumption on G.

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

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

Moreover, let $\upsilon, \tilde\upsilon \in \mathcal{D}_N[0,T]$ be such that $\upsilon_i = \tilde\upsilon_i$ for all $i = 1, ..., N$. 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 award. Let $r > 0$ be the constant interest rate. Then, the payoff

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

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

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

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 [3] for more information. In particular, while $\int_0^T S_t dt$ is continuous with respect to the Skorokhod metric in $\mathcal{C}[0, T]$, it is not Lipschitz continuous in $\mathcal{C}[0, T]$ and it is not even continuous in $\mathcal{D}[0, T]$. Although we assume S to be continuous, since in our analysis we need to consider approximations in $\mathcal{D}[0, T]$, 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 2.6 and 2.8 can be extended to payoffs of the form

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

where H is Lipschitz and $0 < t_1 < ... < t_k \le T$.

Finally, one can extend the results to barrier options, by considering a discretization that is adapted to the barriers. In this paper, we choose not to include this extension to avoid more technicalities. □

2.2. **European Calls.** Let μ be a given probability measure on $\mathbb{R}_+$. At time zero, the investor is allowed to buy any call option with with strike $K \ge 0$, for a price

$$C(K) := \int (x - K)^+ d\mu(x).$$

The probability measure μ is assumed to be derived from observed call prices that are liquidly traded in the market. One may also think that μ describes the probabilistic belief of the firm for the stock price at time T. Then, an arbitrage and an approximation arguments imply that the price of an option 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 $g \in \mathbb{L}^1(\mathbb{R}_+, \mu)$.

Also, since $C(0)$ is the price of a forward, it 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 order to avoid arbitrage, it should satisfy that

$$C(0) = \int x d\mu(x) = S_0 = 1. \tag{2.2}$$

Technically, we also assume that there exists $p > 1$ such that

$$\int x^p d\mu(x) < \infty. \tag{2.3}$$

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 of Hobson 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 (2.3).

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 $v, \tilde{v} \in A$,

$$v_u = \tilde{v}_u, \quad \forall u \in [0,t] \quad \Rightarrow \quad \phi(v)_t = \phi(\tilde{v})_t. \tag{2.4}$$

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 $\chi_{\mathbb{A}}$ is the indicator of the set $\mathbb{A}$. A semi-static portfolio is *admissible,* if there exists $M > 0$ such that

$$Z_t^\pi(S) \geq -M \left(1 + \sup_{0 \leq u \leq t} S_u^p\right), \quad \forall t \in [0,T], \quad S \in \mathcal{C}^+[0,T]. \tag{2.5}$$

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

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

5. The (minimal) *super-hedging cost* of G is defined by,

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

□

Notice that the set of admissible portfolios depend on the exponent p which appears in the assumption (2.3). 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 $\mathbb{S} = (\mathbb{S}_t)_{0\le t\le T}$ be the canonical process given by $\mathbb{S}_t(\omega) := \omega_t$, for all $\omega \in \Omega$. Let $\mathcal{F}_t := \sigma(\mathbb{S}_s,\, 0 \le s \le t)$ be the canonical filtration.

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, below the probability measures need to satisfy this condition as well.

Definition 2.4. A probability measure $\mathbb{P}$ 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 $\mathbb{P}$ and $\mathbb{S}_0 = 1$ $\mathbb{P}$-a.s.

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

Note that if μ satisfies (2.2), then the canonical process $(\mathbb{S}_t)_{t=0}^T$ is a *martingale* (not only a local martingale) under any measure $\mathbb{P} \in \mathbb{M}_\mu$.

Remark 2.5. Observe that (2.2) 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 $\sigma\{W_s | s \le t\}$. 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 (2.2), $M_0 = 1$. Since M is a Brownian martingale, it is continuous. Moreover, since μ has support on the positive real line, $f \ge 0$ and consequently, $M \ge 0$. Then, the distribution of M on the space Ω is an element in $\mathbb{M}_\mu$. □

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

Theorem 2.6. *Assume that the European claim G satisfies the Assumption 2.1 and the probability measure μ satisfies* (2.2) *and* (2.3)*. Then, the minimal super-hedging cost is given by*

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

where $\mathbb{E}_\mathbb{P}$ denotes the expectation with respect to $\mathbb{P}$.

Remark 2.7. One may consider the maximizer, if exists, of the expression

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

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 imposed. This in turn places a restriction on the measures, namely (2.2). 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 [4] for a very clear discussion of the connection between the robust hedging and the 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,

$$\tau_k^{(N)}(S) = \inf\left\{ t > \tau_{k-1}^{(N)}(S) : |S_t - S_{\tau_{k-1}^{(N)}(S)}| = \frac{1}{N} \right\} \wedge T, \tag{2.6}$$

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

$$H^{(N)}(S) = \min\{k \in \mathbb{N} : \tau_k^{(N)}(S) = T\}. \tag{2.7}$$

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 $0 = \tau_0^{(N)}(S) < \tau_1^{(N)}(S) < ... < \tau_{H^{(N)}(S)}^{(N)}(S) = T$. Formally, $\pi := (g, \gamma) \in \mathcal{A}_N$, if it is progressively measurable in the sense of (2.4) 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 $k \geq 1$, $V_N(G) \geq V_{kN}(G) \geq V(G)$. 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 2.8. *Under the assumptions of Theorem 2.6,*

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

2.6. **Proofs of Theorems 2.6 and 2.8.** In order to establish Theorem 2.6 and Theorem 2.8 it is sufficient to prove that

$$\limsup_{N \to \infty} V_N(G) \leq \sup_{\mathbb{P} \in \mathbb{M}_\mu} \mathbb{E}_{\mathbb{P}} \left[G(\mathbb{S}) \right] \tag{2.8}$$

and

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

The first inequality is the difficult one and it will be proved in the last two sections. The second inequality is simpler and we provide its proof here.

