Hedging, Arbitrage, and Optimality with Superlinear Frictions

ARTICLE *in* SSRN ELECTRONIC JOURNAL · AUGUST 2013

DOI: 10.2139/ssrn.2317344

Paolo Guasoni
Dublin City University

Hedging, arbitrage, and optimality with superlinear frictions

Paolo Guasoni[*] Miklós Rásonyi[†]

August 28, 2013

Abstract

In a continuous-time model with multiple assets described by cadlag processes, this paper characterizes superhedging prices, absence of arbitrage, and utility maximizing strategies, under general frictions that make execution prices arbitrarily unfavorable for high trading intensity. With such frictions, dual elements correspond to a pair of a shadow execution price combined with an equivalent martingale measure. For utility functions defined on the real line, optimal strategies exist even if arbitrage is present, because it is not scalable at will.

MSC (2010): 91G10, 91G80.

Keywords: hedging, arbitrage, price-impact, frictions, utility maximization.

[*]Boston University, Department of Mathematics and Statistics, 111 Cummington Street, Boston, MA 02215, USA, and Dublin City University, School of Mathematical Sciences, Glasnevin, Dublin 9, Ireland, email guasoni@bu.edu. Partially supported by the ERC (278295), NSF (DMS-1109047), SFI (07/SK/M1189, 08/SRC/FMC1389), and FP7 (RG-248896).

[†]University of Edinburgh and Rényi Institute, Budapest

1 Introduction

In financial markets, trading moves prices against the trader: buying more, and faster, increases execution prices, while selling does the opposite. This aspect of liquidity, known as market depth [5] or price-impact, is widely documented empirically [19, 11], and has received increasing attention in models of asymmetric information [28], illiquid portfolio choice [35, 20], and optimal liquidation [1, 4, 36]. These models depart from the literature on frictionless markets, where prices are the same for any amount traded. They also depart from transaction costs models, in which prices differ for buying and selling, but are insensitive to quantities.

The growing interest in price-impact has also highlighted a shortage of effective theoretical tools. In these models, what is the analogue of a martingale measure? Which contingent claims are hedgeable, and at what price? What it the optimality condition for utility maximization? In discrete time, several researchers have studied these fundamental questions [2, 32, 31, 18], but extensions to continuous time have proved challenging. This paper aims at filling the gap.

Tackling price-impact in continuous-time requires to clarify two basic concepts, which remain concealed in discrete models: the relevant classes of trading strategies and of pricing functionals. First, to retain price-impact effects in continuous time, execution prices must depend on quantities per unit of time, i.e. trading intensity, rather than on quantities themselves, otherwise price-impact can be avoided with judicious policies [8, 10, 9]. Various classes of trading strategies have appeared in different models ([10] [36]), but a general definition of feasible strategy has not yet emerged. The second key concept is the relevant notion of pricing functional – the analogue of a martingale measure. In the transaction costs literature, such a pricing functional are identified as a consistent prices system, a pair $(\tilde{S}, Q)$ of a price $\tilde{S}$ evolving within the bid-ask spread, and a probability Q under which $\tilde{S}$ is a martingale. Such a definition suggests that with frictions, passing to the risk-neutral setting requires both a change in the probability measure, and a change in the price process.

Superlinear frictions, such as price-impact models, entail that execution prices become arbitrarily unfavorable as traded quantities grow per unit of time: buying too fast is impossibly expensive, and selling intolerably punitive. As a result, trading is feasible only at finite rates – the number of shares is absolutely continuous. This feature sets apart superlinear frictions from frictionless markets, in which the number of shares are predictable processes, and from transaction cost models, in which they have finite variation.

Finite trading rates have two central implications: first, portfolio values are well-defined for asset prices that follow general cadlag processes, not only for semimartingales. Second, immediate portfolio liquidation is impossible, and therefore the usual notion of admissibility, based on a lower bound for liquidation values, is inappropriate with superlinear frictions. We define a *feasible* strategy as any trading policy with finite trading rate and trading volume, without any lower bounds on portfolio values. In frictionless markets, or with transaction costs, this approach would fail for two reasons: first, such a class would not be closed in any reasonable sense, as a block trade is approximated by intense trading over small time intervals. Second, portfolios unbounded from below allow doubling strategies, which lead to arbitrage even with martingale prices.

Neither issue arises with superlinear frictions. Block trades are infeasible, even in the limit, as intense trading incurs exorbitant costs: put differently, bounded losses imply bounded *trading volume* (Lemma 3.4). The bound on trading volume in turn yields the closedness of the payoffs of feasible strategies (Proposition 3.5), and the martingale property of portfolio values under shadow execution prices, which excludes arbitrage through doubling strategies (Lemma 5.4).

Arbitrage is also different from frictionless or transaction cost models. Unlike these settings, in which an arbitrage opportunity scales freely, superlinear frictions imply that scaling trading rates results in a less than proportional scaling of payoffs. In fact, we prove a stronger result, whereby *all* payoffs are dominated by a single random variable, the *market bound*, which depends on the friction and on the asset price only (Lemma 3.5). This bound implies, in particular, that price-impact defeats arbitrage, if pursued on a large scale.

All these definitions and properties come together – and are sustained by – the main superhedging result, Theorem 3.7, which characterizes the initial asset positions that can dominate a given claim through trading, in terms of shadow execution prices. The main message of this theorem is that the superhedging price of a claim is the supremum of its expected value under a martingale measure for an execution price, *minus* a penalty, which reflects how far the shadow price is from the base price. The penalty depends on the *dual friction*, introduced by [18] in discrete time, and is zero for any equivalent martingale measure of the asset price. Importantly, the theorem is valid even if there are no martingale measures, or if the price is not a semimartingale.

The superhedging theorem, which does not assume absence of arbitrage, characterizes a large class of models that do not admit arbitrage of the second kind (strategies that lead to a sure minimal gain) even in limited amounts. As for transaction costs, this class contains any price process that satisfies the conditional full support property [22], including fractional Brownian motion.

We conclude the paper addressing utility maximization with superlinear frictions. First, a general theorem guarantees that optimal solutions exist – even with arbitrage, which must be chosen optimally, lest price-impact offset gains. Second, optimal strategies are identified by a version of the familiar first-order condition that the marginal utility of the optimal payoff be proportional to the stochastic discount factor. Some technicalities aside, price-impact leads to a novel condition, which prescribes that the stochastic discount factor makes the shadow execution price, not the base price, a martingale. In models of transaction costs this criterion formally reduces to the usual shadow price approach for optimality [26].

The rest of the paper proceeds with section 2, which describes the model in detail. The main theoretical tools are developed in section 3, which proves the market bound, the trading volume bound, the closedness of the payoff space, and the main superhedging result. Section 4 discusses the implications for arbitrage of the second kind, and its absence with prices with conditional full support. Section 5 concludes with the results on utility maximization.

2 The Model

Let $T > 0$ denote a finite time horizon, and consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ with $\mathcal{F}_0$ trivial, satisfying the usual hypotheses as well as $\mathcal{F} = \mathcal{F}_T$. $\mathcal{O}$ denotes the optional sigma-field on $\Omega \times [0,T]$. The market includes a safe asset S^0, used as numeraire, hence $S^0_t \equiv 1$, $t \in [0,T]$, and d risky assets, described by càdlàg, adapted processes $(S^i_t)^{1\le i\le d}_{t\in[0,T]}$. Henceforth S denotes the d-dimensional process with components S^i, $1 \le i \le d$, the concatenation xy of two vectors x, y of equal dimensions denotes their scalar product, and $|x|$ denotes the Euclidean norm of x. For a $(d+1)$-dimensional vector x, its coordinates are denoted by $x^0, \dots, x^d$.

The next definition identifies those strategies for which the number of shares changes over time at some finite rate ϕ, hence the number of shares is a.s. differentiable.

