Risk Minimization under Transaction Costs

Paolo Guasoni
Scuola Normale Superiore
P.zza dei Cavalieri, 7
56126 Pisa - Italy
guasoni@sns.it

November 12, 1999

Abstract

We study the general problem of an agent wishing to hedge some liability at a fixed date in a market with a risky asset, where proportional transaction costs are present. Other frictions, such as incomplete information and constraints on strategies, may also be present.

We obtain a minimization problem on a space of predictable processes with finite variation. Borrowing a technique from Calculus of Variation, on this space we look for a convergence where minimizing sequences are compact, and risk is lower semicontinuous. This leads to the class of convex decreasing risk functionals, containing *coherent risk measures*.

If the risky asset is a martingale, we show the existence of an optimal strategy for the problems of *shortfall minimization* and *utility maximization*. Finally, we make a partial extension of these results to the semimartingale case.

1 Introduction

The problem of hedging under transaction costs has raised interest among academics in several fields, resulting in a large number of papers on this subject, ranging from purely theoretical studies to lots of empirical papers, concerned with the evaluation of specific trading strategies.

Remarkably, market frictions generate highly nontrivial problems on both the theoretical and the applied sides, leaving the practitioner without a general theory, and the theoretician with few helpful heuristic arguments.

On the applied side, Leland [18] went directly to the main question faced by options traders, and proposed an explicit formula for *adjusted volatility*, to be used for hedging at fixed time intervals. His work sparked interest in different directions: Kabanov and Safarian [?] proved rigorously under what conditions Leland's strategies converge to a perfect replication. On the other hand, Hodges and Neuberger [15] called for a formalization of the problem within stochastic control theory, arguing that the Leland's strategies are in no sense optimal.

In this framework, Davis, Panas, and Zariphopoulou [8] gave the first mathematically rigorous analysis of the problem of hedging a European claim in the Black-Scholes model with proportional transaction costs. Indeed, their proof of the existence of an optimal strategy relies on the viscosity theory, which provides solutions in a weak sense to a wide range of PDEs. In fact, optimal strategies turn out to miss any differentiability property, as it was already recognized in the consumption-investment problem by Davis and Norman [7]. This casts a shadow on analogous problems where the underlying is not a Markov process, since they might no longer be written in terms of PDEs.

Meanwhile, the mathematical theory of derivatives in markets with incomplete information has proceeded at an impressive pace, settling some fundamental questions such as the link between martingale measures and absence of arbitrage, and the identification of the right spaces of strategies where optima should be looked for. When transaction costs are introduced, these questions seem still open, at least in their generality.

This paper addresses the problem of hedging a claim H expiring at a fixed date T in a market with a risky asset X, where proportional transaction costs are present. We also allow other frictions, such as incomplete information, and constraints on strategies. An optimal strategy is defined as a minimizer of a risk functional ρ, mapping the profit an loss distribution at expiration to a real number. This framework encompasses different problems, such as hedging a claim expiring at time T, or maximizing utility from terminal wealth.

The paper is organized as follows: in section 2 we describe in details our model for a market with transaction costs. Section 3 is devoted to the spaces of admissible strategies, where we look for a convergence providing both sufficiently many compact sets, and reasonable (semi)continuity properties for the portfolio value. Section 4 contains the main results on the existence of optimal strategies in markets with incomplete information and transaction costs, but only when X is a martingale. Section 5 handles the constrained case, while the last section discusses the problems arising when X is a semimartingale. Assuming the existence of a martingale measure (which is somewhat arbitrary in presence of transaction costs), we extend some of the existence results to the semimartingale case.

We conclude recalling in the appendix a few not-so-popular results in functional analysis and stochastic integration which are used extensively in the present paper.

The author wishes to thank M. Pratelli for his constant advice and encouragement through the all the stages of this research.

2 The Model

Consider a simple market with a risky asset X and a riskless asset B. X is a continuous martingale defined on a probability space $(\Omega, \mathcal{F}, P)$, adapted with respect to a given filtration $\{\tilde{\mathcal{F}}_t\}_{t\in[0,T]}$, which satisfies the usual hypotheses. We also assume that $\mathcal{F} = \tilde{\mathcal{F}}_T$. If B is a deterministic (as we shall assume), without loss of generality we can set $B = 1$, since it is the same as replacing X by $\frac{X}{B}$.

In this market, an agent starts with some initial capital c, and faces a certain contingent liability H at time T. His/Her goal is to set up a self-financing portfolio which minimizes the total risk at time T.

Since we only have two assets, the self-financing condition implies that a trading strategy is uniquely determined by the number of shares θ_t invested in the risky asset at time t.

The agent can make trading decisions only based on partial information, therefore the stochastic process θ_t must be adapted to the filtration $\mathcal{F}_t$, which is generally smaller than $\tilde{\mathcal{F}}_t$.

In this market continuous trade is allowed without constraints (constrained problems will be considered in Section 5), but proportional transaction costs are present. A simple (i.e. a piecewise constant) strategy is defined as follows:

$$\theta_t = \sum_{i=1}^{n} \theta^i 1_{[\tau_i,\tau_{i+1})}$$

where $\tau_0 \equiv 0$, $\tau_1 < \cdots < \tau_n \leq T$ is a finite set of stopping times with respect to $\mathcal{F}^X$, and each random variable θ^i is τ_i-measurable (in particular, θ^0 is a constant). The cost of such a strategy up to time t is given by:

$$C_t(\theta) = k_0|\theta_0| + \sum_{\tau\leq t} k_{\tau_i}|\theta_{\tau_i} - \theta_{\tau_{i-1}}| \tag{1}$$

where the process k_t is adapted to $\mathcal{F}_t$, and determines the cost scheme. For example, costs are proportional to the number of shares if k_t is a constant, or to the amount traded if $k_t = kX_t$, for some constant k.

If we approximate a general θ with simple strategies θ^n, and we let the mesh of the partition converge to zero, we have:

$$C_t(\theta^n) \to k_0|\theta_0| + \int_{(0,t)} k_s d|D\theta|_s \tag{2}$$

where $|D\theta|(\omega)$ denotes the variation measure associated to $\theta(\omega)$. Note that the integral is defined pathwise, in the usual Stiltjes sense.

This argument provides a convenient extension for the cost of a general trading strategy: in fact we take the right-hand side of (2) as the definition of total transaction cost.

Note that $C_T(\theta)$ does not include the cost at time T, which may or may not be present, depending on the settlement provisions of particular contracts. Since we are interested in general hedging problems, we choose not to take the final cost into account.

For convenience, we assume that $\theta_t = 0$ for $t < 0$ and $\theta_t = \theta_T$ for $t > T$. With this convention, denoting the interval $[0, T]$ by I, we have

$$C_T(\theta) = \int_I k_t d|D\theta_t|$$

In this setting, the portfolio value at time t associated to the trading strategy θ_t is given by the initial capital, plus the trading gain, minus the transaction cost, namely:

$$V_t^c(\theta) = c + G_t(\theta) - C_t(\theta)$$

where

$$G_t(\theta) = \int_0^t \theta_s dX_s \qquad \text{and} \qquad C_t(\theta) = \int_{[0,t]} k_s d|D\theta|_s$$

From the simple inequality:

$$C_T(\theta) \geq (\inf_{t \in I} k_t)|D\theta|(I)$$

we see that, if the infimum of k_t is a.s. strictly positive (as it is usually the case), then $E[C_T(\theta)] < \infty$ implies $P(|D\theta|(I) < \infty) = 1$. In other words, almost every path of θ must be of bounded variation.

Since our goal is to hedge a liability H, with $E[|H|] < \infty$, there is no need to consider strategies of unbounded variation. In fact, any of them would lead to an even bigger liability at expiration.

Now we need a mean for ranking strategies. Indeed, available options include expected utility, VaR, risk measures and many other criteria, none of which is universally preferred to the others. For a thorough discussion on this topic, we refer the interested reader to Artzner, Delbaen, Eber and Heath [2], where the class of *coherent risk measures* is proposed on a number of economic justifications.