Let $\mathbb{P} \in \mathbb{M}_\mu$ and let $\pi = (g, \gamma)$ be super-replicating. Since γ is progressively measurable in the sense of (2.4), the stochastic integral

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

is defined with respect to $\mathbb{P}$. Also $\mathbb{P}$ is a martingale measure and G satisfies the lower bound (2.5). Hence, the above stochastic integral is a $\mathbb{P}$ local–martingale. Moreover,

$$\int_0^t \gamma_u(\mathbb{S})d\mathbb{S}_u \geq G(\mathbb{S}) \geq -M(1+\sup_{0\leq t\leq T}|\mathbb{S}_t|^p).$$

Also in view of (2.3) and the Doob-Kolmogorov inequality,

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

Therefore, $\mathbb{E}_{\mathbb{P}} \int_0^T \gamma_u(\mathbb{S})d\mathbb{S}_u \leq 0$. Since π is super-replicating, this together with (2.5) implies that

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

where in the last equality we again used that the distribution of $\mathbb{S}_T$ under $\mathbb{P}$ equals to μ. This completes the proof of the lower bound. Together with (2.8), which will be proved in the last sections, 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 it was done in [12, 21]. Let us briefly recall this notion. Let $\mathcal{Q}$ be the set of all martingale measures $\mathbb{P}$ on the canonical space $\mathcal{C}^+[0,T]$ under which the canonical process $\mathbb{S}$ satisfies $\mathbb{S}_0 = 1, \mathbb{P}$-a.s., has quadratic variation and satisfies $\mathbb{E}_{\mathbb{P}} \sup_{0\leq t\leq T} \mathbb{S}_t < \infty$. 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

$$\int_0^t \gamma_u d\mathbb{S}_u, \quad t \in [0,T]$$

exists for any probability measure in the set $\mathbb{P} \in \mathcal{Q}$ and satisfies (2.5) $\mathbb{P}$-a.s. We refer the reader to [21] 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

$$g(\mathbb{S}_T) + \int_0^T \gamma_u d\mathbb{S}_u \geq G(\mathbb{S}), \quad \mathbb{P} \text{ a.s.},$$

for all $\mathbb{P} \in \mathcal{Q}$. 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,

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

From simple arbitrage arguments it follows that

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

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

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

Now if $\mathbb{P} \in \mathbb{M}_\mu$, then the two terms with λ are equal. So we first restrict the measures to the set $\mathbb{M}_\mu$ and then use Theorem 2.6, to arrive at

$$\begin{aligned} V(G) &\geq V_{qs}(G) \geq \inf_{\lambda\in\mathbb{L}^1(\mathbb{R}_+,\mu)} \sup_{\mathbb{P}\in\mathcal{Q}} \mathbb{E}_{\mathbb{P}}\left(G(\mathbb{S}) - \lambda(\mathbb{S}_T) + \int \lambda d\mu\right) \\ &\geq \sup_{\mathbb{P}\in\mathcal{Q}} \inf_{\lambda\in\mathbb{L}^1(\mathbb{R}_+,\mu)} \mathbb{E}_{\mathbb{P}}\left(G(\mathbb{S}) - \lambda(\mathbb{S}_T) + \int \lambda d\mu\right) \\ &\geq \sup_{\mathbb{P}\in\mathbb{M}_\mu} \mathbb{E}_{\mathbb{P}}\left[G(\mathbb{S})\right] = V(G). \end{aligned}$$

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* (2.2),(2.3). *Then,*

$$\begin{aligned} V(G) &= V_{qs}(G) = \sup_{\mathbb{P}\in\mathbb{M}_\mu} \mathbb{E}_{\mathbb{P}}\left[G(\mathbb{S})\right] \\ &= \inf_{\lambda\in\mathbb{L}^1(\mathbb{R}_+,\mu)} \sup_{\mathbb{P}\in\mathcal{Q}} \mathbb{E}_{\mathbb{P}}\left(G(\mathbb{S}) - \lambda(\mathbb{S}_T) + \int \lambda d\mu\right) \\ &= \sup_{\mathbb{P}\in\mathcal{Q}} \inf_{\lambda\in\mathbb{L}^1(\mathbb{R}_+,\mu)} \mathbb{E}_{\mathbb{P}}\left(G(\mathbb{S}) - \lambda(\mathbb{S}_T) + \int \lambda d\mu\right). \end{aligned}$$

4. Proof of the main results

The rest of the paper is devoted to the proof of (2.8).