Definition 2.1. *A* feasible strategy *is a process ϕ in the class*

$$\mathcal{A} := \left\{ \phi : \phi \text{ is a } \mathbb{R}^d\text{-valued, optional process}, \int_0^T |\phi_u| du < \infty \text{ a.s.} \right\}. \tag{1}$$

In this definition, the process ϕ represents the *trading rate*, that is, the speed at which the number of shares in each asset changes over time, and the condition $\int_0^T |\phi_u| du < \infty$ means that *absolute turnover* (the cumulative number of shares bought or sold) remains finite in finite time.

The above definition compares to that of admissible strategies in frictionless markets as follows. On one hand, it relaxes the solvency constraint typical of admissibility, since a feasible strategy can lead to negative wealth. On the other hand, this definition restricts the number of shares to be differentiable in time, while usual admissible strategies an have an arbitrarily irregular number of shares.[1]

With this notation, in the absence of frictions the self-financing condition would imply a position at time T in the safe asset (henceforth, cash) equal to[2]:

$$z^0 - \int_0^T S_t \phi_t dt, \tag{2}$$

where z^0 represents the initial capital, and the integral reflects the cost of purchases and the proceeds of sales. For a given trading strategy ϕ, frictions reduce the cash position, by making purchases more expensive, and sales less profitable. With a similar notation to [18], we model this effect by introducing a function G, which summarizes the impact of frictions on the execution price at different trading rates:

Assumption 2.2 (Friction)**.** *Let $G : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}_+$ be a $\mathcal{O} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable function, such that $G(\omega, t, \cdot)$ is convex with $G(\omega, t, x) \geq G(\omega, t, 0)$ for all ω, t, x. Henceforth, set $G_t(x) := G(\omega, t, x)$, i.e. the dependence on ω is dropped, and t is used as a subscript.*

With this definition, for a given strategy $\phi \in \mathcal{A}$ and an initial asset position $z \in \mathbb{R}^{d+1}$, the resulting positions at time $t \in [0,T]$ in the risky and safe assets are defined as:

$$V_t^i(z, \phi) := z^i + \int_0^t \phi_u^i du \quad 1 \leq i \leq d, \tag{3}$$

$$V_t^0(z, \phi) := z^0 - \int_0^t \phi_u S_u du - \int_0^t G_u(\phi_u) du. \tag{4}$$

The first equation merely says that the cumulative number of shares V_t^i in the i-th asset is given by the initial number of shares, plus subsequent flows. The second equation contains the new term involving the friction G, which summarizes the impact of trading on execution prices. The convexity of $x \mapsto G_t(x)$ implies that trading twice as fast for half the time may only increase execution costs – speed is expensive. The condition $G(\omega, t, x) \geq G(\omega, t, 0)$ means that

[1] In the definition of feasible strategy an optional trading rate leads to a continuous, hence predictable, number of shares, as for usual admissible strategies.

[2] By the càdlàg property of S_t, the function $S_t(\omega), t \in [0,T]$ is bounded for almost every $\omega \in \Omega$, hence the integral in (2) is finite a.s. for each ϕ satisfying $\int_0^T |\phi_t| dt < \infty$ a.s.

inactivity is always cheaper than any trading activity. Indeed, for most models in the literature $G(\omega, t, 0) = 0$, but the above definition allows for $G(\omega, t, 0) > 0$, which is interpreted as a cost of participation in the market, such as the fees charged by exchanges to trading firms. Finally, note that in general V_t^0 may take the value $-\infty$ for some (unwise) strategies.

With a single risky asset, the above specification is equivalent to assuming that a trading rate of ϕ_t implies an execution price equal to

$$\tilde{S}_t = S_t + G_t(\phi_t)/\phi_t \tag{5}$$

which is (since G is positive) higher than S_t when buying, and lower when selling. Thus, $G \equiv 0$ boils down to a frictionless market, while proportional transaction costs correspond to $G_t(x) = \varepsilon S_t |x|$ with some $\varepsilon > 0$. Yet, this paper focuses on neither of these settings, which entail either zero or linear costs, but rather on superlinear frictions, defined as those that satisfy the following conditions.

Assumption 2.3 (Superlinearity). *There is $\alpha > 1$ and an optional process H such that*[3]

$$\inf_{t\in[0,T]} H_t > 0 \quad a.s., \tag{6}$$

$$G_t(x) \geq H_t |x|^\alpha, \quad \text{for all } \omega, t, x, \tag{7}$$

$$\int_0^T \sup_{|x|\leq N} G_t(x) dt < \infty \quad a.s. \text{ for all } N > 0, \tag{8}$$

$$\sup_{t\in[0,T]} G_t(0) \leq K \text{ for some constant } K. \tag{9}$$

Condition (7) is the central assumption of superlinearity, and prescribes that trading twice as fast for half the time increases costs by a minimum positive proportion. Condition (6) requires that frictions never disappear, and (8) that they remain finite in finite time. By (9), the participation cost must be uniformly bounded in $\omega \in \Omega$. In summary, these conditions characterize nontrivial, finite, superlinear frictions.

The most common examples in the literature are, with one risky asset, the friction $G_t(x) := \Lambda |x|^\alpha$ for some $\Lambda > 0$, $\alpha > 1$ and, in multiasset models, the friction $G_t(x) := x'\Lambda x$ for some symmetric, positive-definite, $d \times d$ square matrix Λ.

Remark 2.4. Our results remain valid assuming that (7) holds for $|x| \geq M$ only, with some $M > 0$. Such an extension requires only minor modifications of the proofs, and may accommodate models for which a low trading rate incurs either zero or linear costs.

3 Superhedging and Dual Characterization of Payoffs

Despite their similarity to models of frictionless markets and transaction costs, superlinear frictions lead to a surprisingly different structure of attainable payoffs, as shown in this section. First, note that the class of feasible strategies considered above, while still well-defined even in a model without frictions or with proportional transaction costs, is virtually useless in these settings, where such a class is not closed in any reasonable sense. Indeed, optimal policies in

[3] We implicitly assume that $\inf_{t\in[0,T]} H_t$ is a random variable, which is always the case if H is progressively measurable, in addition to optional.

such models are not smooth, as the number of shares follows a diffusion in a typical frictionless model, and a nondifferentiable function of finite variation in a transaction costs setting.

With superlinear frictions, feasible strategies are closed in a strong sense. The intuitive reason is that approximating a nonsmooth strategy would require trading at increasingly high speed, and hence infinite turnover.

3.1 The Market Bound

Superlinear frictions lead to a striking boundedness property: for a fixed initial position, all payoffs of feasible strategies are bounded above by a single random variable $B < \infty$, the *market bound*, which depends on the friction G and on the price S, but *not* on the strategy. This property clearly fails in frictionless markets, where any payoff with zero initial capital can be scaled arbitrarily, and therefore admits no uniform bound. In such markets, a much weaker boundedness property holds: Corollary 9.3. of [12] shows that the set of payoff of x-admissible strategies is bounded in L^0 *if* the market is arbitrage-free, in the sense that the condition (NFLVR) holds, and a similar result holds with transaction costs under the (RNFLVR) property [21].