We shall consider risk functionals $\rho : L^p \mapsto \mathbb{R}$, where ρ satisfies the following properties:

i) ρ is convex;

ii) if $X \leq Y$ a.e., then $\rho(X) \geq \rho(Y)$ (ρ is decreasing);

iii) ρ has the Fatou property. In other words, if $X_n \to X$ a.e., then

$$\rho(X) \leq \liminf_{n\to\infty} \rho(X_n)$$

In particular, this class includes the *shortfall*, for $\rho(X) = E[X^-]$, and σ-additive *coherent risk measures* (see Delbaen [9] for details), for $\rho(X) = \sup_{P\in\mathcal{P}} E_P[-X]$, where $\mathcal{P}$ is a set of probabilities. *Utility maximization* can also be embedded in this framework, minimizing the functional $\rho(X) = -E[U(X)]$, where U is the utility function.

Remark 2.1. In presence of transaction costs, some authors prefer to define a strategy as a pair of increasing processes (L_t, M_t), denoting the cumulative purchases and sales. For instance, in Cvitanić and Karatzas [5] L_t and M_t denote amounts of money transferred, while in Davis, Panas, and Zariphopoulou [8] numbers of shares traded.

It is easily seen that these definitions are equivalent to ours: in fact, the difference between two increasing processes is obviously of bounded variation. Viceversa, any function of bounded variation (and hence any process) can be represented as a difference of two increasing functions, which are uniquely determined under the assumption that they do not simultaneously increase (see AFP [1] for details). From the financial point of view, this is natural to ask, since it prevents opposite transactions from taking place at the same time. Furthermore, for a $\theta \in BV([0,T])$, the associated increasing functions can be explicitly written as:

$$f(t) = D\theta^+([0,t]) \quad \text{and} \quad g(t) = D\theta^-([0,t])$$

where $D\theta^+$ and $D\theta^-$ are respectively the positive and negative parts of the Radon measure $D\theta$.

Remark 2.2. In our framework, the costs associated to the purchase and sale of the risky asset are equal. This is not the case in several papers on proportional transaction costs, where the costs for buying and selling one share of the risky asset are respectively λX_t and μX_t (for example, see Davis, Panas, and Zariphopoulou [8] and Cvitanić and Karatzas [5]). This feature can be embedded in our model via the alternative definition of $C_T(\theta)$:

$$C_T^{\lambda,\mu}(\theta) = \int_I k_t d(\lambda D\theta^+ + \mu D\theta^-)$$

Since we have that $|D\theta| = D\theta^+ + D\theta^-$, for $\lambda = \mu$ the above definition boils down to (2), up to a rescaling of k. In fact, the quantity $\lambda D\theta^+ + \mu D\theta^-$ shares all the relevant properties of $|D\theta|$, such as weak lower semicontinuity, and all the discussion that follows on $|D\theta|$ should be easy to adapt to the case $\lambda \neq \mu$.

In practice, given two processes for the bid and the ask prices, one can always obtain a model where $\lambda = \mu$, choosing X_t as the midpoint between the bid and the ask, and $2k_t$ equal to the bid-ask spread. Since this convention greatly simplifies notation, throughout the paper we shall only consider the case $\lambda = \mu$.

Remark 2.3. The pointwise definition (1) of variation can be modified into the following (much less intuitive), which is invariant up to sets of Lebesgue measure zero:

$$|D\theta|(\omega) = \sup_{\substack{\phi \in C_c^1(0,T) \\ \|\phi\|_\infty \leq 1}} \int_I \theta_s(\omega)\phi'(s)ds$$

In fact it can be shown that for each θ there exists a representative such that the expression above coincides with the (generally higher) pointwise variation. For details, see AFP [1].

3 Strategies with Transaction costs

In the literature on incomplete markets, the following spaces of strategies are often considered, especially for $p = 2$:

$$\Theta^p = \{\theta : \theta \ \mathcal{F}_t\text{-predictable}, \ G_T(\theta) \in L^p(P)\}$$

Θ^p can be endowed with the topology induced by the map $\theta \mapsto \int_0^T \theta_t dX_t$. It is clear that $G_T(\Theta^p)$ is a linear subspace of L^p: it turns out that it is also closed, as it follows from Theorems 6.9 and 6.10, in the cases $p > 1$ and $p = 1$ respectively.

The spaces Θ^p are usually the natural candidates for a hedging problem in a frictionless market, but transaction costs force a much narrower set of admissible strategies. As we argued in the previous section, in a market with proportional costs we should only consider strategies with finite variation. This leads us to define the following spaces:

$$\Theta^p_C = \{\theta \in \Theta^p, \ C_T(\theta) \in L^p(P)\}$$

endowed with the norm:

$$\|\cdot\|_{\Theta^p_C} : \ \theta \to \left(\left\| \int_0^T \theta_t dX_t \right\|_p^p + \left\| \int_I k_t d|D\theta| \right\|_p^p \right)^{\frac{1}{p}} =$$

$$= \left(\|G_T(\theta)\|_p^p + \|C_T(\theta)\|_p^p \right)^{\frac{1}{p}}$$

We begin our discussion with the following result:

Proposition 3.1. *Let X be a continuous martingale, and k a continuous, $\mathcal{F}^X$-adapted process, such that $\tilde{k}(\omega) = \min_{t \in I} k_t(\omega) > 0$ for a.e. ω. Then Θ_C^p is a Banach space for all $p \geq 1$.*

The proof requires a few lemmas, which will be useful in the sequel:

Lemma 3.2. *For a fixed ω, let $\theta^n(\omega)_t \to \theta(\omega)_t$ for a.e. t, and $|D\theta(\omega)^n|(I) < C$ uniformly in n. Then $D\theta(\omega)^n \rightharpoonup D\theta(\omega)$ in the weak star topology of Radon measures.*

Proof. For all $\phi \in C_c^\infty[0,T)$, we have:

$$\nu(\phi) = \lim_{n \to \infty} \int_{[0,T)} \phi_t dD\theta_t^n = -\lim_{n \to \infty} \int_I \phi_t' \theta_t^n dt = -\int_I \phi_t' \theta_t dt = D\theta(\phi)$$

It remains to show that the distribution $D\theta$ is in fact a Radon measure, and this follows from the inequality:

$$D\theta(\phi) \leq \sup_{t \in I} |\phi(t)| \limsup_{n \to \infty} |D\theta^n|(I) \leq C \sup_{t \in I} |\phi(t)|$$

which completes the proof. □

The following is a standard result in measure theory (see for instance AFP [1]):

Lemma 3.3. *Let $\mu^n \rightharpoonup \mu$, where μ^n, μ are Radon measures on I, and convergence is meant in the weak star sense. Then $|\mu| \leq \liminf_{n \to \infty} |\mu^n|$.*

The next lemma states that, roughly speaking, taking limits can only reduce transaction costs. Intuitively, this is the case because in the limit strategy some transactions may cancel out, while no new ones can arise. As we shall see later, this has interesting consequences on the type of risk functions and constrained problems that we are able to solve.

Lemma 3.4. *If θ^n is bounded in Θ_C^p, and $\theta_t^n \to \theta_t$ a.e. in $dtP(d\omega)$ then:*

$$C_T(\theta) \leq \liminf_{n \to \infty} C_T(\theta^n) \qquad \text{for a.e. } \omega \tag{3}$$

and

$$\|C_T(\theta)\|_p \leq \liminf_{n \to \infty} \|C_T(\theta^n)\|_p \tag{4}$$

for all $p \geq 1$.