4.1. Reduction to bounded claims.

We start with a reduction to claims that are bounded from above. Towards this result, we first prove a technical result. 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 subsection 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 (2.6), (2.7), respectively, and set

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

We next define a portfolio $(g^{(N,K)}, \gamma^{(N,K)}) \in \mathcal{A}_N$ as follows. For $t \in (\tau_k, \tau_{k+1}]$ and $k = 0, 1, ..., n-1$, set

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

and with $c_p := p/(p-1)$ define $g^{(N,K)}$ by

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

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

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

where

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

Therefore, $\pi^{(N,K)} := (g^{N,K)}, \gamma^{(N,K)})$ satisfies (2.5) and super-replicates α_K. Hence,

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

Also, in view of (2.3),

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

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* (2.8) *for claims G that are non-negative, bounded from above and satisfying the Assumption 2.1.*

Proof. We proceed in two steps. First suppose that (2.8) holds for nonnegative claims that are bounded from above. Then, the conclusions of Theorem 2.6 and Theorem 2.8 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 (2.8) holds for G_K. Therefore,

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

In view of Assumption 2.1,

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

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

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

By the linearity of the market, this inequality implies that

$$V_N(G) \le V_N(G_K) + V_N\left((L\|S\| + G(0) - K)\,\chi_{\{L\|S\|+G(0)\ge K\}}\right).$$

Moreover, in view of the previous lemma,

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

Using these, we conclude that

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

Hence, (2.8) 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 (2.8) holds, i.e.,

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

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

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

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

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

In view of Lemma 4.1,

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

We combine the above inequalities to conclude that

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

This exactly (2.8). □

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 (2.6) and (2.7), 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

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

where as usual $sign(x) = 1$ if $x \geq 0$, and $sign(x) = -1$ for $x < 0$.

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

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

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 $\mathbb{D}[0,T]$ 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

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

be the sets of possible differences between two consecutive jump times. Next, we define a subset $\mathbb{D}^{(N)}$ 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. $f(0) = 1$,
2. f is piecewise constant with jumps at times $t_1, ..., t_n$, where
$$t_0 = 0 < t_1 < t_2 < ... < t_n < T,$$
3. for any $k = 1, ..., n$, $|f(t_k) - f(t_{k-1})| = 1/N$,
4. for any $k = 1, ..., n$, $t_k - t_{k-1} \in U_k^{(N)}$. □

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.

We now define an approximation

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

as follow. Recall $\tau_k = \tau_k^{(N)}(S)$, $n = H^{(N)}(S)$ from above and also from (2.6), (2.7). Set $\hat{\tau}_0 := 0$, and for $k = 1, ..., n$, define

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

It is clear that $0 = \hat{\tau}_0 < \hat{\tau}_1 < ... < \hat{\tau}_n < T$ and $\hat{\tau}_k < \tau_k$ for all $k = 0, \dots, n$.

We now define $F^{(N)}(S)$ by,

$$F_t^{(N)}(S) = [1 - S_{\tau_1}] + \sum_{k=0}^{n-2} S_{\tau_{k+1}} \chi_{[\hat{\tau}_k, \hat{\tau}_{k+1})}(t) + \left[S_{\tau_{n-1}} + \frac{1}{N} sign(S_T - S_{\tau_{n-1}}) \right] \chi_{[\hat{\tau}_{n-1}, T]}(t), \tag{4.2}$$

where we set $\hat{\tau}_{-1} = 0$. Observe that the value of the k-th 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,

$$F_{\hat{\tau}_m}^{(N)} - F_{\hat{\tau}_{m-1}}^{(N)} = S_{\tau_{m+1}} - S_{\tau_m}, \qquad \forall\ m = 1, \dots, n-2, \quad \text{and} \tag{4.3}$$
$$F_{\hat{\tau}_{n-1}}^{(N)} - F_{\hat{\tau}_{n-2}}^{(N)} = \frac{1}{N} sign\left(S_T - S_{\tau_{n-1}}\right).$$

This shift is essential in order to deal with some delicate questions of adeptness 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 (2.4). 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. We also point out 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* (4.2). *For any G satisfying the Assumption 2.1 with the constant L,*

$$\left|G(S) - G(F^{(N)}(S))\right| \le \frac{5L\|S\|}{N}, \quad \forall\ S \in \mathcal{C}^+[0,T].$$

Proof. Define the map

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

First recall Assumption 2.1 and observe that $\hat{S}$ and $\hat{F} = \hat{F}^{(N)}$ are like the functions υ and $\tilde{\upsilon}$ in that definition. Hence,

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

For $k < n$,

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

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,

$$\sum_{k=1}^{n} |\Delta\tau_k - \Delta\hat{\tau}_k| \le \sum_{k=1}^{\infty} \frac{1}{2^k N} = \frac{1}{N}. \tag{4.4}$$

Combining the above inequalities, we arrive at

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

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

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

Finally, we observe that by construction,

$$\left\|\hat{F} - F\right\| \le \frac{2}{N}, \quad \Rightarrow \quad \left|G(F) - G(\hat{F})\right| \le \frac{2L}{N}.$$

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, (4.4) 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 $\{1/(i2^kN) : i = 1, 2, \dots\}$ in $U_k^{(N)}$ is to make sure that $\Delta\hat{\tau}_k > 0$. For probabilistic reasons (i.e. adaptability), we want $\hat{\tau}_k \le \tau_k$. 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 (2.8) is a duality result for probabilistic problems. We first introduce this structure and then relate it to the problem V_N.