A central tool in this analysis is the function G^*, the Fenchel-Legendre conjugate of G, which we call *dual friction.* Its importance was first recognized by [18], who used it to derive a superhedging result in discrete time. G^* is defined as [4]

$$G_t^*(y) := \sup_{x \in \mathbb{R}^d} (xy - G_t(x)), \ y \in \mathbb{R}^d, \ t \in [0,T] \ , \tag{10}$$

and the typical case $d = 1$, $G_t(x) = \Lambda |x|^\alpha$ leads to $G_t^*(y) = \frac{\alpha-1}{\alpha} \alpha^{\frac{1}{1-\alpha}} \Lambda^{\frac{1}{1-\alpha}} y^{\frac{\alpha}{\alpha-1}}$ (in particular, $G_t^*(y) = y^2/(4\Lambda)$ for $\alpha = 2$). With this notation, observe that:

Lemma 3.1. *Under Assumption 2.3, the random variable $B := \int_0^T G_t^*(-S_t)dt$ is finite almost surely.*

Proof. Consider first the case $d = 1$. Then, by direct calculation,

$$G_t^*(y) \le \sup_{r \in \mathbb{R}} \{ry - H_t |r|^\alpha\} = \frac{\alpha - 1}{\alpha} \alpha^{\frac{1}{1-\alpha}} H_t^{\frac{1}{1-\alpha}} y^{\frac{\alpha}{\alpha-1}} \tag{11}$$

Noting that $\sup_{t \in [0,T]} |S_t|$ is finite a.s. by the càdlàg property of S, and knowing that $\inf_{t \in [0,T]} H_t$ is a positive random variable, it follows that

$$\sup_{t \in [0,T]} G_t^*(-S_t) < \infty \text{ a.s.},$$

which clearly implies the statement. If $d > 1$, then note that

$$G_t^*(y) \le \sup_{r \in \mathbb{R}^d} \left(\sum_{i=1}^d r^i y^i - H_t |r|^\alpha \right) \le \sum_{i=1}^d \sup_{r \in \mathbb{R}^d} \left(r^i y^i - (H_t/d)|r|^\alpha \right) \le \sum_{i=1}^d \sup_{x \in \mathbb{R}} \left(x y^i - (H_t/d)|x|^\alpha \right) \tag{12}$$

and the conclusion follows from the scalar case. □

[4] Note that the supremum can be taken over $\mathbb{Q}^d$, hence G^* is $\mathcal{O} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable. Note also that, under Assumption 2.3, $G_t^*(\cdot)$ is a finite, convex function satisfying $G_t^*(x) \ge -K$ for all x, see the proof of Lemma 3.1.

The key observation is that:

Lemma 3.2. *Under Assumption 2.3, any $\phi \in \mathcal{A}$ satisfies*

$$V_T^0(z, \phi) \leq z^0 + B \ a.s.$$

Proof. Indeed, this follows from 4, the definition of G_t^*, and Lemma 3.1. □

Since $B < \infty$ a.s, it is impossible to achieve a scalable arbitrage, that is an arbitrarily large profit from zero initial capital with positive probability. Even if an arbitrage exists, amplifying it too much backfires, because the superlinear friction eventually overrides profits. Yet, arbitrage opportunities can exist in limited size (cf. section 4 below).

3.2 Trading Volume Bound

For $Q \sim P$, denote by $L^1(Q)$ the Banach space of $(d+1)$-dimensional, Q-integrable random variables; $L^0(A)$ denotes the set of A-valued random variables for some subset A of a Euclidean space, equipped with the topology of convergence in probability. $E_Q X$ denotes the expectation of a random variable X under Q. From now on, fix $1 < \beta < \alpha$, where α is as in Assumption 2.3. Let γ be the conjugate number of β, defined by

$$\frac{1}{\beta} + \frac{1}{\gamma} = 1.$$

The next definition identifies a class of reference probability measures with integrability properties that fit the friction G. As it will be clear later, this class identifies those probabilities under which some shadow execution price may have the martingale property.

Definition 3.3. *$\mathcal{P}$ denotes the set of probabilities $Q \sim P$, such that*

$$E_Q \int_0^T H_t^{\beta/(\beta-\alpha)} (1 + |S_t|)^{\beta\alpha/(\alpha-\beta)} dt < \infty.$$

$\tilde{\mathcal{P}}$ denotes the set of probabilities $Q \in \mathcal{P}$ such that

$$E_Q \int_0^T |S_t| dt < \infty \quad \text{and} \quad E_Q \int_0^T \sup_{|x| \leq N} G_t(x) dt < \infty \text{ for all } n \geq 1$$

For a (possibly multidimensional) random variable W, define

$$\mathcal{P}(W) := \{Q \in \mathcal{P} : E_Q|W| < \infty\}, \quad \tilde{\mathcal{P}}(W) := \{Q \in \tilde{\mathcal{P}} : E_Q|W| < \infty\}.$$

Under Assumption 2.3, note that [14, page 266], $\tilde{\mathcal{P}}(W) \neq \emptyset$ for all W.

The next lemma shows that, if a payoff has a finite negative part under some of these probabilities, then its trading rate must also be integrable. There is no analogue to such a result in frictionless markets, but transaction costs [21, Lemma 5.5] lead to a similar property, whereby any admissible strategy must satisfy an upper bound on its total variation. In both cases, the intuition is that, with frictions, excessive trading causes unbounded losses. Hence, a bound on losses translates into one for trading volume.

In the sequel, x_- denotes the negative part of $x \in \mathbb{R}$.

Lemma 3.4. *Let $Q \in \mathcal{P}$ and $\phi \in \mathcal{A}$ be such that $E_Q \xi_- < \infty$, where*

$$\xi := -\int_0^T S_t \phi_t dt - \int_0^T G_t(\phi_t) dt.$$

Then

$$E_Q \int_0^T |\phi_t|^\beta (1 + |S_t|)^\beta dt < \infty.$$

Proof. For ease of notation, set $T := 1$. Define $\phi_t(n) := \phi_t 1_{\{|\phi_t| \le n\}} \in \mathcal{A}$, $n \in \mathbb{N}$. As $n \to \infty$, clearly $\phi_t(n) \to \phi_t$ for all t and $\omega \in \Omega$, and the random variables

$$\xi_n := -\int_0^1 S_t \phi_t(n) dt - \int_0^1 G_t(\phi_t(n)) dt = \tag{13}$$

$$-\sum_{i=1}^d \int_0^1 S_t^i \phi_t^i(n) [1_{\{S_t^i \le 0, \phi_t^i \le 0\}} + 1_{\{S_t^i > 0, \phi_t^i \le 0\}} + 1_{\{S_t^i \le 0, \phi_t^i > 0\}} + 1_{\{S_t^i > 0, \phi_t^i > 0\}}] dt \tag{14}$$

$$-\int_0^1 G_t(\phi_t(n)) dt \tag{15}$$

converge to ξ a.s. by monotone convergence. (Note that each of the terms with an indicator converges monotonically, and that $G_t(0) \le G_t(x)$ for all x.) Hölder's inequality yields

$$\int_0^1 |\phi_t(n)|^\beta (1 + |S_t|)^\beta dt = \int_0^1 |\phi_t(n)|^\beta H_t^{\beta/\alpha} \frac{1}{H_t^{\beta/\alpha}} (1 + |S_t|)^\beta dt \le \tag{16}$$

$$\left[\int_0^1 |\phi_t(n)|^\alpha H_t dt\right]^{\beta/\alpha} \left[\int_0^1 \left(\frac{1}{H_t^{\beta/\alpha}} (1 + |S_t|)^\beta\right)^{\alpha/(\alpha-\beta)} dt\right]^{(\alpha-\beta)/\alpha} \le$$

$$\left[\int_0^1 G_t(\phi_t(n)) dt\right]^{\beta/\alpha} \left[\int_0^1 \left(\frac{1}{H_t^{\beta/\alpha}} (1 + |S_t|)^\beta\right)^{\alpha/(\alpha-\beta)} dt\right]^{(\alpha-\beta)/\alpha}.$$

All these integrals are finite by Assumption 2.3 and the càdlàg property of S. Now, set

$$m := \left[\int_0^1 \left(\frac{1}{H_t^{\beta/\alpha}} (1 + |S_t|)^\beta\right)^{\alpha/(\alpha-\beta)} dt\right]^{(\alpha-\beta)/\alpha},$$

and note that, by Jensen's inequality,

$$\int_0^1 |\phi_t(n)| (1 + |S_t|) dt \le [\int_0^1 |\phi_t(n)|^\beta (1 + |S_t|)^\beta dt]^{1/\beta}. \tag{17}$$