Proof. By assumption, for a.e. ω, $\theta^n_t(\omega) \to \theta_t(\omega)$ for a.e. t. To prove (3), we show that for all subsequences n_j for which $C_T(\theta^{n_j}(\omega))$ converges, we have:

$$C_T(\theta(\omega)) \leq \lim_{j\to\infty} C_T(\theta^{n_j}(\omega)) \tag{5}$$

If $C_T(\theta^{n_j}(\omega)) \to \infty$, then (5) is trivial. If not, then $C_T(\theta^{n_j}(\omega)) < M(\omega)$ for all j and hence

$$|D\theta^{n_j}(\omega)|(I) < \frac{M(\omega)}{\tilde{k}(\omega)}$$

Lemma 3.2 implies that $D\theta^{n_j}(\omega) \rightharpoonup D\theta(\omega)$. By Lemma 3.3, we obtain:

$$C_T(\theta(\omega)) = \int_I k_t d|D\theta(\omega)| \leq \lim_{j\to\infty} \int_I k_t d|D\theta^{n_j}(\omega)| = \lim_{j\to\infty} C_T(\theta^{n_j}(\omega))$$

and (3) follows. For (4), notice that:

$$\|C_T(\theta(\omega))\|_p^p = E\left[\left(\int_I k_t d|D\theta(\omega)|\right)^p\right] \leq E\left[\liminf_{n\to\infty}\left(\int_I k_t d|D\theta^n(\omega)|\right)^p\right] \leq$$
$$\leq \liminf_{n\to\infty} E\left[\left(\int_I k_t d|D\theta^n(\omega)|\right)^p\right] = \liminf_{n\to\infty} \|C_T(\theta(\omega))\|_p^p < \infty$$

where the last inequality follows from the uniform boundedness of θ^n in Θ^p_C, and the previous one holds by Fatou's Lemma. $\square$

Proof of Proposition 3.1. i) Θ^p_C is a linear subspace of Θ^p, therefore $\|G_T(\theta)\|_p$ is a norm. Hence it is sufficient to prove that $\|C_T(\theta)\|_p$ is also a norm, and that the space is complete.

Trivially, $\|C_T(\lambda\theta)\|_p = |\lambda|\,\|C_T(\theta)\|_p$. Let $\theta, \eta \in \Theta^p_C$. We have:

$$\|C_T(\theta+\eta)\|_p = \left\|\int_I k_t d|D(\theta+\eta)|_t\right\|_p \leq$$
$$\leq \left\|\int_I k_t d|D\theta| + \int_I k_t d|D\eta|_t\right\|_p \leq \|C_T(\theta)\|_p + \|C_T(\eta)\|_p$$

It remains to show that Θ^p_C is complete. Let θ^n be a Cauchy sequence for Θ^p_C: since it is also Cauchy in Θ^p, which is complete, we can assume that $\theta^n \to \theta$ in Θ^p. We need to show that $\|C_T(\theta)\|_p < \infty$.

By assumption, $C_T(\theta^n - \theta) \to 0$ in L^p thus, up to a subsequence, convergence holds a.e. Since $C_T(\theta^n - \theta) = \int_0^T k_t d|D(\theta^n - \theta)|$, and k_t is strictly positive, it follows that, for a.e. ω, $|D(\theta^n - \theta)|$ converges to 0 in the weak star topology of Radon measures. As a result,

$$\lim_{n\to\infty} |D(\theta^n - \theta)(\omega)|([0,T]) = 0 \quad \text{for a.e. } \omega$$

Finally, we observe that:

$$\|\theta_s^n - \theta_s\|_\infty \leq |D(\theta^n - \theta)|([0,T])$$

which implies that $\theta^n - \theta$ converges to 0 in L^∞ for a.e. ω.

By Lemma 3.4, $\|C_T(\theta)\|_p < \infty$, and the proof is complete. $\square$

The following is a straightforward consequence of Lemma 3.4, which is nonetheless worth noting:

Corollary 3.5. *If θ^n is bounded in Θ^p_C, and $\theta^n_t \to \theta_t$ a.e. in $dtP(d\omega)$ then:*

$$V_T^c(\theta) \geq \limsup_{n\to\infty} V_T^c(\theta^n)$$

for $dtP(d\omega)$-a.e. ω and in L^p.

Remark 3.6. Θ^p_C is not a Hilbert space even for $p = 2$. In fact, it easily checked that the equality $|\theta+\eta|^2+|\theta-\eta|^2 = 2|\theta|^2+2|\eta|^2$, which is valid in any Hilbert space, is not satisfied from the deterministic strategies $\theta_t = 1_{\{t<\frac{1}{3}T\}}$ and $\eta_t = 1_{\{t\geq\frac{2}{3}T\}}$.

Remark 3.7. Θ^p_C is not *separable.* To see this, observe that the set of deterministic strategies $\{\theta^x\}_{x\in I}$, where $\theta^x_t = 1_{\{t\geq x\}}$, is uncountable, and $\|\theta^x - \theta^y\|_p >$ 1 for all $x \neq y$ and for all $p \geq 1$.

The following inequality states a continuous immersion of Θ^r_C into Θ^p, for $r > p$, provided that $\frac{\langle X\rangle_T^{\frac{1}{2}}}{\tilde{k}}$ is sufficiently integrable.

Proposition 3.8. *For any $p, q \geq 1$, we have:*

$$\|G_T(\theta)\|_p \leq \|C_T(\theta)\|_{pq} \left\| \frac{\langle X\rangle_T^{\frac{1}{2}}}{\tilde{k}} \right\|_{pq'}$$

where $q' = \frac{q}{q-1}$.

Proof. For any p, we have, by the Burkholder-Davis-Gundy inequality:

$$E\left[|G_T(\theta)|^p\right] \leq E\left[\left(\int_0^T \theta_t^2 d\langle X\rangle_t\right)^{\frac{p}{2}}\right] \leq E\left[\left(\sup_{t\leq T}|\theta_t|\right)^p \langle X\rangle_T^{\frac{p}{2}}\right] \leq$$
$$\leq E\left[|D\theta|(I)^p \langle X\rangle_T^{\frac{p}{2}}\right] \leq E\left[C_T(\theta)^p \left(\frac{\langle X\rangle_T^{\frac{1}{2}}}{\tilde{k}(\omega)}\right)^p\right] \leq$$
$$\leq E\left[C_T(\theta)^{pq}\right]^{\frac{1}{q}} E\left[\left(\frac{\sqrt{\langle X\rangle_T}}{\tilde{k}(\omega)}\right)^{pq'}\right]^{\frac{1}{q'}}$$

and raising both sides to the power $\frac{1}{p}$ the thesis follows. $\square$

Remark 3.9. Proposition 3.8 admits a simple financial interpretation. The transaction cost needed for a gain with a high moment of order p is bounded if the asset itself has a sufficiently high moment of the same order (a buy and hold strategy does the job). Otherwise, the strategy itself must amplify the swings of the market. In this case, the less the market is volatile, the higher the moment of the strategy.

Remark 3.10. While for $p > 1$ the $\mathcal{H}^p$ norm of a uniformly integrable martingale is equivalent to the L^p norm of its final value, it is well-known that this is no longer true for $p = 1$. In this case, the estimate in Proposition 3.8 holds not only for the L^1 norm of $G_T(\theta)$, but also for its $\mathcal{H}^1$ norm, as it is immediately seen by the proof.

Remark 3.11. The integrability condition for $\tilde{X}$ in 3.8 seems not too restrictive. For example, if k is a constant, it boils down to an integrability condition for X. In most models of financial markets, the asset X has finite moments of any order, therefore is is automatically satisfied.

On the other hand, if $k_t = kX_t$, the moments of $\tilde{k}$ can be computed in terms of X_t, and integrability can be obtained via the Hölder inequality.

Remark 3.7 suggests that the norm topology in Θ^p_C is too restrictive to provide sufficient compactness on the space of strategies. The following lemma provides a partial remedy for our situation:

Lemma 3.12. *Let $G_T(\theta^n) \to G_T(\theta)$ in the L^p-norm. Then:*

i) if $p > 1$, up to a subsequence $\theta^n \to \theta$ a.e. in $d\langle X\rangle_t P(d\omega)$;

ii) if $p = 1$, there exists some η^n, convex combinations of stoppings of θ^n, such that $\eta^n \to \theta$ a.e. in $d\langle X\rangle_t P(d\omega)$.