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). Denote by $\hat{\mathbb{S}} = (\hat{\mathbb{S}}_t)_{0\le t\le T}$ 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 $\hat{\mathbb{P}}^{(N)}\left(\mathbb{D}^{(N)}\right) = 1$ and $\hat{\mathbb{P}}^{(N)}(\{f\}) > 0$ for all $f \in \mathbb{D}^{(N)}$. Let $\hat{\mathcal{F}}_t^{(N)}$, $t \in [0, T]$ by the filtration which is generated by the process $\hat{\mathbb{S}}$ and satisfying the usual assumptions (right continuous and contains $\hat{\mathbb{P}}^{(N)}$ null sets). Under the measure $\hat{\mathbb{P}}$, the canonical map $\hat{\mathbb{S}}$ has finitely many jumps. Let

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

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

Then, 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 *càdlàg* progressively measurable stochastic process $(\hat{\gamma}_t)_{t=0}^T$. Thus , there exists a map

$$\phi : \mathbb{D}[0, T] \to \mathbb{D}[0, T],$$

so that it is progressively measurable in the sense of (2.4) and $\hat{\gamma} = \phi(\hat{\mathbb{S}})$, $\hat{\mathbb{P}}$-a.s. Since we may always work with ϕ instead of $\hat{\gamma}$, there is no loss of generality in assuming that $\hat{\gamma}$ is itself is progressively measurable in the sense of (2.4).

We now give 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 $\hat{\gamma} : \mathbb{D}[0, T] \to \mathbb{D}[0, T]$ is *càdlàg*, progressively measurable and $h : A^{(N)} \to \mathbb{R}$.
2. A semi-static portfolio is $\hat{\mathbb{P}}^{(N)}$*-admissible,* if h is bounded and there exists $M > 0$ such that

$$\int_0^t \hat{\gamma}_{u\text{-}} d\hat{\mathbb{S}}_u \ge -M, \quad \hat{\mathbb{P}}^{(N)} - a.s., \quad t \in [0, T], \tag{4.5}$$

where $\hat{\gamma}_{u\text{-}}$ is the left limit of the *càdlàg* function $\hat{\gamma}$ at time u.

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

$$h(\hat{\mathbb{S}}_T) + \int_0^T \hat{\gamma}_{u\text{-}} d\hat{\mathbb{S}}_u \geq G(\hat{\mathbb{S}}), \quad \hat{\mathbb{P}}^{(N)} - a.s. \tag{4.6}$$

□

4.4. **Approximating μ.** Recall the set $\mathcal{A}_N$ of portfolios used in the definition of V_N in subsection 2.5.

Next we provide a connection between the probabilistic super-replication and the discrete robust problem. However, h in 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,

$$\int h d\mu^{(N)} = \int g^{(N)} d\mu. \tag{4.7}$$

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

4.5. **Probabilistic super-replication.** We now introduce the super-replication problem by requiring that the inequalities in (4.6) hold $\hat{\mathbb{P}}^{(N)}$-almost surely. Let G be a European claim as before and N be a positive integer. Then, the probabilistic super-replication problem is given by,

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

Recall that in the probabilistic structure, admissibility and related notions are defined in Definition 4.6.

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 $\hat{\pi} = (g, \hat{\gamma})$ in sense of Definition 4.6. We lift this portfolio to a semi-static portfolio $\pi^{(N)} = (g^{(N)}, \gamma^{(N)}) \in \mathcal{A}_N$ as defined in the subsection 2.5.

Indeed, let $g^{(N)}$ be as in subsection 4.4 and let $\hat{\gamma}_{t\text{-}}$ be the left limit of the *càdlàg* function $\hat{\gamma}$ at time t. We now define $\gamma^{(N)} : \mathcal{C}^+[0,T] \to \mathcal{D}[0,T]$ by

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

where $\tau_k = \tau_k^{(N)}(S)$ is as in (2.6), n is as in (4.1) and $F^{(N)}(S)$, $\hat{\tau}_k := \hat{\tau}_k(S)$ is as in (4.2). Notice that n, the number of jumps of S, is by construction exactly one more than the number of jumps of $F^{(N)}$. Also notice that

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

Lemma 4.7. *For any probabilistic semi-static portfolio $(g,\hat{\gamma})$, $\gamma^{(N)}$ defined above is progressively measurable in the sense of* (2.4).