Note also that if $x \ge 1$ and $x \ge 2^{\frac{\beta}{\alpha-\beta}} m^{\frac{\alpha}{\alpha-\beta}}$ then $x^{1/\beta} - (x/m)^{\alpha/\beta} \le x - 2x = -x$. This observation, applied to

$$x := \int_0^1 |\phi_t(n)|^\beta (1 + |S_t|)^\beta dt$$

implies that $\xi_n \leq -x$ on the event $\{x \geq 2^{\frac{\beta}{\alpha-\beta}} m^{\frac{\alpha}{(\alpha-\beta)}} + 1\}$. Thus,

$$\int_0^1 |\phi_t(n)|^\beta (1 + |S_t|)^\beta dt \leq (\xi_n)_- + 2^{\frac{\beta}{\alpha-\beta}} m^{\frac{\alpha}{(\alpha-\beta)}} + 1, \text{ a.s.}$$

Letting n tend to ∞, it follows that

$$\int_0^1 |\phi_t|^\beta (1 + |S_t|)^\beta dt \leq \xi_- + 2^{\frac{\beta}{\alpha-\beta}} m^{\frac{\alpha}{(\alpha-\beta)}} + 1, \tag{18}$$

which implies the the claim, since $E_Q \xi_- < \infty$ by assumption, and $E_Q m^{\frac{\alpha}{\alpha-\beta}} < \infty$ from $Q \in \mathcal{P}$. □

3.3 Closed Payoff Space

The central implication of the previous result is that the class of multivariate payoffs superhedged by a feasible strategy, defined as $C := [\{V_T(0, \phi) : \phi \in \mathcal{A}\} - L^0(\mathbb{R}_+^{d+1})] \cap L^0(\mathbb{R}^{d+1})$, is closed in a rather strong sense. This is the key property, which confirms that this class of strategies is suitable for a superhedging result.

Proposition 3.5. *Under Assumption 2.3, the set $C \cap L^1(Q)$ is closed in $L^1(Q)$ for all $Q \in \mathcal{P}$ such that $\int_0^T |S_t| dt$ is Q-integrable.*

Proof. Take $T = 1$ for simplicity, and assume that $\rho_n := \xi_n - \eta_n \to \rho$ in $L^1(Q)$ where $\eta_n \in L^0(\mathbb{R}_+^{d+1})$ and $\xi_n = V_1(0, \psi(n))$ for some $\psi(n) \in \mathcal{A}$ are such that $\rho_n \in L^1(Q)$. Up to a subsequence, this convergence takes place a.s. as well.

Lemma 3.4 implies that $E_Q \int_0^1 |\psi_t(n)|^\beta (1 + |S_t|)^\beta dt$ must be finite for all n since $(\xi_n)_- \leq (\rho_n)_-$ and the latter is in $L^1(Q)$. Applying (18) with the choice $\phi := \psi(n)$ we get

$$\int_0^1 |\psi_t(n)|^\beta (1 + |S_t|)^\beta dt \leq (\rho_n)_- + 2^{\frac{\beta}{\alpha-\beta}} m^{\frac{\alpha}{(\alpha-\beta)}} + 1.$$

Now, since $Q \in \mathcal{P}$, and the sequence ρ_n is bounded in $L^1(Q)$ because it is convergent in $L^1(Q)$, it follows that

$$\sup_{n \geq 1} E_Q \int_0^1 |\psi_t(n)|^\beta (1 + |S_t|)^\beta dt < \infty. \tag{19}$$

Consider $\mathbb{L} := L^1(\Omega, \mathcal{F}, Q; \mathcal{B})$, the Banach space of $\mathbb{B}$-valued Bochner-integrable functions, where $\mathbb{B} := L^\beta([0,1], \mathbb{B}([0,1]), Leb)$ is a separable and reflexive Banach space. The functions $\psi.(n) : \Omega \to \mathbb{B}$ are easily seen to be weakly measurable, hence also strongly measurable by the separability of $\mathbb{B}$. By (19), the sequence $\phi.(n)$ is bounded in $\mathbb{L}$, so Lemma 15.1.4 in [12] yields convex combinations

$$\tilde{\psi}.(n) = \sum_{j=n}^{M(n)} \alpha_j(n) \psi.(n)$$

which converge to some $\tilde{\psi}. \in \mathbb{L}$ a.s. in $\mathbb{B}$-norm.

By the bound in (19), $\sup_n E_Q \int_0^1 |\phi_t(n)|(1+|S_t|)dt < \infty$. Now apply Lemma 9.8.1 of [12] to the sequence $\tilde{\psi}.(n)$ in the space of $(d+1)$-dimensional random variables $L^1(\Omega \times [0,1], \mathcal{O}, \nu)$, where ν is the measure defined by

$$\nu(A) := \int_{\Omega \times [0,1]} 1_A(\omega, t)(1+|S_t|)dtdQ(\omega)$$

for $A \in \mathcal{O}$ (which is finite by the choice of Q). This lemma yields convex combinations $\hat{\psi}.(n)$ of the $\tilde{\psi}.(n)$ such that $\hat{\psi}.(n)$ converges to $\psi.$ almost everywhere in ν, and hence almost everywhere in $P \times Leb$. This shows, in particular, that ψ is $\mathcal{O}$-measurable.

In particular, since $\tilde{\psi}.(n)$ converge a.s. in $\mathbb{B}$-norm, also $\hat{\psi}.(n) \to \tilde{\psi}$ a.s. in $\mathbb{B}$-norm, so $\psi = \tilde{\psi}$, $P \times Leb$-a.e. and hence $\tilde{\psi}. \to \psi.$ a.s. in $\mathbb{B}$-norm.

Define $\tilde{\xi}_n := \sum_{i=n}^{M(n)} \alpha_j(n)\xi_j$ and $\tilde{\eta}_n := \sum_{i=n}^{M(n)} \alpha_j(n)\eta_j$. It holds that $\lim_{n\to\infty} \int_0^T \tilde{\psi}_t(n) S_t dt = \int_0^T \psi_t S_t dt$ almost surely, and also

$$\lim_{n\to\infty} \tilde{\xi}_n^i = \lim_{n\to\infty} \int_0^T \tilde{\psi}_t^i(n)dt = \int_0^T \psi_t^i dt \text{ a.s. for } i = 1, \ldots, d.$$

Hence, $\tilde{\eta}_n^i \to \eta^i$ a.s. with $\eta^i := \int_0^T \tilde{\psi}_t^i dt - \rho^i \in L^0(\mathbb{R}_+)$. By the convexity of G_t,

$$\begin{aligned}
\rho^0 &= \lim_{n\to\infty} (\tilde{\xi}_n^0 - \tilde{\eta}_n^0) \\
&\le \limsup_{n\to\infty} \left[-\int_0^1 \tilde{\psi}_t(n) S_t dt - \int_0^1 G_t(\tilde{\psi}_t(n))dt - \tilde{\eta}_n^0 \right] \\
&\le \limsup_{n\to\infty} \left[-\int_0^1 \tilde{\psi}_t(n) S_t dt - \int_0^1 G_t(\psi_t)dt - \int_0^1 G_t(\tilde{\psi}_t(n))dt + \int_0^1 G_t(\psi_t)dt - \tilde{\eta}_n^0 \right] \\
&= -\int_0^1 \tilde{\psi}_t S_t dt - \int_0^1 G_t(\psi_t)dt + \limsup_{n\to\infty} \left[-\int_0^1 G_t(\tilde{\psi}_t(n))dt + \int_0^1 G_t(\psi_t)dt - \tilde{\eta}_n^0 \right].
\end{aligned}$$

Now Fatou's lemma and $\tilde{\eta}_n \in L^0(\mathbb{R}_+^{d+1})$ imply that the limit superior is in $-L^0(\mathbb{R}_+)$ (note that $G_t(\cdot)$ is continuous by convexity), hence there is $\eta^0 \in L^0(\mathbb{R}_+)$ such that

$$\rho^0 = -\int_0^1 \psi_t S_t dt - \int_0^1 G_t(\psi_t)dt - \eta^0,$$

which proves the proposition. □

The closedness property above is in fact stronger than closedness in probability, as the following corollary shows.