Proof. $i)$ The Burkholder-Davis-Gundy inequality yields:

$$E\left[|G_T(\theta^n) - G_T(\theta)|^p\right] \geq c_p E\left[\left|\int_0^T (\theta^n_t - \theta_t)^2 d\langle X\rangle_t\right|^{\frac{p}{2}}\right] \tag{6}$$

For some positive constant c_p. Since the left-hand side converges to zero, it follows that $\int_0^T (\theta^n_t - \theta_t)^2 d\langle X\rangle_t$ also converges to zero in probability, and $\theta^n \to \theta$ in the measure $d\langle X\rangle_t(\omega)P(d\omega)$. Up to a subsequence, convergence holds a.e.

$ii)$ The situation is more delicate here, because (6) is not true in L^1. Denote by $\tau_k = \inf\{t : \ |G_T(\theta)| \geq k\}$. The stopped martingales $G_{T\wedge\tau_k}(\theta)$ clearly converge to $G_T(\theta)$ almost surely. For each k, since $G_{T\wedge\tau_k}(\theta) \in \mathcal{H}^1$, we can apply Corollary 6.13 , and obtain that for some stopping times $T_{n,k}$, $G_{T_{n,k}}(\theta^n) \rightharpoonup G_{T\wedge\tau_k}(\theta)$ in $\sigma(\mathcal{H}^1, BMO)$. Up to a sequence of convex combinations $\xi^k_n = \sum_{j=n}^{M_n} \beta^{k,n}_j G_{T_{j,k}}(\theta^j)$, we can assume that $\xi^n_k \to G_{T\wedge\tau_k}(\theta)$ in the strong topology of $\mathcal{H}^1$. Observe also that if $k' < k$, then $(\xi_k^n)^{\tau_{k'}} \to G_{T\wedge\tau_{k'}}(\theta)$: hence, by a diagonalization argument, we consider the sequence η^n such that $G_T(\eta^n) = \xi_n^n$, which satisfies the condition $G_{T\wedge\tau_k}(\eta^n) \to G_{T\wedge\tau_k}(\theta)$ in the $\mathcal{H}^1$ norm for all k. In other words:

$$\lim_{n\to\infty} E\left[\left(\int_0^{\tau_k} (\eta_t^n - \theta_t)^2 d\langle X\rangle_t\right)^{\frac{1}{2}}\right] = 0$$

This means that $\eta^n \to \theta$ in the measure $d\langle X^{\tau_k}\rangle_t P(d\omega)$, and up to a subsequence, a.e. Since $d\langle X^{\tau_k}\rangle_t P(d\omega)$ is an sequence of measures increasing to $d\langle X\rangle_t P(d\omega)$, we conclude that $\eta^n \to \theta$ a.e. in $d\langle X\rangle_t P(d\omega)$. $\square$

4 Existence of Optimal Strategies

This section contains the main existence results for optimal hedging strategies in unconstrained incomplete markets with proportional transaction costs.

In general, the existence of a minimum requires two basic ingredients: relative compactness of minimizing sequences (up to some transformation which leaves them minimizing), and lower semicontinuity of the functional.

Compactness is obviously much easier in L^p spaces with $p > 1$, since it coincides with boundedness, but this kind of information is rare to obtain in applications. On the other hand, some measures of risk (maximization of utility and minimization of shortfall) seem to provide natural bounds on the L^1 norms of optimizing portfolios at expiration.

Moreover, in the next section we shall see that if X is a semimartingale, then a minimization problem can be reduced through a change of measure to a problem in L^1. This shows that the L^1 case is both mathematically more challenging, and the most relevant in applications.

We also need the functional $F : \theta \mapsto \rho(V_T^c(\theta) - H)$ to be lower semicontinuous (shortly l.s.c.). Corollary 3.5 shows that in general $V_T^c(\cdot)$ is upper semicontinuous with respect to a.s. convergence in $dtP(d\omega)$, but not necessarily continuous. This means that we need a decreasing ρ to ensure the semicontinuity of F. Also, we are going to take convex combinations of minimizing strategies, and we need a convex ρ to leave them minimizing. This leads to the following

Definition 4.1. We define a *convex decreasing risk functional* as a function $\rho : L^p \mapsto \mathbb{R}$, satisfying the following properties:

i) ρ is convex;

ii) if $X(\omega) \leq Y(\omega)$ for a.e. ω, then $\rho(X) \geq \rho(Y)$ (ρ is decreasing);

iii) ρ has the Fatou property. Namely, if $X_n \to X$ a.e., then

$$\rho(X) \leq \liminf_{n\to\infty} \rho(X_n)$$

We easily see that the above definition provides the desired properties of semicontinuity and convexity:

Lemma 4.2. *Let ρ be a convex decreasing functional, $c > 0$ and $H \in L^p(\Omega, \mathcal{F}_T)$. Denoting by $F : \theta \mapsto \rho\left(V_T^c(\theta) - H\right)$, if $\theta^n \to \theta$ a.e. in $dtP(d\omega)$, we have:*

i) F is convex;

ii) F is l.s.c. with respect to a.s. convergence in $dtP(d\omega)$.

Proof. $i)$ Since ρ is convex decreasing, and the function $V_T^c(\cdot)$ is concave, it follows that $F = \rho(\cdot - H) \circ V_T^c(\cdot)$ is convex.

$ii)$ By Lemma 3.5, we have:

$$V_T^c(\theta) \geq \limsup_{n\to\infty} V_T^c(\theta^n)$$

Since ρ is decreasing:

$$\rho(V_T^c(\theta) - H) \leq \rho(\limsup_{n\to\infty} V_T^c(\theta^n) - H)$$

and finally, by the Fatou property of ρ:

$$\rho(V_T^c(\theta) - H) \leq \liminf_{n\to\infty} \rho(V_T^c(\theta^n) - H)$$

□

For convex decreasing functionals we are going to prove an existence result on bounded sets of Θ_C^p. Examples include σ-additive *coherent risk measures* (see Delbaen [9] for details).

A special class of these functionals consists of those which can be written as $\rho(X) = E\left[\nu(X)\right]$, where $\nu : \mathbb{R} \mapsto \mathbb{R}$ is a convex decreasing function. In this case, we show that an optimal strategy exists in the whole space Θ_C^1, since minimizing sequences are automatically bounded. Both the problems of *shortfall* minimization and *utility* maximization belong to this class.

Throughout this section, we assume that the two measures $d\langle X\rangle_t$ and dt are equivalent. This is necessary to draw inference on $C_T(\theta^n)$, which depends on convergence with respect to the measure $dtP(d\omega)$, from the convergence of $G_T(\theta^n)$, which provides information in the measure $d\langle X\rangle_t P(d\omega)$. This assumption implies that X cannot have intervals of constancy, and that its bracket process $\langle X\rangle_t$ cannot exhibit a Cantor-ladder behavior. In practice, it is satisfied by all models defined by means of diffusions, even with Hölder coefficients or volatility jumps.

We start with risk minimization in Θ^p_C, with $p > 1$. In this case boundedness is among the assumptions, therefore the minimum will generally depend on the particular bound considered.

The next lemma provides a class of weakly compact sets:

Lemma 4.3. *For $C > 0$ and $p > 1$, the set*

$$B_{C,D} = \{\theta : \|G_T(\theta)\|_p \leq C, \|C_T(\theta)\|_p \leq D\}$$

is weakly compact in Θ^p_C for $D \in \mathbb{R}^+ \cup \{+\infty\}$.