Proof. Let $S, \tilde{S} \in \mathcal{C}^+[0,T]$ be such that $S_u = \tilde{S}_u$ for all $u \le t$ for some $t \in [0,T]$. 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

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

so that $0 \le k_t(S) \le n$ 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 $\gamma_t^{(N)}(S) = \gamma_t^{(N)}(\tilde{S}) = 0$. We now assume that $k_t(S) > 0$ and use the definition of $\hat{\tau}_k$ to conclude that

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

Hence,

$$\gamma_t^{(N)}(S) = \hat{\gamma}_{\theta\text{-}} \left(F^{(N)}(S)\right), \quad \gamma_t^{(N)}(\tilde{S}) = \hat{\gamma}_{\theta\text{-}} \left(F^{(N)}(\tilde{S})\right).$$

Moreover, the definition of $F^{(N)}$ implies that

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

Therefore, by the progressive measurability 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 subsection 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

$$G^{(N)}(S) := G(S) - \frac{6L\|S\|}{N}.$$

We first show that

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

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 $\gamma^{(N)}:\mathcal{C}^+[0,T]\to\mathcal{D}[0,T]$ such that the semi-static portfolio $\pi^{(N)}:=(g^{(N)},\gamma^{(N)})$ is admisible, belongs to $\mathcal{A}_N$ and super-replicates $G^{(N)}$ in the sense of Definition 2.3.

Let $g^{(N)}$ be as in subsection 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 (2.5) for every $t\in[0,T]$. Fix such a time point $t\in[0,T]$. In view of (4.5), there exist M so that

$$\int_{[0,t]}\hat{\gamma}_{u\text{-}}(F)dF_u\geq -M,\quad \forall\ t\in[0,T],\ \ \forall\ F\in\mathbb{D}^{(N)}.$$

We claim that

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

for every $k\leq n-1$. Indeed, we use (4.3) and the definitions to compute that

$$\begin{aligned}\int_{[0,\hat{\tau}_{k-1}]}\hat{\gamma}_{u\text{-}}(F)dF_u &= \sum_{m=1}^{k-1}\hat{\gamma}_{\hat{\tau}_m\text{-}}(F)\left(F_{\hat{\tau}_m}-F_{\hat{\tau}_{m-1}}\right)=\sum_{m=1}^{k-1}\hat{\gamma}_{\hat{\tau}_m\text{-}}(F)\left(S_{\tau_{m+1}}-S_{\tau_m}\right)\\ &= \sum_{m=1}^{k}\gamma^{(N)}_{\tau_{m+1}}(F)\left(S_{\tau_{m+1}}-S_{\tau_m}\right)=\int_{\tau_1}^{\tau_k}\gamma_u^{(N)}(S)dS_u\\ &= \int_0^{\tau_k}\gamma_u^{(N)}(S)dS_u.\end{aligned}$$

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 $S\in\mathcal{C}^+[0,T]$, 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,

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

and

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

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

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

and since

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

there exists $\lambda\in[0,1]$ (depending on t) 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 progressively measurable $\hat{\gamma}_u(F) = \hat{\gamma}_u(\tilde{F})$ for all $u \leq \hat{\tau}_k$. Therefore,

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

Also both $F, \tilde{F} \in \mathbb{D}^{(N)}$, and $\hat{\mathbb{P}}(F), \hat{\mathbb{P}}(\tilde{F}) > 0$. Using (4.5) we may conclude that

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

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

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,

$$\begin{aligned} g^{(N)}(S_T) \quad + \quad & \int_0^T \gamma_u^{(N)}(S) dS_u \\ = \quad & \hat{\lambda}\left[h(F_T) + \int_0^T \hat{\gamma}_{u\text{-}}(F) dF_u\right] + (1-\hat{\lambda})\left[h(\bar{F}_T) + \int_0^T \hat{\gamma}_{u\text{-}}(\bar{F}) d\bar{F}_u\right] \\ \geq \quad & \hat{\lambda} G(F) + (1-\hat{\lambda}) G(\bar{F}). \end{aligned}$$

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

$$\left|G(S) - G(\bar{F})\right| \geq |G(S) - G(F)| + \left|G(F) - G(\bar{F})\right| \leq \frac{6L\|S\|}{N}.$$

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

$$V_N(G - 6L\|S\|/N) \leq \hat{V}_N(G).$$

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

$$V_N(G) \leq V_N(6L\|S\|/N) + V_N(G - 6L\|S\|/N).$$

Therefore,

$$V_N(G) \leq V_N(6L\|S\|/N) + \hat{V}_N(G).$$

Finally, by Lemma 4.1,

$$\limsup_{N\to\infty} V_N(6L\|S\|/N) = 0.$$

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 were introduced in subsection 4.3. We consider two classes of measures on this set.

Definition 4.9. 1. We say that a probability measures $\mathbb{Q}$ 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

$$\sum_{k=0}^{\infty} \left| \mathbb{Q}\left(\hat{\mathbb{S}}_T = k/N\right) - \mu^{(N)}\left(\{k/N\}\right)\right| < \frac{K}{N}. \tag{4.8}$$

□

The following follows from known duality results. We will combine it with 4.8 and Proposition 5.1, that will be proved in the next section to complete the proof of the inequality (2.8).