Corollary 3.6. *Under Assumption 2.3, the set C is closed in probability.*

Proof. Let $\rho_n \in C$ tend to ρ in probability. Up to a subsequence, convergence also holds almost surely. There exists $Q \in \mathcal{P}$ (see page 266 of [14]) such that $\rho, \sup_n |\rho - \rho_n|, \int_0^T |S_t| dt$ are all Q-integrable. Then $\rho_n \to \rho$ in $L^1(Q)$ as well and Proposition 3.5 implies that $\rho \in C$. □

$\mathbb{R}$	$\mathbb{R}^d$	$\mathbb{R}^{d+1}$
	$\bar{x} = (x_1/x_0, \dots, x_d/x_0)1_{\{x^0 \neq 0\}}$	$x = (x_0, x_1, \dots, x_d)$
	$\tilde{x} = (x_1, \dots, x_d)$	$\hat{x} = (1, x_1, \dots, x_d)$
c		$\check{c} = (c, 0, \dots, 0)$

Table 1: Summary of vector notation.

3.4 Superhedging

Finally, the main superhedging theorem. To the best of our knowledge, Theorem 3.7 is the first dual characterization in continuous-time of hedgeable contingent claims with price-impact. Results in discrete time include [2, 32, 31, 18]. Our result is inspired, in particular, by Theorem 3.1 of [18] for finite probability spaces.

Note that both terminal claims and initial endowments are multivariate, for a good reason. With price impact, which forces finite trading rates, thereby prohibiting instant purchases or sales, even in the Black-Scholes model it is impossible to buy one share of the risky asset for a sure price in finite time. Thus, superhedging of general claims in terms of cash yields mostly trivial results.

In the multivariate notation below, inequalities among vectors are understood component-wise: $x \leq y$ means that $x_i \leq y_i$ for all i. Also, for a $(d+1)$-dimensional vector x, define $\bar{x}$ as the d-dimensional vector with $\bar{x}^i = (x^i/x^0)1_{\{x^0 \neq 0\}}$, $i = 1, \dots, d$, while $\hat{x}$ denotes the $(d+1)$-dimensional vector with coordinates $\hat{x}^i = x^i$, $i = 1, \dots, d$ and $\hat{x}^0 = 1$. (See Table 1 for a summary of notation.)

Theorem 3.7. *Let $W \in L^0(\mathbb{R}^{d+1})$, $z \in \mathbb{R}^{d+1}$ and let Assumption 2.3 hold. There exists $\phi \in \mathcal{A}$ such that $V_T(z, \phi) \geq W$ a.s. if and only if*

$$Z_0 z \geq E_Q(Z_T W) - E_Q \int_0^T Z_t^0 G_t^*(\bar{Z}_t - S_t)dt, \tag{20}$$

for all $Q \in \mathcal{P}$ and for all $\mathbb{R}_+^{d+1}$-valued bounded Q-martingales Z with $Z_0^0 = 1$ satsifying $Z_t^i = 0$, $i = 1, \dots, d$ on $\{Z_t^0 = 0\}$.

The proof of the theorem in fact yields also the following slightly different version, in terms of bounded martingales only.

Theorem 3.8. *Let $W \in L^0(\mathbb{R}^{d+1})$, $z \in \mathbb{R}^{d+1}$ and let Assumption 2.3 hold. Fix a reference probability $Q \in \tilde{\mathcal{P}}(W)$. There exists $\phi \in \mathcal{A}$ such that $V_T(z, \phi) \geq W$ a.s. if and only if*

$$Z_0 z \geq E_Q(Z_T W) - E_Q \int_0^T Z_t^0 G_t^*(\bar{Z}_t - S_t)dt, \tag{21}$$

for all $\mathbb{R}_+^{d+1}$-valued bounded Q-martingales Z with $Z_0^0 = 1$ satsifying $Z_t^i = 0$, $i = 1, \dots, d$ on $\{Z_t^0 = 0\}$. □

Defining $dQ'/dQ := Z_T^0$ one can state Theorem 3.8 in the following form, in which martingale probabilities Q are replaced by stochastic discount factors Z:

Corollary 3.9. *Let $W \in L^0(\mathbb{R}^{d+1})$, $z \in \mathbb{R}^{d+1}$ and Assumption 2.3 hold. Fix a reference probability $Q \in \tilde{\mathcal{P}}(W)$. There exists $\phi \in \mathcal{A}$ such that $V_T(z,\phi) \geq W$ a.s. if and only if*

$$\hat{Z}_0 z \geq E_{Q'}(\hat{Z}_T W) - E_{Q'} \int_0^T G_t^*(Z_t - S_t)dt, \tag{22}$$

for all $Q' \ll P$ with bounded dQ'/dQ and for all $\mathbb{R}_+^d$-valued Q'-martingales Z such that $(dQ'/dQ)Z_T$ is bounded. □

Finally, in the case of a finite Ω Theorem (3.8) reduces to a simple version, without any integrability conditions:

Theorem 3.10. *Let Ω be finite. Let $W \in L^0(\mathbb{R}^{d+1})$, $z \in \mathbb{R}^{d+1}$ and let Assumption 2.3 hold. Fix any reference probability $Q \sim P$. There exists $\phi \in \mathcal{A}$ such that $V_T(z,\phi) \geq W$ a.s. if and only if*

$$Z_0 z \geq E_Q(Z_T W) - E_Q \int_0^T Z_t^0 G_t^*(\bar{Z}_t - S_t)dt, \tag{23}$$

for all $\mathbb{R}_+^{d+1}$-valued Q-martingales Z with $Z_0^0 = 1$, and satisfying $Z_t^i = 0$, $i = 1,\dots,d$ on $\{Z_t^0 = 0\}$. □

Proof of Theorem 3.7. For a $(d+1)$-dimensional vector x, $\tilde{x}$ denotes the d-dimensional vector $\tilde{x}^i := x^i$, $i = 1,\dots,d$ (cf. Table 1). First, assume that $V_T(z,\phi) \geq W$. Take $Q \in \mathcal{P}(W)$ and a bounded Q-martingale Z with nonnegative components (more generally, it is enough to assume that $Z_T W$ is Q-integrable and that $Z_T \in L^\gamma(Q)$), satisfying $Z_t^i = 0$, $i = 1,\dots,d$ on $\{Z_t^0 = 0\}$.

Note that $E_Q|W| < \infty$ and $W^0 \leq z + \int_0^T [-\phi_t S_t - G_t(\phi_t)]\, dt$ because $V_T(z,\phi) \geq W$, hence Lemma 3.4 implies

$$E_Q \int_0^T |\phi_t|^\beta (1 + |S_t|)^\beta dt < \infty. \tag{24}$$

Again, since $V_T(z,\phi) \geq W$, it follows that

$$Z_T(W - z) \leq \int_0^T \left[-Z_T^0 \phi_t S_t - Z_T^0 G_t(\phi_t) + \tilde{Z}_T \phi_t\right] dt. \tag{25}$$

By (24), Fubini's theorem applies and the properties of conditional expectations imply that

$$\begin{aligned}
E_Q(Z_T W) &\leq z E_Q Z_T + E_Q \int_0^T \left[-Z_T^0 \phi_t S_t - Z_T^0 G_t(\phi_t) + \tilde{Z}_T \phi_t\right] dt \\
&= z Z_0 + \int_0^T E_Q(-Z_T^0 \phi_t S_t - Z_T^0 G_t(\phi_t) + \tilde{Z}_T \phi_t) dt \\
&= z Z_0 + \int_0^T E_Q(-Z_t^0 \phi_t S_t - Z_t^0 G_t(\phi_t) + \tilde{Z}_t \phi_t) dt \\
&= z Z_0 + \int_0^T E_Q(-Z_t^0 \phi_t S_t - Z_t^0 G_t(\phi_t) + Z_t^0 \bar{Z}_t \phi_t) dt \\
&\leq z Z_0 + E_Q \int_0^T Z_t^0 G_t^*(\bar{Z}_t - S_t) dt,
\end{aligned}$$

which proves the first implication of this theorem.