Proof. Let $\theta^n \in B_{C,D}$. For $p > 1$, Θ^p is a reflexive Banach space (it is isometric to a closed subspace of L^p, which is reflexive). Hence, the set

$$B_C = \{\theta : \|G_T(\theta)\|_p \leq C\}$$

is weakly compact in Θ^p, and up to a subsequence $G_T(\theta^n) \rightharpoonup G_T(\theta) \in B_C$. Since $B_{C,D}$ is convex, by Theorem 6.7 there exists a sequence $\eta^n \in B_{C,D}$ of convex combinations of θ^n, such that $G_T(\eta^n) \to G_T(\theta)$ in L^p. By Lemma 3.12, it follows that, up to a subsequence, $\eta^n \to \theta$ a.e. in $d\langle X\rangle_t P(d\omega)$, and by Lemma 3.4, we conclude that $\theta \in B_{C,D}$. □

Proposition 4.4. *Let ρ be a convex decreasing functional, $c > 0$ and $H \in L^p(\Omega, \mathcal{F}_T)$. For any $M > 0$ let us denote*

$$\Theta^p_{C,M} = \{\theta \in \Theta^p_C,\ \|G_T(\theta)\|_p \leq M\}$$

Then the following minimum problem admits a solution:

$$\min_{\theta \in \Theta^p_{C,M}} \rho\left(V^c_T(\theta) - H\right)$$

Proof. Let θ^n be a minimizing sequence, so that $F(\theta^n) \to \inf_{\theta \in \Theta^p_{C,M}} F(\theta)$. Since $\Theta^p_{C,M}$ is weakly compact by Lemma 4.3, up to a subsequence we can assume that $\theta^n \rightharpoonup \theta \in \Theta^p_{C,M}$. Then, by Theorem 6.7, there exists a sequence of convex combinations $\eta^n = \sum_{k=n}^\infty \alpha^n_k \theta^k$, such that $\eta^n \to \theta$ in the strong topology. By Lemma 3.12, we can assume up to a subsequence that $\eta^n \to \theta$ in the $d\langle X\rangle_t P(d\omega)$-a.s. convergence, and by Jensen's inequality:

$$F(\eta^n) \leq \sum_{k=n}^{\infty} \alpha^n_k F(\theta^k) \leq \max_{n \leq k} F(\theta^k)$$

Passing to the limit:

$$\lim_{n\to\infty} F(\eta^n) \leq \lim_{n\to\infty} \max_{n \leq k} F(\theta^k) = \lim_{n\to\infty} F(\theta^n)$$

Finally, by the semicontinuity of F, we obtain:

$$F(\theta) \leq \lim_{n\to\infty} F(\theta^n)$$

hence θ is a minimum. $\square$

We now turn to risk minimization in Θ^1_C. As mentioned before, this case has the advantage that minimizing sequences are bounded for some problems considered in applications. On the other hand, a few mathematical issues arise: a bounded sequence in L^1 does not necessarily converge, even in a weak sense, and the L^1 norm of a uniformly integrable martingale is not equivalent to the $\mathcal{H}^1$ norm.

It turns out that the first problem can be overcome through a result of Komlós [17] (see also Schwartz [20], for a shorter proof). We can circumvent the latter at the price of using stopping as a further transformation on minimizing sequences, besides extracting subsequences and taking convex combinations.

We start with the existence result for general convex decreasing risks:

Proposition 4.5. *Let ρ be a convex decreasing functional, $c > 0$ and $H \in L^1(\Omega, \mathcal{F}_T)$. For any $M > 0$ the following minimum problem admits a solution:*

$$\min_{\theta\in\Theta^1_{C,M}} E\left[\rho\left(V^c_T(\theta) - H\right)\right]$$

Proof. By Komlós' Theorem (6.8), up to a subsequence of convex combinations $\eta^n = \sum_{k=n}^{M_n} \alpha^n_k \theta^k$ we can assume that $G_T(\eta^n) \to \gamma$ a.e. and in L^1, and by Yor's Theorem 6.10, there exists some $\theta \in \Theta^1$ such that $\gamma = G_T(\theta)$.

To see that $\theta \in \Theta^1_C$, first we apply Lemma 3.12, obtaining a sequence $\zeta^n = \sum_{j=n}^{M_n} \beta^n_j (\eta^j)^{T_{j,n}}$, such that $\zeta^n \to \theta$ a.e. Then Lemma 3.4 implies that $C_T(\theta) \in L^1(P)$, as required.

By Jensen's inequality, we have:

$$F(\zeta^n) \leq \sum_{j=n}^{M_n} \beta^n_j (\eta^j)^{T_{j,n}} \leq \max_{n\leq j} F((\eta^j)^{T_{j,n}})$$
$$F(\eta^n) \leq \sum_{k=n}^{M_n} \alpha^n_k F(\theta^k) \leq \max_{n\leq k} F(\theta^k)$$

and, passing to the limit:

$$\lim_{n\to\infty} F(\zeta^n) \leq \lim_{n\to\infty} \max_{n\leq j} F((\eta^j)^{T_{j,n}}) = \lim_{n\to\infty} F(\eta^n)$$
$$\lim_{n\to\infty} F(\eta^n) \leq \lim_{n\to\infty} \max_{n\leq k} F(\theta^k) = \lim_{n\to\infty} F(\theta^n)$$

Finally, by the semicontinuity of F, we obtain:

$$F(\theta) \leq \lim_{n\to\infty} F(\theta^n)$$

hence θ is a minimum. $\square$

In the special case of ρ being the expectation of a convex decreasing function ν, it turns out that the optimal strategy in Θ^1_C coincides with that in $\Theta^1_{C,M}$, for some value of M. This is shown in the following

Proposition 4.6. *Let $\nu : \mathbb{R} \mapsto \mathbb{R}$ a strictly convex (at least in one point) decreasing function, $c > 0$ and $H \in L^1(\Omega, \mathcal{F}_T)$. Then the problem*

$$\min_{\theta\in\Theta^1_C} E\left[\nu\left(V^c_T(\theta) - H\right)\right]$$

admits a solution.

Proof. Let θ^n be a minimizing sequence, so that $F(\theta^n) \to \inf_{\theta\in\Theta_{C,M}} F(\theta)$. Since ν is strictly convex, we have that:

$$\nu(x) \geq a + bx^- - (b-\varepsilon)x^+ \tag{7}$$

which implies

$$E\left[Y^+\right] \leq \frac{1}{\varepsilon}\left(E\left[\nu(Y)\right] - a + bE\left[Y\right]\right)$$

for any integrable random variable Y. Substituting $Y = V^c_T(\theta^n) - H$, we get:

$$E\left[(V^c_T(\theta^n) - H)^+\right] \quad \leq \quad \frac{1}{\varepsilon}\left(E\left[\nu(V^c_T(\theta^n) - H)\right] - a + bE\left[V^c_T(\theta^n) - H\right]\right)$$

The first term in the right-hand side is bounded by assumption, since θ^n is a minimizing sequence. The second term is also bounded, because $E\left[V^c_T(\theta^n)\right] = c - E\left[C_T(\theta^n)\right] \leq c$ and H is integrable. Therefore $E\left[(V^c_T(\theta^n) - H)^+\right]$ is bounded and the inequality

$$E\left[V^c_T(\theta^n)^+\right] \leq E\left[(V^c_T(\theta^n) - H)^+\right] + E\left[H^+\right]$$

implies that $E\left[V^c_T(\theta^n)^+\right]$ is bounded. In a similar fashion, (7) yields:

$$E\left[Y^-\right] \leq \frac{1}{b}(E\left[\nu(Y)\right] + (b-\varepsilon)E\left[Y^+\right] - a)$$

Substituting again $Y = V^c_T(\theta^n) - H$, and using the foregoing result, we conclude that $\sup_n E\left|V^c_T(\theta^n)\right| < \infty$. Also:

$$E\left|V^c_T(\theta^n)\right| \geq E\left[-V^c_T(\theta^n)\right] = -c + E\left[C_T(\theta^n)\right]$$

which implies that $\sup_n E\left[C_T(\theta^n)\right] < \infty$, and hence $\sup_n E\left[|G_T(\theta^n)|\right] < \infty$. Now that boundedness is shown, the thesis follows from Proposition 4.5 for a suitable M.

$\square$

Example 4.7 (Shortfall risk). In Proposition 4.6, choosing $\nu(x) = x^-$, we obtain the existence of a *shortfall minimizing* strategy, that is a solution of the problem

$$\max_{\theta \in \Theta^p_{C,M}} E\left[(H - V^c_T(\theta))^+\right]$$

Without transaction costs, but only assuming that X is a semimartingale, this problem has been solved for European options by Cvitanic and Karatzas [6] in a complete market, and by Cvitanic [4] in incomplete and constrained markets. In both cases, they use the duality approach, as opposed to the Neyman-Pearson lemma approach, employed by Föllmer and Leukert [14] to solve the same problem in an unconstrained incomplete market.