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

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

Proof. Fix N and define the set

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

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 $h \in \mathcal{Z}$ such that

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

By construction, the support of the measure $\hat{\mathbb{P}}$ is $\hat{\mathbb{D}}^{(N)}$. Also all elements of $\hat{\mathbb{D}}^{(N)}$ are piece-wise constant. Thus, $\hat{\mathbb{P}}$ almost surely, the canonical process $\hat{\mathbb{S}}$ is trivially a semi-martingale. Hence, we may use the results of the seminal paper [10]. In particular, by Theorem 5.7 in [10], for $x > \sup_{\mathbb{Q}\in\mathbb{M}_N} \mathbb{E}_{\mathbb{Q}}\left(G(\hat{\mathbb{S}}) - h(\mathbb{S}_T)\right)$, there exists an *admissible* portfolio strategy $\hat{\gamma}$ such that

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

Thus, $(h + x, \hat{\gamma})$ satisfies (4.5)–(4.6), and $\hat{V}_N(G) \leq \mathbb{V} + \epsilon$. We now let ϵ to zero to conclude that

$$\hat{V}_N(G) \leq \inf_{h\in Z} \sup_{\mathbb{Q}\in\mathbb{M}_N} \left(\mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{S}}) - h(\hat{\mathbb{S}}_T)) + \int h d\mu^{(N)}\right). \tag{4.9}$$

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 $f : A^{(N)} \to \mathbb{R}$ equiped 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

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

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 2, in [4] 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 (4.9) yields,

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

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

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

In view of (4.10),

$$\begin{aligned}
\hat{V}_N(G) &\leq \sup_{\mathbb{Q}\in\mathbb{M}_N} \left(\mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{S}})) + \int h_n^{\mathbb{Q}} d\mu^{(N)} - \mathbb{E}_{\mathbb{Q}} h^{\mathbb{Q}}(\hat{\mathbb{S}}_T)\right) \\
&= \sup_{\mathbb{Q}\in\mathbb{M}_N} \left\{\mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{B}})) - N\sum_{k=0}^{\infty} \left|\mathbb{Q}\left(\hat{\mathbb{S}}_T = k/N\right) - \mu^{(N)}(\{k/N\})\right|\right\}
\end{aligned}$$

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

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

Since G is bounded by K, this implies that

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

However, since $G \geq 0$, we may assume that $\mathbb{Q} \in \mathbb{M}_N^{(K)}$. Then,

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

□

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 (2.8) 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* (2.2)-(2.3). *Then,*

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

We prove this result 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. Indeed, 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 continuous function f_K with $f_K(0) = 0$, so that for any $\mathbb{Q} \in \mathbb{M}_N^{(K)}$ and $\epsilon > 0$,*

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

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

Jump times. Since the probability measure $\mathbb{Q}$ is supported on the set $\mathbb{D}^{(N)}$, the the canonical process $\hat{\mathbb{S}}$ is a purely jump process under $\mathbb{Q}$, with a finite number of jumps. Introduce the jump times by setting $\sigma_0 = 0$ and for $k > 0$,

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

Next we introduce the largest random time

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

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

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

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 $i = 1, ..., m$,

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

Wiener space. Let $(\Omega^W, \mathcal{F}^W, P^W)$ be a complete probability space together with a standard $m+2$–dimensional Brownian motion $\left\{W_t = \left(W_t^{(1)}, W_t^{(2))}, ..., W_t^{(m+2)}\right)\right\}_{t=0}^{\infty}$, 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) $\tau_1 \le \tau_2 \le ... \le \tau_m$ such that the distribution (under the Wiener measure P^W) of the random vector $(\tau_1, ..., \tau_m, Z_{\tau_1}, ..., Z_{\tau_m})$ is equal to the distribution of the random vector $(\sigma_1, ..., \sigma_m, \hat{\mathbb{S}}_{\sigma_1}, ..., \hat{\mathbb{S}}_{\sigma_m})$ under the measure $\mathbb{Q}$. Namely,

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

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

Construction of τ'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 $k = 1, ..., m$, define the functions $\Psi_k, \Phi_k : \mathbb{T}^k \times \{-1, 1\}^{k-1} \to [0, 1]$ by

$$\Psi_k(\vec{\alpha}_k; \vec{\beta}_{k-1}) := \mathbb{Q}\left(\sigma_k - \sigma_{k-1} \ge \alpha_k \;\middle|\; A\right), \tag{5.3}$$

where

$$A := \left\{\sigma_i - \sigma_{i-1} = \alpha_i,\ \hat{\mathbb{S}}_{\sigma_i} - \hat{\mathbb{S}}_{\sigma_{i-1}} = \beta_i/N,\ \ i \le k-1\right\},$$

and

$$\Phi_k(\vec{\alpha}_k; \vec{\beta}_{k-1}) = \mathbb{Q}\left(\hat{\mathbb{S}}_{\sigma_k} - \hat{\mathbb{S}}_{\sigma_{k-1}} = 1/N \;\middle|\; B\right), \tag{5.4}$$

where

$$B = \left\{\sigma_k < T, \sigma_j - \sigma_{j-1} = \alpha_j,\ \hat{\mathbb{S}}_{\sigma_i} - \hat{\mathbb{S}}_{\sigma_{i-1}} = \beta_i/N,\ j \le k, i \le k-1\right\}.$$