To prove the reverse implication, suppose there is no ϕ such that $V_T(z,\phi) \geq W$, which means that $W - z \notin C$. Fix $Q \in \tilde{\mathcal{P}}(W)$. The set $C \cap L^1(Q)$ is closed in $L^1(Q)$ by Proposition 3.5. The Hahn-Banach theorem then provides a nonzero, bounded $(d+1)$-dimensional random variable $\tilde{Z}$ such that

$$E_Q[\tilde{Z}(W-z)] > \sup_{X \in C \cap L^1(Q)} E_Q[\tilde{Z}X]. \tag{26}$$

Since $-L^0(\mathbb{R}^{d+1}) \subset C$, $\tilde{Z} \geq 0$ a.s, otherwise the supremum would be infinity. Define now the (deterministic) processes $\psi(n,i)$ for all $n \in \mathbb{N}$ and $i = 1, \ldots, d$ by setting $\psi_t^i(n,i) := n$, $\psi_t^j(n,i) = 0$, $j \neq i$ for all $t \in [0,T]$.

We claim that $E_Q\tilde{Z}^0 > 0$. Otherwise, for some $i > 0$ one should have $E_Q\tilde{Z}^i > 0$. By Assumption 2.3 $\psi(n,i) \in \mathcal{A}$. By the choice of Q, we even have $V_T(0,\psi(n,i)) \in C \cap L^1(Q)$ and $E_Q\tilde{Z}V_T(0,\psi(n,i)) = nTE_Q\tilde{Z}^i \to \infty$ as $n \to \infty$, which is impossible by (26). So we conclude that $E_Q\tilde{Z}^0 > 0$. Up to a positive multiple of Z, $E_Q\tilde{Z}^0 = 1$. Define $Z_t := E_Q[\tilde{Z}|\mathcal{F}_t]$, $t \in [0,T]$.

We also claim that, for all $i = 1, \ldots, d$,

$$(P \times Leb)(A_i) = 0, \text{ where } A_i := \{(\omega,t) : Z_t^0(\omega) = 0\} \setminus \{(\omega,t) : Z_t^i(\omega) = 0\}). \tag{27}$$

If this were not the case for some i, define $\psi^i(n,i) := n1_{A_i}$, $\psi^j(n,i) := 0$, $j \neq i$. Clearly, $\psi(n,i) \in \mathcal{A}$ and $V_T(0,\psi(n,i)) \in C \cap L^1(Q)$ while $E_QV_T(0,\psi(n,i)) \to \infty$, $n \to \infty$, which is absurd, proving (27).

By the measurable selection theorem applied to the measure space $(\Omega \times [0,T], \mathcal{O}, P \otimes Leb)$ (see Proposition III.44 in [13]), there is an optional process $\tilde{\chi}(n)$ such that

$$-K \leq \tilde{\chi}_t(n)[\bar{Z}_t - S_t] - G_t(\tilde{\chi}_t(n)) \leq G_t^*(\bar{Z}_t - S_t) \wedge n$$

and

$$\tilde{\chi}_t(n)[\bar{Z}_t - S_t] - G_t(\tilde{\chi}_t(n)) \geq [G_t^*(\bar{Z}_t - S_t) \wedge n] - \frac{1}{n},$$

for $(P \times Leb)$-almost every (ω,t). Here K denotes the bound for $\sup_{t\in[0,T]} G_t(0)$ from (9). Now define $\chi_t(n) := \tilde{\chi}_t(n)1_{\{|\tilde{\chi}_t(n)| \leq N(n)\}}$ where $N(n)$ is chosen such that $(P \times Leb)(|\tilde{\chi}_t(n)| > N(n)) \leq 1/n^2$. By Assumption 2.3, $\chi(n) \in \mathcal{A}$ and by the choice of Q, $V_T(0,\chi(n)) \in C \cap L^1(Q)$. By construction,

$$\lim_{n\to\infty} \chi_t(n)[\bar{Z}_t - S_t] - G_t(\chi_t(n)) = G_t^*(\bar{Z}_t - S_t), \quad (P \times Leb) - \text{a.e.}$$

Since Z_T is bounded, Fatou's lemma implies that

$$\begin{aligned}
\liminf_{n\to\infty} E_Q Z_T V_T(0,\chi(n)) &= \liminf_{n\to\infty} E_Q \int_0^T \chi_t(n)[\tilde{Z}_T - Z_T^0 S_t] - Z_T^0 G_t(\chi_t(n))]dt \qquad (28)\\
&= \liminf_{n\to\infty} E_Q \int_0^T \chi_t(n)[\tilde{Z}_t - Z_t^0 S_t] - Z_t^0 G_t(\chi_t(n))dt \\
&= \liminf_{n\to\infty} E_Q \int_0^T \chi_t(n) Z_t^0[\bar{Z}_t - S_t] - Z_t^0 G_t(\chi_t(n))dt \\
&\geq E_Q \int_0^T Z_t^0 G_t^*(\bar{Z}_t - S_t)dt. \qquad (29)
\end{aligned}$$

From (26) we infer that

$$
\begin{aligned}
zZ_0 < \limsup_{n\to\infty} [E_Q(WZ_T) - E_Q Z_T V_T(0,\chi(n))] = \\
E_Q(WZ_T) - \liminf_{n\to\infty} E_Q Z_T V_T(0,\chi(n)) \le \\
E_Q(WZ_T) - E_Q \int_0^T Z_t^0 G_t^*(\bar{Z}_t - S_t)dt.
\end{aligned}
$$

This concludes the proof. □

Remark 3.11. The proof of Theorems 3.7 and 3.8 also shows that the statement remains valid by replacing the class of bounded martingales with the class of Q-martingales with $Z_T \in L^\gamma(Q)$ such that $Z_T W$ is Q-integrable.

For a real number c, denote by $\check{c}$ the $(d+1)$-dimensional vector $(c, 0, \ldots, 0)^T$ (cf. Table 1). The next corollary specializes Theorem 3.7 to the situation in which a claim in cash is hedged from an initial cash position only.

Corollary 3.12. *Let $W \in L^0(\mathbb{R})$, $c \in \mathbb{R}$ and let Assumption 2.3 hold. There exists $\phi \in \mathcal{A}$ such that $V_T^0(\check{c}, \phi) \ge W$ a.s. and $V_T^i(\check{c}, \phi) \ge 0$, $i = 1, \ldots, d$ if and only if*

$$
c \ge E_Q(Z_T^0 W) - E_Q \int_0^T Z_t^0 G_t^*(\bar{Z}_t - S_t)dt, \tag{30}
$$

for all $Q \in \mathcal{P}(W)$ and for all $\mathbb{R}_+^{d+1}$-valued bounded Q-martingales Z_t with $Z_0^0 = 1$ satisfying $Z_t^i = 0$, $i = 1, \ldots, d$ on $\{Z_t^0 = 0\}$. □

To understand the meaning of (30), it is helpful to consider its statement in the frictionless case, at least formally[5]. If S_t itself is a Q-martingale, then with the choice of $Z_t^0 := 1$, $Z_t^i := S_t^i$, $i = 1, \ldots, d$ the penalty term with G^* vanishes. It follows that, in order to super-replicate W, the initial endowment c must be greater than or equal to the supremum of $E_Q W$ over the set of equivalent martingale measures for S_t. This shows that our findings are formally consistent with well-known superhedging theorems for frictionless markets. The results are similarly consistent with superhedging theorems for proportional transaction costs [24], formally obtained with $G = \varepsilon S_t |x|$.

4 Arbitrage

The superhedging result proved in the previous section holds regardless of arbitrage. As a result, it can *detect* arbitrage, because any positive payoff that is superhedged for strictly less than zero is an arbitrage. Such opportunities, which start from an insolvent position and, by clever trading, yield a solvent one, are known in the literature as arbitrage of the second kind, and date back to [23]. (see also [25] in the context of large financial markets). This definition is used with markets frictions in [16, 17], and, more recently, in [33, 15, 7, 6, 30].