Choosing $\nu(x) = (x^-)^p$, with $p > 1$, one obtains a solution for the problem studied by Pham [19] in discrete time.

Example 4.8 (Utility maximization). Let U be a concave increasing function. The *utility maximization* problem

$$\max_{\theta \in \Theta^p_{C,M}} E\left[U\left(V^c_T(\theta) - H\right)\right]$$

admits a solution. In fact, apply Proposition 4.6, with $\nu(x) = -U(x)$. This problem has been studied for European options in a Markovian model by Hodges and Neuberger [15] and developed more rigorously by Davis, Panas, and Zariphopoulou [8]: in both papers, a stochastic control problem is considered, and the assumptions on the model lead to a Hamilton-Jacobi-Bellmann equation which can be solved in a weak sense. In more general models, the same problem has been studied by Cvitanic and Karatzas [5] with the convex duality approach.

Remark 4.9. The variable H needs not be a function of X_T alone: in fact we only require that it is $\mathcal{F}$-measurable. This means that the existence result is valid for a general path-dependent option, as long as its exercise is fixed at time T. This excludes American-type options.

5 Problems with Constraints

In this section we study the problem of hedging with constraints on the space of strategies. Essentially, we consider two types of constraints: those on the position in the risky asset (such as limits on short-selling), and those on the portfolio value (such as margin requirements).

The existence of a constrained minimum depends on two conditions: the stability of the restricted set of admissible strategies under the transformations used on minimizing sequences, and its closedness in the topology where the risk functional is lower semicontinuous.

In this setting, it becomes evident that the more transformations are used in the proof of the unconstrained problem, the smaller is the set of tractable constraints. Since in the case of L^p we only take convex combinations of strategies, it follows that we can obtain an existence result for constraints of the type $\theta_t \in A_t$, where $A_t(\omega)$ is a closed convex subset of $\mathbb{R}$.

On the contrary, in the L^1 case we also use stopped strategies, hence $A_t(\omega)$ will have to be a closed convex containing zero. At any rate, it seems that these conditions are not restrictive for most applications.

Since a convex closed set in $\mathbb{R}$ is indeed a (possibly unbounded) interval, we shall describe a constraint by means of two processes m_t and M_t, taking values in $\mathbb{R} \cup \{-\infty\}$ and $\mathbb{R} \cup \{+\infty\}$ respectively, and representing the lower and upper bounds of the strategy at time t. Of course, both m_t and M_t must be predictable with respect to $\mathcal{F}_t$.

The next proposition establishes the existence of a minimum for a constraint of the type $\theta_t \in A_t$.

Proposition 5.1. *Let m_t and M_t two predictable processes, such that $m_t \leq \theta_t \leq M_t$. Denoting by*

$$\Gamma^p(m, M) = \{\theta_t \in \Theta_C^p : m_t \leq \theta \leq M_t \text{ for } dtP(d\omega)\text{-a.e.}\}$$

if $\Gamma^p(m, M)$ is not empty, we have that:

i) if $p > 1$, then for all $K > 0$ the minimum problem

$$\min_{\theta \in \Theta_{C,K}^p \cap \Gamma^p(m,M)} \rho\left(V_T^c(\theta) - H\right)$$

admits a solution.

ii) if $p = 1$, and $m_t \leq 0 \leq M_t$ $dtP(d\omega)$-a.e., then the minimum problem

$$\min_{\theta \in \Gamma^1(m,M)} \rho\left(V_T^c(\theta) - H\right)$$

admits a solution.

Proof. Reread the proof of Propositions 4.4 and 4.5, observing that convex combinations of strategies in $\Gamma^p(m, M)$ remain in $\Gamma^p(m, M)$ and that (for 4.6) a strategy in $\Gamma^1(m, M)$, if stopped, remains in $\Gamma^1(m, M)$. Finally, if $\theta^n \to \theta$ a.e. in $dtP(d\omega)$, and $\theta^n \in \Gamma^p(m, M)$, then $\theta \in \Gamma^p(m, M)$. $\square$

Example 5.2 (Short-selling). For $m \equiv 0$ and $M \equiv +\infty$, the constraint above becomes $\theta_t \geq 0$ for all t, which amounts to forbid the short sale of X. Notice that in this case the constraint is deterministic.

Example 5.3 (Margin requirements). Despite the generality of Proposition 5.1, there are some relevant constraints of financial markets which look awkward to embed in the framework above. For instance, consider a solvability condition, which in our notation looks like:

$$V_t^c(\theta) - k_t|\theta_t| \geq -l_t$$

where l_t is the maximum credit line available at time t. This constraint is clearly stable with respect to the operations considered, but it does not fit well in Proposition 5.1, where the processes m and M should be written explicitly. Indeed, in this example it can be shown that:

$$(m_t, M_t) = \begin{cases} (-\infty, +\infty) & \text{if } t < \tau \\ (0, 0) & \text{if } t \geq \tau \end{cases}$$

where τ is defined as:

$$\tau = \inf\{t : \; V_t^c(\theta) - k_t|\theta_t| = -l_t\}$$

In other words, the agent is unconstrained until the solvency limit is hit, then the position must be closed for the rest of the period.

In these cases, where the processes m and M are defined implicitly by an inequality, it makes more sense to check directly that the constraint is stable under convex combinations, limits a.e. in $dtP(d\omega)$ and, in the L^1 case, under stopping.

6 The Semimartingale Case

In this section we discuss the problems arising in the more realistic case where X is a semimartingale, and give a partial extension of the foregoing results to this setting.

It is well-known that in a frictionless incomplete market the absence of arbitrage (and thus the well-posedness of a hedging problem) is equivalent to the existence of an equivalent (local) martingale measure (see Delbaen and Schachermayer [11] for a general version of this result). This is no longer true in presence of transaction costs, which may add downside risk to potential arbitrage opportunities, effectively excluding most of them.

In discrete time, Jouini and Kallal [16] show that a market is arbitrage-free if and only if there exists a martingale measure for a process $\tilde{X}$, lying between the bid and the ask prices (respectively $X_t + k_t$ and $X_t - k_t$, in our notation).

In fact, the existence of a martingale measure is a stronger condition than the absence of arbitrage, since it implies that, removing frictions, the same market remains arbitrage-free. Nevertheless, most market models considered in the literature do admit equivalent martingale measures, and here we shall make this additional assumption.

Definition 6.1. We define the following sets of martingale measures:

$$\mathcal{M}_q^e(P) = \{Q \sim P : \frac{dQ}{dP} \in L^q(P), X \text{ is a } Q\text{-local martingale}\}$$

The second issue deals with the space of admissible strategies: if X is a semimartingale, the space:

$$G_T(\Theta^p) = \left\{\theta : \int_0^T \theta_t dX_t \ \in L^p(P)\right\}$$

is generally not closed, unless additional assumptions are made on X. For $p = 2$, Delbaen, Monat, Schachermayer, Schweizer and Stricker [10] established a necessary and sufficient condition for the closedness of $G_T(\Theta^p)$. Unfortunately, this condition fails to hold for some stochastic volatility models (see for instance Biagini, Guasoni and Pratelli [3] for an example), suggesting that the choice of the space $G_T(\Theta^p)$ may not be satisfactory.

In this spirit, Delbaen and Schachermayer [12] have proposed a different space K_p with better closure properties. We briefly summarize a few facts:

Definition 6.2. Let K_p^s be the set of bounded simple integrals with respect to X (as defined in section 2). K_p denotes the closure of K_p^s in the norm topology of $L^p(P)$.

Remark 6.3. If X is a martingale, it is easily seen that $K_p = G_T(\Theta^p)$. This allows to generalize the definition of Θ_C^p to the semimartingale case.