As usual we set $\mathbb{Q}(\cdot|\emptyset) \equiv 0$. Next, for $k \le m$, 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,

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

and

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

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 $\tau_0 \equiv 0$ and define the random variables $\tau_1, ..., \tau_m, Y_1, ..., Y_m$ by the following recursive relations

$$\begin{aligned} \tau_1 &= \sum_{k=1}^{\infty} t_k \chi_{\{W^{(1)}_{t_k} - W^{(1)}_{t_{k+1}} > \Theta_1(t_k)\}} \prod_{j=k+1}^{\infty} \chi_{\{W^{(1)}_{t_j} - W^{(1)}_{t_{j+1}} < \Theta_1(t_j)\}}, \\ Y_1 &= 2\chi_{\{W^{(2)}_{\{\tau_1\}} > \Gamma_1(\tau_1)\}\}} - 1, \end{aligned} \tag{5.7}$$

and for $i > 1$

$$\begin{aligned} \tau_i &= \tau_{i-1} + \Delta_i \\ Y_i &= \chi_{\{\tau_i<T\}}\left(2\chi_{\{W^{(i+1)}_{\tau_i} - W^{(i+1)}_{\tau_{i-1}} > \Gamma_i(\vec{\Delta\tau}_i, \vec{Y}_{i-1})\}} - 1\right), \end{aligned}$$

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,

$$\begin{aligned} A_i &:= \{|Y_{i-1}| > 0\}, \\ B_{i,k} &:= \{W^{(1)}_{t_k+\tau_{i-1}} - W^{(1)}_{t_{k+1}+\tau_{i-1}} > \Theta_i(\vec{\tau}_{i-1}, t_k; \vec{Y}_{i-1})\}, \\ C_{i,k} &:= \bigcap_{j=k+1}^{\infty} \{W^{(1)}_{t_j+\tau_{i-1}} - W^{(1)}_{t_{j+1}+\tau_{i-1}} < \Theta_i(\vec{\Delta\tau}_{i-1}, t_j; \vec{Y}_{i-1})\}. \end{aligned}$$

Since t_k is decreasing with $t_1 = T$, $\tau_1 \le \tau_2 \le ... \le \tau_m$ 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)–(5.8), the strong Markov property and the independency of the Brownian motion increments it follows that

$$\begin{aligned} (5.8)\qquad & P^W(\tau_k - \tau_{k-1} \ge \alpha_k \big| (\vec{\Delta\tau}_{k-1}; \vec{Y}_{k-1}) = (\vec{\alpha}_{k-1}; \vec{\beta}_{k-1})) \\ &= P^W\left(\bigcap_{j=m}^{\infty}\left(W^{(1)}_{t_j+\tau_{k-1}} - W^{(1)}_{t_{j+1}+\tau_{k-1}} < \Theta_k(\vec{\alpha}_{k-1}, t_j; \vec{\beta}_{k-1})\right)\right) \\ &= \prod_{j=m}^{\infty} P^W\left(W^{(1)}_{t_j+\tau_{k-1}} - W^{(1)}_{t_{j+1}+\tau_{k-1}} < \Theta_k(\vec{\alpha}_{k-1}, t_j; \vec{Y}_{k-1})\right) \\ &= \Psi_k(\vec{\alpha}_k, \vec{\beta}_{k-1}), \end{aligned}$$

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

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

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

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

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

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

where $\Delta\hat{\mathbb{S}}_k = \hat{\mathbb{S}}_{\sigma_k} - \hat{\mathbb{S}}_{\sigma_{k-1}}$, $k \le m$.

Continuous martingale. Set

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

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

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

By the construction of Y and τ'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 $\mathbb{R}$ the Prokhorov's metric is defined by

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

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 the Prokhorov's metric see [19], Chapter 3, Section 7).

Let ν_1 and ν_2, be the distributions of $\hat{\mathbb{S}}_{\sigma_m}$ and $\hat{\mathbb{S}}_T$ respectively, under the measure $\mathbb{Q}$. In view of (5.1), $d(\nu_1,\nu_2) < \epsilon$. Moreover, (4.8) implies that $d(\nu_2, \mu^{(N)}) < \frac{K}{N}$ 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 [19] and Theorem 1 in [20] 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,

$$\sup_{\mathbb{P}\in\mathbb{M}_\mu} \mathbb{E}_{\mathbb{P}}\left[G(\mathbb{S})\right] \ge E^W(G(\Gamma)). \tag{5.13}$$

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 and Λ, and the Holder inequality yields

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

where $p > 1$ is as (2.3) and $q = p/(p-1)$.