[5]The theorem does not apply to the frictionless case because $G = 0$ does not satisfy Assumption 2.3, and feasible strategies differ from admissible strategies

Definition 4.1. *An arbitrage of the second kind is a strategy $\phi \in \mathcal{A}$, such that $V_T(\check{c}, \phi) \geq 0$ for some $c < 0$. Absence of arbitrage of the second kind* (NA2) *holds if no such opportunity exists.*

Note that this definition requires that S has positive components. Otherwise, a a non-negative position in an asset with negative price (as $V_T(\check{c}, \phi) \geq 0$ stipulates) cannot be interpreted as solvent.

The following theorem is a direct consequence of Corollary 3.12 and Remark 3.11.

Theorem 4.2. *Let Assumption 2.3 hold. Then,* (NA2) *holds if and only if, for all $\varepsilon > 0$, there exists $Q \in \mathcal{P}$ and an $\mathbb{R}_+^{d+1}$-valued Q-martingale Z with $Z_T \in L^\gamma(Q)$ such that $E_Q \int_0^T Z_t^0 G_t^*(\bar{Z}_t - S_t)dt < \varepsilon$.*

A broad class of models enjoys the (NA2) property. Let $D \subset (0, \infty)^d$ be nonempty, open and convex. We denote by $C[t, T](D)$ (resp. $C_x[t, T](D)$) the set of continuous functions f from $[t, T]$ to D (resp. satisfying $f(t) = x$). Both spaces are equipped with the Borel sets of the topology induced by the uniform metric. Recall that a continuous stochastic process S on $[t, T]$ can be understood as a $C[t, T](D)$-valued random variable, and its support is defined in this (metric) space.

Definition 4.3. *A process S has conditional full support in D (henceforth, CFS-D) if $S \in C[0, T](D)$ a.s. and*

$$\text{supp} P(S|_{[t,T]} \in \cdot | \mathcal{F}_t) = C_{S_t}[t, T](D) \quad \textit{a.s. for all } t \in [0, T].$$

Theorem 4.4. *Let Assumption 2.3 hold with $H_t := H$ constant. If S has the* CFS-D *property, then* (NA2) *holds.*

Proof. It follows from [29] that for all ε there is $Q \sim P$ and a Q-martingale M_t evolving in $D \subset \mathbb{R}_+^d$ such that $|S_t - M_t| < \varepsilon$ a.s. for all t. Define $Z_t^i := M_t^i$ for $i = 1, \ldots, d$ and $Z_t^0 := 1$ for all t.

In [29] (see also [22]) it is shown that S_T and hence Z_T are in $L^2(Q)$. A closer inspection of the proof reveals that in fact there exist $Z_T \in L^p(Q)$ for arbitrarily large p. Take $p := \max\{\gamma, \alpha\beta/(\alpha - \beta)\}$. Then Q is `easily seen` to be in $\mathcal{P}$ and Z_T is in $L^\gamma(Q)$. The estimate (11) in Lemma 3.1 implies that

$$E_Q \int_0^T G_t^*(\bar{Z}_t - S_t)dt = E_Q \int_0^T G_t^*(M_t - S_t)dt \leq \int_0^T \ell(\varepsilon)dt \leq T\ell(\varepsilon)$$

for a continuous (deterministic) function ℓ, which clearly tends to 0 as $\varepsilon \to 0$. Now the claim follows by Theorem 4.2. □

Theorem 4.4 has an immediate implication for fractional Brownian motion. The arbitrage properties of fractional Brownian motion have long been delicate: in a frictionless setting it admits arbitrage of the second kind [34] but, with transaction costs, it does not even have arbitrage of the first kind [22]. With price-impact, the above theorem implies that it does not admit arbitrage of the second kind, since it satisfies the CFS-D property [22]. Whether arbitrage of the first kind (a positive, and possibly strictly positive, payoff from nothing) is still an open question.

5 Utility Maximization

This section discusses utility maximization in price-impact models. The first result (Theorem 5.1 below) shows that optimal strategies exist under a simple integrability assumption, which is easy to check in practice. In particular, optimal strategies exist regardless of arbitrage, since such opportunities are necessarily limited. Put differently, the budget equation is nonlinear, therefore one cannot add to an optimal strategy an arbitrage opportunity, and expect the resulting wealth to be the sum.

The second result establishes the first-order condition for utility maximization, which provides a simple criterion for optimality, and helps understand the differences with the corresponding results for frictionless markets and transaction costs. In particular, it shows that the analogue of a shadow price for price-impact models is a hypothetical frictionless price for which the optimal strategy would coincide with the execution price of the same strategy in the original price-impact model. This notion reduces to that of shadow price for markets with transaction costs.

Importantly, these results consider only utilities defined on the real line, such as exponential utility, but exclude power and logarithmic utilities, which are defined only for positive values. This setting is consistent with the definition of feasible strategies, which do not constrain wealth to remain positive. One technical challenge to optimality in such a setting is to show that wealth processes are martingales (or supermartingales) with respect to martingale measures, and Lemma 5.4 below implies such a property for any feasible strategy. Finally, since the focus is on utility functions defined on a single variable, and with price impact there is no scalar notion of portfolio value, the results below assume for simplicity that all strategies begin and end with cash only.

Let W be an arbitrary real-valued random variable (representing a random endowment) and $c \in \mathbb{R}$ the investor's initial capital.

Theorem 5.1. *Let $U : \mathbb{R} \to \mathbb{R}$ be concave and nondecreasing, and let $E|U(c + B + W)| < \infty$ hold. Under Assumption 2.3, there is $\phi^* \in \mathcal{A}'(u)$ such that*

$$EU(V_T^0(\check{c}, \phi^*) + W) = \sup_{\phi \in \mathcal{A}'(u,c)} EU(V_T^0(\check{c}, \phi) + W),$$

where $\mathcal{A}'(u, c) = \{\phi \in \mathcal{A} : V_T^i(\check{c}, \phi) = 0,\ i = 1, \ldots, d,\ EU_-(V_T^0(\check{c}, \phi) + W) < \infty\}$.

This theorem applies, in particular, for U bounded above and W bounded below.

Proof. Corollary 3.6 implies that

$$C' := \check{c} + \big(C \cap \{X : X^i = 0 \text{ a.s.},\ i = 1, \ldots, d\}\big)$$

is closed in probability.

Let $\phi(n)$ be a sequence in $\mathcal{A}'(U, c)$ with

$$\lim_{n\to\infty} EU(V_T^0(\check{c}, \phi(n)) + W) = \sup_{\phi \in \mathcal{A}'(U,c)} EU(V_T^0(\check{c}, \phi) + W).$$

Since $V_T^0(\check{c}, \phi(n)) \leq c + B$ a.s. for all n, by Lemma 9.8.1 of [12] there are convex combinations such that $\sum_{j=n}^{M(n)} \alpha_j(n) V_T^0(\check{c}, \phi(j)) \to V$ a.s. for some $[-\infty, c + B]$-valued random variable V.

By convexity of G, we have that for $\tilde{\phi}(n) := \sum_{j=n}^{M(n)} \alpha_j(n)\phi(j)$,

$$V_T^0(\check{c}, \tilde{\phi}(n)) \geq \sum_{j=n}^{M(n)} \alpha_j(n) V_T^0(\check{c}, \phi(j)),$$

so $\sum_{j=n}^{M(n)} \alpha_j(n) V_T(\check{c}, \phi(j)) \in C'$ for each n.