A short version of the main result of Delbaen and Schachermayer [12] sounds as follows:

Theorem 6.4. *Let $1 \leq p \leq \infty$, and $p' = \frac{p}{p-1}$. If X is a continuous semimartingale locally in $L^p(P)$ such that $\mathcal{M}_{p'}^e(P) \neq \emptyset$, and $f \in L^p(P)$, the following conditions are equivalent:*

i) $f \in K_p$*;*

ii) There exists a X-integrable predictable process θ such that $G_t(\theta)$ is a uniformly integrable Q-martingale for each $Q \in \mathcal{M}_{p'}^e(P)$, and $G_t(\theta)$ converges to f in the $L^1(Q)$ norm (as t converges to T);

iii) $E_Q[f] = 0$ *for each* $Q \in \mathcal{M}_{p'}^e(P)$.

Since K_p replaces $G.(\Theta^p)$ when X is a semimartingale, the definition of Θ_C^p can be extended as follows:

$$\Theta_C^p(P) = \{\theta : G.(\theta) \in K_p, C_T(\theta) \in L^p(P)\}$$

We now see when a minimization problem of the type:

$$\min_{\theta \in \Theta^p_C(P)} \rho(V^c_T(\theta) - H)$$

fits into the framework outlined in the previous sections for the martingale case. The idea is to find an auxiliary problem under Q, equivalent to the above problem under P. We have the following:

Proposition 6.5. *Let $1 \leq p \leq \infty$, X a continuous semimartingale locally in $L^p(P)$, and $Q \in \mathcal{M}^e_{p'}(P)$. Then the problem:*

$$\min_{\theta \in \Theta^p_{C,M}(P)} \rho(V^c_T(\theta) - H) \tag{S}$$

admits a solution for any $M > 0$.

Proof. Let $\{Q_i\}_i$ a countable set of martingale measures, such that $\{\frac{dQ_i}{dP}\}_i$ is dense in $\{\frac{dQ}{dP} : Q \in \mathcal{M}^e_{p'}(P)\}$ in the $L^{p'}(P)$ norm.

By the Hölder inequality, the identity map $Id : K_p(P) \mapsto G_T(\Theta^1(Q_i))$ is a continuous operator for all i. Note also that if $C_T(\theta) \in L^p(P)$, then $C_T(\theta) \in L^1(Q_i)$.

Let θ^n be a minimizing sequence for (S). Since the set of simple strategies is dense both in $K_p(P)$ and $\Theta^1_{C,M'}(Q_i)$, it follows that θ^n is a minimizing sequence for the problem:

$$\min_{\theta \in \Theta^1_{C,M'}(Q_i)} \rho(V^c_T(\theta) - H) \tag{M}$$

for a suitable M'. By Proposition 4.5 we can extract a minimizing sequence of convex combinations of θ^n converging to a minimizer θ of (M). Since minimizing sequences are stable under convex combinations, we can take further subsequences of convex combinations converging to θ in $L^1(Q_i)$ for any finite set of i. By a diagonalization argument, we obtain a η^n such that $\eta^n \to \theta$ in $L^1(Q_i)$ for all i.

If $Q \in \mathcal{M}^e_{p'}(P)$, we can assume up to a subsequence that $\frac{dQ_i}{dP} \to \frac{dQ}{dP}$ in $L^{p'}(P)$. By construction, for all i we have that:

$$E_{Q_i}[G_T(\theta)] = 0$$

and as $i \to \infty$, we obtain that $E_Q[G_T(\theta)] = 0$ for any $Q \in \mathcal{M}^e_{p'}(P)$. By Theorem 6.4, it follows that $\theta \in K_p$. Also, Lemma 3.4 implies that $C_T(\theta) \in L^p(P)$. Therefore $\theta \in \Theta^p_C(P)$and the proof is complete. □

Remark 6.6. The diagonalization procedure in Proposition 6.5 is necessary since the map $T : K_p(P) \mapsto G_T(\Theta^1(Q))$ is generally not onto (see Delbaen and Schachermayer [12], Remark 2.2 c) for details).

Appendix

We recall here a few results in functional analysis and stochastic integration that we use in the main text. Unlike the previous sections, here T denotes a generic stopping time.

This result dates back to Banach, and is well-known:

Theorem 6.7. *Let x_n be a relatively weakly compact sequence in a Banach space V. Then there exists a sequence of convex combinations $y_n = \sum_{i=n}^{\infty} \alpha_i^n x_n$ such that y_n converges in the norm topology of V.*

Bounded sets are relatively compact in L^p spaces for $p > 1$. For L^1 this is not the case, since its weak star closure leads to the space of Radon measures. Nonetheless, relative compactness can be recovered through convex combinations, as shown by the following:

Theorem 6.8 (Komlós). *Let X_n be a sequence of random variables, such that $\sup_n E|X_n| < \infty$. Then there exists a subsequence X'_n and a random variable $X \in L^1$ such that, for each subsequence X''_n of X'_n,*

$$\frac{1}{n}\sum_{i=1}^{n} X''_n \to X \ a.e.$$

The next type of results states when a sequence of stochastic integrals converges to a stochastic integral. Again, the situation is different for $p > 1$ and $p = 1$. The first case is a classic result of stochastic integration, and dates back to Kunita-Watanabe for $p = 2$:

Theorem 6.9 (Kunita-Watanabe). *Let X be a continuous local martingale, θ^n a sequence of X-integrable predictable stochastic processes such that each $\int_0^t \theta_s^n dX_s$ is a L^p-bounded martingale, and such that the sequence $\int_0^\infty \theta_s^n dX_s$ converges to a random variable G in the norm topology of L^p.*

Then there is an $\mathcal{F}^X$-predictable stochastic process θ such that $\int_0^t \theta_s dX_s$ is an L^p-bounded martingale, and such that $\int_0^t \theta_s dX_s = G$.

The case $p = 1$ is due to Yor [21]:

Theorem 6.10 (Yor). *Let X be a continuous local martingale, θ^n a sequence of X-integrable predictable stochastic processes such that each $\int_0^t \theta_s^n dX_s$ is a uniformly integrable martingale, and such that the sequence $\int_0^\infty \theta_s^n dX_s$ converges to a random variable G in the norm topology of L^1 (or even in the $\sigma(L^1, L^\infty)$ topology).*

Then there is an $\mathcal{F}^X$-predictable stochastic process θ such that $\int_0^t \theta_s dX_s$ is a uniformly integrable martingale, and such that $\int_0^t \theta_s dX_s = G$.

The main difference between $p > 1$ and $p = 1$ is that in the latter case the norms $L^p : M \mapsto E\left[M_\infty^p\right]$ and $\mathcal{H}^p : M \mapsto \langle M \rangle_\infty^{\frac{p}{2}}$ are no longer equivalent. Yor's idea is to reduce the L^1 case to $\mathcal{H}^1$ by stopping arguments.

In fact, Theorem 6.10 is a consequence of a more general result (see [21], Theorem 2.4 page 277), which gives further information on converging sequences of uniformly integrable martingales. We report this result with its proof, and an easy corollary used in this paper.

Note that a different proof of 6.10 can be found in Delbaen and Schachermayer [13], with an excellent exposition of the properties of the space $\mathcal{H}^1$.

Theorem 6.11 (Yor). *Let A be a set of uniformly integrable martingales, and denote*

$$\Phi(A) = \{Y^T : \ Y \in A, \ Y^T \in \mathcal{H}^1\}$$

Let Y be a uniformly integrable martingale, $Y^n \in A$ and $Y^n \to Y$ in L^1 (or even weakly in $\sigma(L^1, L^\infty)$). Then, for all stopping times T such that $Y^T \in \mathcal{H}^1$, we have that Y^T belongs to the closure of $\Phi(A)$ in the weak topology $\sigma(H^1, BMO)$.

Proof. We separate the proof into three steps.

Step 1: we first show that if the theorem holds for $Y \in \mathcal{H}^1$ and $T = \infty$, then it holds in general. If $Y^n \rightharpoonup Y$ in $\sigma(L^1, L^\infty)$, then $(Y^n)^T \rightharpoonup Y^T$, and we can apply the theorem to $(Y^n)^T, Y^T$ (since $Y^T \in \mathcal{H}^1$ by assumption) obtaining that Y^T belongs to the closure of $\Phi(\{(Y^n)^T\}_n)$, which is smaller than $\Phi(A)$.