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

$$\mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{S}})) - \mathbb{E}_{\mathbb{Q}}(G(\tilde{\mathbb{S}})) \leq K\epsilon \tag{5.15}$$
$$|E^W(G(Z)) - E^W(G(\hat{Z}))| \leq LE^W\|Z - \hat{Z}\| \leq \frac{L}{N}.$$

Recall that by (5.2), $(\hat{Z}, P^W) = (\tilde{\mathbb{S}}, \mathbb{Q})$. Thus, $E^W(G(\hat{Z})) = \mathbb{E}_{\mathbb{Q}}(G(\tilde{\mathbb{S}}))$. This together with (5.15) yields

$$\mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{S}})) \leq \frac{L}{N} + K\epsilon + E^W(G(Z)). \tag{5.16}$$

From Assumption 2.1, the Doob inequality, (5.13)–(5.14) and (5.16) we obtain

$$\begin{aligned}\sup_{\mathbb{P}\in\mathbb{M}_\mu} \mathbb{E}_{\mathbb{P}}[G(\mathbb{S})] &\geq E^W(G(\Gamma)) \\ &\geq E^W(G(Z)) - L\epsilon^{1/2q} - KE^W(\chi_{\{\|\Gamma - Z\| > \epsilon^{1/2q}\}}) \\ &\geq \mathbb{E}_{\mathbb{Q}}(G(\hat{\mathbb{S}})) - \frac{L}{N} - f_K(\epsilon),\end{aligned}$$

where

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

□

References

[1] L. Ambrosio and A. Pratelli, *Existence and stability results in the L^1 theory of optimal transportation* In *Optimal transportation and applications* (Martina Franca, 2001), volume 1813 of Lecture Notes in Math., pages 123160. Springer, Berlin, (2003)

[2] B. Acciaio, M. Beiglbock, F. Penkner, W. Schachermayer and J. Temme, *A Trajectorial Interpretation of Doob's Martingale Inequalities,* to appear in Ann. Appl. Prob.

[3] P. Billingsley, *Convergence of Probability Measures,* Wiley, New York, (1968).

[4] M. Beiglbock, P. Henry-Labordère and F. Penkner, *Model–independent bounds for option prices: a mass transport approach*, preprint.

[5] H. Brown, D. Hobson and L.C.G. Rogers, *Robust hedging of barrier options,* Math.Finance, **11**, 285–314, (2001).

[6] P. Carr and R. Lee, *Hedging Variance Options on Continuous Semimartingales,* Finance and Stochastics, **14**, 179–207, (2010).

[7] A.M.G. Cox and J. Obloj, *Robust pricing and hedging of double no-touch options*, Finance and Stochastics, **15**, 573-605, (2011).

[8] A.M.G. Cox and Wang, *Root's Barrier: Construction, Optimality and Applications to Variance Options*, Annals of Applied Probability, forthcoming, (2012).

[9] M.H.A. Davis, J. Obloj and V. Raval *Arbitrage bounds for prices of weighted variance swaps*, Mathematical Finance, forthcoming, (2012).

[10] F. Delbaen and W. Schachermayer, *A general version of the fundamental theorem of asset pricing*, Math. Annalen, **300**, 463 – 520, (1994).

[11] N. El Karoui and M-C. Quenez, *Dynamic programming and pricing of contingent claims in an incomplete market*, SIAM J. Control and Opt., **33/1**, 29 – 66, (1995).

[12] A. Galichon, P. Henry-Labordère and N. Touzi, *A stochastic control approach to no-arbitrage bounds given marginals, with an application to Lookback options,* preprint.

[13] D. Hobson, *Robust hedging of the lookback option,* Finance and Stochastics, **2**, 329–347, (1998).

[14] D. Hobson, *The Skorokhod Embedding Problem and Model-Independent Bounds for Option Prices,* Paris–Princeton Lectures on Mathematical Finance, Springer, (2010).

[15] L.V. Kantorovich, *On the transfer of masses*, Dokl. Akad. Nauk. SSSR (in Russian), **37**, 227–229, (1942).

[16] L.V. Kantorovich, *On a problem of Monge*, Uspekhi Mat. Nauk. **3**, 225–226, (1948).

[17] P. Henry-Labordère and N. Touzi, *Maximum Maximum of Martingales given Marginals*, preprint.

[18] G. Monge, *Memoire sur la Theorie des Déblais et des Remblais*, Histoire de lAcad. des Sciences de Paris, (1781).

[19] A.N. Shiryaev, *Probability,* Springer-Verlag, New York, (1984).

[20] A.V. Skorokhod, *On a representation of random variables,* Theory Probab. Appl, **21**, 628–632, (1976).

[21] H.M. Soner, N. Touzi, and J. Zhang. *Dual Formulation of Second Order Target Problems*, Annals of Applied Probability, to appear, (2012).

[22] H.M. Soner, N. Touzi, and J. Zhang. *Wellposedness of Second Order Backward SDEs*, Probability Theory and Related Fields, **153**, 149–190, (2012)

[23] C. Villani. *Topics in optimal transportation*, volume 58 of Graduate Studies in Mathematics. AMS, Providence, RI, (2003).

[24] C. Villani. *Optimal Transport. Old and New*, volume 338 of Grundlehren der mathematischen Wissenschaften. Springer, (2009).

DEPARTMENT OF STATISTICS, HEBREW UNIVERSITY OF JERUSALEM, ISRAEL. E.MAIL: YAN.DOLINSKY@MAIL.HUJI.AC.IL

DEPARTMENT OF MATHEMATICS, ETH ZURICH & SWISS FINANCE INSTITUTE. E.MAIL: HMSONER@ETHZ.CH