Step 2: we further reduce the proof to the case of $Y^n \in \mathcal{H}^1$. For all n, there exists a sequence of stopping times $S_k^n \to \infty$ such that $(Y^n)^{S_k^n} \in \mathcal{H}^1$. Denoting by $Z_n = (Y^n)^{S_k^n}$, by martingale convergence there exists some $k = k_n$ such that $\|Y_\infty^n - Z_\infty^n\|_{L^1} = \|Y_\infty^n - E\left[Y_\infty^n | \mathcal{F}_{S_k^n}\right]\|_{L^1} \le \frac{1}{n}$. Therefore, $Z^n \in \mathcal{H}^1$ for all n, and we have, for all $g \in L^\infty$:

$$|E\left[(Z_\infty^n - Y_\infty)g\right]| \le \|Z_\infty^n - Y_\infty^n\|_{L^1} \|g\|_{L^\infty} + |E\left[(Y_\infty^n - Y_\infty)g\right]|$$

therefore Z_∞^n converges weakly in $\sigma(L^1, L^\infty)$ to Y_∞. Also, the set $\Phi(\{Z^n\}_n)$ is smaller then $\Phi(A)$.

Step 3: we now give the proof under the assumptions $Y^n, Y \in \mathcal{H}^1, T = \infty$. It is sufficient to show, for any finite subset $(U^1, \dots, U^d) \in BMO$, that there exists some stopping time T such that:

$$|E\left[[(Y^n)^T - Y, U^i]_\infty\right]| < \varepsilon \quad \text{for all } i$$

We take $T = \inf\{t : \sum_{i=1}^d |U^i| \ge k\}$, choosing k such that:

$$E\left[\int_{(T,\infty]} |d[Y, U^i]_s|\right] < \frac{\varepsilon}{2} \quad \text{for all } i$$

which is always possible by the Fefferman inequality, since $[Y, U^i]$ has integrable variation, $Y \in \mathcal{H}^1$, and $U^i \in BMO$. Therefore it remains to show that, for some fixed T, and for all $i = 1, \dots, d$:

$$\lim_{n\to\infty} E\left[[(Y^n)^T, U^i]_\infty\right] = \lim_{n\to\infty} E\left[[Y^n, (U^i)^T]_\infty\right] = E\left[[Y, (U^i)^T]_\infty\right]$$

U^i is bounded in $[0, T)$, but it belongs to BMO, it has bounded jumps, hence it is also bounded on $[0, T]$. As a result, $(U^i)^T$ is bounded. The local martingale $[Y^n, (U^i)^T] - Y^n(U^i)^T$ hence belongs to the class D, and we have:

$$E\left[[Y^n, (U^i)^T]_\infty\right] = E\left[Y^n_\infty U^i_T\right] \quad \text{and} \quad E\left[[Y, (U^i)^T]_\infty\right] = E\left[Y_\infty U^i_T\right]$$

Finally, $U^i_T \in L^\infty$, and the thesis follows from the assumption $Y^n \rightharpoonup Y$ in $\sigma(L^1, L^\infty)$. ☐

Remark 6.12. The statement of Theorem 6.11 with $A = \{Y^n\}_n$ says that, for any $Y^\tau \in \mathcal{H}^1$, there exists a sequence of indices n_k, and a sequence of stopping times τ_k such that $Y^{\tau_k}_{n_k} \rightharpoonup Y^\tau$ in $\sigma(H^1, BMO)$. However, *a priori* the sequence n_k may not tend to infinity, and the the stopping times τ_k may not converge almost surely to τ.

The proof provides more information on these issues. For the first, note that in fact $n_k = k$. For the latter, let us look more closely to the three steps.

Step 1 simply shows that τ_k may be chosen such that $\tau_k \leq \tau$ a.e.

In Step 2, we have $S^n_k \to \infty$, and k_n must be sufficiently high. Therefore we can replace it with some higher k_n such that the condition $P(S^n_{k_n} < n) < \frac{1}{n}$ is satisfied as well.

Likewise, in Step 3 the stopping time T needs to be sufficiently high, hence it may be chosen to satisfy the condition $P(T < n) < \frac{1}{n}$. By a diagonalization argument, we can select a sequence $(Y^n)^{T_n}$ such that $(Y^n)^{T_n} \rightharpoonup Y$ in $\sigma(H^1, BMO)$, and $T_n \to \infty$ a.e.

The previous Remark provides the following:

Corollary 6.13. *If $Y^n \rightharpoonup Y$ in $\sigma(L^1, L^\infty)$, and $Y^\tau \in \mathcal{H}^1$ for some stopping time τ, then there exists a subsequence n_k and a sequence of stopping times $\tau_k \to \tau$ a.e. such that $(Y^{n_k})^{\tau_k} \rightharpoonup Y^\tau$ in $\sigma(\mathcal{H}^1, BMO)$.*

References

[1] L. Ambrosio, N. Fusco, and D. Pallara, *Special functions of bounded variation and free discontinuity problems*, Oxford University Press, 1999.

[2] P. Artzner, F. Delbaen, J.M. Eber, and D. Heath, *Coherent measures of risk*, Mathematical Finance **9** (1999), no. 3, 203–228.

[3] F. Biagini, P. Guasoni, and M. Pratelli, *Mean-variance hedging for stochastic volatility models*, Preprint, April 1999.

[4] J. Cvitanic, *Minimizing expected loss of hedging in incomplete and constrained markets*, SIAM Journal of Control and Optimization (1999), To appear.

[5] J. Cvitanic and I. Karatzas, *Hedging and portfolio optimization under transaction costs: a martingale approach*, Mathematical Finance **6** (1996), 133–165.

[6] ———, *On dynamic measures of risk*, Finance and Stochastics (1999).

[7] M.H.A. Davis and A.R. Norman, *Portfolio selection with transaction costs*, Mathematics for Operations Research **15** (1990), 676–713.

[8] M.H.A. Davis, V. Panas, and T. Zariphopoulou, *European option pricing with transaction costs*, SIAM Journal of Control and Optimization **31** (1993), no. 2, 470–493.

[9] F. Delbaen, *Coherent risk measures on general probability spaces*, Preprint, November 1998.

[10] F. Delbaen, P. Monat, Schachermayer W., M. Schweizer, and C. Stricker, *Weighted norm inequalities and hedging in incomplete markets*, Finance and Stochastics **1** (1997), no. 3, 181–227.

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

[12] ———, *Attainable claims with p-th moments*, Annales Institute H. Poincaré de Probabilites and Statistiques **32** (1996), no. 6, 743–763.

[13] ———, *A compactness principle for bounded sequences of martingales with applications*, Preprint, March 1998.

[14] H. Föllmer and P. Leukert, *Efficient hedging: Cost versus shortfall risk*, Finance and Stochastics **4** (1999), no. 2.

[15] Hodges and Neuberger, *Optimal replication of contingent claims under transaction costs*, Review of Futures Markets **8** (1989), 222–239.

[16] E. Jouini and H. Kallal, *Martingales and arbitrage in securities markets with transaction costs*, Journal of Economic Theory **66** (1995), no. 1, 178–197.

[17] J. Komlós, *A generalization of a problem of steinhaus*, Acta Mathematica Academiae Scientiarum Hungaricae **18** (1967), no. 1-2, 217–229.

[18] H.E. Leland, *Option pricing and replication with transaction costs*, Journal of Finance **40** (1985), 1283–1301.

[19] H. Pham, *Dynamic l^p hedging in discrete time under cone constraints*, Preprint, 1998.

[20] M. Schwartz, *New proofs of a theorem of komlós*, Acta Mathematica Hungarica **47** (1986), no. 1-2, 181–185.

[21] M. Yor, *Sous-espaces denses dans l^1 ou h^1 et reprèsentation des martingales*, Séminaire de Probabilité XII, Lecture Notes in Mathematics, vol. 649, Springer-Verlag, 1978, pp. 265–309.