By the concavity of U,

$$EU\left(\sum_{j=n}^{M(n)} \alpha_j(n) V_T^0(\check{c}, \phi(j)) + W\right) \geq \sum_{j=n}^{M(n)} \alpha_j(n) EU(V_T^0(\check{c}, \phi(j))).$$

Fatou's lemma implies that $EU(V) \geq \sup_{\phi \in \mathcal{A}'(u)} EU(V_T^0(\check{c}, \phi) + W)$, in particular, V is finite-valued and hence $\check{V} \in C'$ by the convexity and closedness of C'. It follows that $V = V_T^0(\check{c}, \phi^*) - Y^0$ for some $\phi^* \in \mathcal{A}'(U, c)$ and $Y \in L_+^0$. Clearly, $EU(V_T^0(\check{c}, \phi^*) + W - Y^0) = \sup_{\phi \in \mathcal{A}'(U,c)} Eu(V_T^0(\check{c}, \phi) + W)$. Necessarily, $EU(V_T^0(\check{c}, \phi^*) + W) = \sup_{\phi \in \mathcal{A}'(U,c)} EU(V_T^0(\check{c}, \phi) + W)$ as well.[6] This completes the proof. □

Remark 5.2. The proofs of Theorem 5.1 and Proposition 3.5 use Lemmata 9.8.1 and 15.1.4 in [12]. They could be replaced, with minor modifications, with Komlós's theorem [27] and its extensions [3, 37].

While the previous result shows the existence of optimal strategies, the next theorem provides a sufficient conditions for a strategy's optimality, through a variante of the usual first order condition.

Theorem 5.3. *Let Assumption 2.3 hold, and*

a) let U be concave, continuously differentiable, with U' strictly decreasing, and

$$U(x) \leq -C|x|^{\delta}, \ x \leq 0, \tag{31}$$

for some $C > 0$ and $\delta > 1$;

b) denoting by $\tilde{U}$ the convex conjugate function of U, i.e.

$$\tilde{U}(y) := \sup_{x \in \mathbb{R}} \{U(x) - xy\}, \ y > 0,$$

assume that $\tilde{U}''(y)$ exists and is strictly positive for all $y > 0$;

c) let W be a bounded random variable;

d) let $Q \in \mathcal{P}$ be such that

$$dQ/dP \in L^{\eta}, \tag{32}$$

where $(1/\eta) + (1/\delta) = 1$. Let Z be a càdlàg process with $Z_T \in L^{\gamma'}$ for some $\gamma' > \gamma$;

e) let $G_t(\cdot)$ be $P \times Leb$-a.s. twice continuously differentiable in x and $G_t''(x) > 0$ for all x;

[6] Note that U can be constant on an (infinite) interval hence $Y^0 \neq 0$ is possible.

f) let ϕ^* be a feasible strategy such that, for some $y^* > 0$, the following conditions hold:

i) Z is a Q-martingale;

ii) $U'(V_T^0(x,\phi^*) + W) = y^*(dQ/dP)$;

iii) $Z_t = S_t + G_t'(\phi_t^*)$ a.s. in $P \times Leb$;

iv) $E_Q\left(V_T^0(x,\phi^*) - \int_0^T G_t^*(Z_t - S_t)dt\right) = x$.

Then the strategy ϕ^* is optimal for the problem

$$\max_{\phi \in \mathcal{A}'(U,c)} E\left[U(V_T^0(x,\phi) + W)\right]. \tag{33}$$

Proof. For any feasible strategy $(\phi_t)_{t\geq 0}$ such that $\Phi_t = \int_0^t \phi_s ds$ satisfies $\Phi_T = 0$, the final payoff equals

$$V_T^0(x,\phi) = x - \int_0^T S_t\phi_t dt - \int_0^T G_t(\phi_t)dt. \tag{34}$$

Let Z_t be as in the statement of the Theorem, and rewrite the above payoff as:

$$V_T^0(x,\phi) = x - \int_0^T Z_t\phi_t dt + \int_0^T (Z_t - S_t)\phi_t dt - \int_0^T G_t(\phi_t)dt.$$

By definition of G_t^* it follows that:

$$V_T^0(x,\phi) \leq x + \int_0^T Z_t\phi_t dt + \int_0^T G_t^*(Z_t - S_t)dt, \tag{35}$$

and equality holds if $Z_t - S_t = G_t'(\phi_t)$, $P \times Leb$-a.s., that is, when iii) holds.

It follows from Lemma 5.4 that:

$$0 \leq E_Q\left[\left(x - V_T^0(x,\phi) + \int_0^T G_t^*(Z_t - S_t)dt\right)\right] \tag{36}$$

Thus, for any payoff $V_T^0(x,\phi) + W$ and any $y > 0$ the following holds:

$$\begin{aligned} E\left[U(V_T^0(x,\phi) + W)\right] \leq & E\left[U(V_T^0(x,\phi) + W) + y(dQ/dP)\left(x - V_T^0(x,\phi) + \int_0^T G_t^*(Z_t - S_t)dt\right)\right] \\ \leq & E\left[\tilde{U}(y(dQ/dP)) + y(dQ/dP)\left(\int_0^T G_t^*(Z_t - S_t)dt + W\right)\right] + yx. \end{aligned} \tag{37}$$

Equality holds if both conditions ii) and iii) are satisfied. Since the above inequality holds for any $y > 0$, it follows that:

$$\sup_{\phi \in \mathcal{A}'(U,c)} E\left[U(V_T^0(x,\phi) + W)\right] \leq \inf_{y>0}\left(E\left[\tilde{U}(y(dQ/dP)) + y(dQ/dP)\left(\int_0^T G_t^*(Z_t - S_t)dt + W\right)\right] + yx\right) \tag{38}$$

The infimum on the right-hand side is achieved at y^* if the following condition holds:

$$E_Q\left[-\tilde{U}'(y^*(dQ/dP)) - \left(\int_0^T G_t^*(Z_t - S_t)dt + W\right)\right] = x. \tag{39}$$

Since $-\tilde{U}' = (U')^{-1}$, the above condition, combined with ii), reduces to

$$E_Q\left(V_T^0(x,\phi) - \int_0^T G_t^*(Z_t - S_t)dt\right) = x \tag{40}$$

which coincides with condition iv). Thus, if conditions i), ii), iii) and iv) hold for ϕ^* then, by (37),

$$E\left[U(V_T^0(x,\phi^*) + W)\right] = E\left[\tilde{U}(y^*(dQ/dP)) + y^*(dQ/dP)\left(\int_0^T G_t^*(Z_t - S_t)dt + W\right)\right] + y^*x.$$

For all $\phi \in \mathcal{A}'(U,c)$

$$E\left[U(V_T^0(x,\phi) + W)\right] \leq E\left[\tilde{U}(y^*(dQ/dP)) + y^*(dQ/dP)\left(\int_0^T G_t^*(Z_t - S_t)dt + W\right)\right] + y^*x,$$

by (38). Hence the strategy ϕ^* is indeed optimal. □

Lemma 5.4. *Under the assumptions of the previous Theorem, any $\phi \in \mathcal{A}'(U,c)$ satisfies*

$$E_Q \int_0^T \phi_t Z_t dt = 0.$$

Proof. Assume $T = 1$. Define

$$\Phi_t^+ := \int_0^t (\phi_s)_+ ds, \quad \Phi_t^- := \int_0^t (\phi_s)_- ds.$$

We will show $E_Q \int_0^1 Z_t d\Phi_t^+ = E_Q \int_0^1 Z_t d\Phi_t^- = 0$.

Since $\phi \in \mathcal{A}'(U,c)$, (31) and Hölder's inequality imply that $E_Q[V_1^0(x,\phi)]_- < \infty$, hence Lemma 3.4 implies that

$$E_Q \int_0^1 |\phi_t|^\beta (1 + |S_t|)^\beta dt < \infty,$$

a fortiori,

$$E_Q(\Phi_1^+)^\beta = E_Q\left(\int_0^1 (\phi_t)_+ dt\right)^\beta < \infty. \tag{41}$$

Define $\Phi_t^+(n) := \Phi^+(k_n(t))$ where

$$k_n(t) := \max\{i \in \mathbb{N} : \frac{i}{n} \leq t\}.$$

and observe that $d\Phi_t^+(n) \to d\Phi_t^+$ a.s. in the sense of weak convergence of measures on $\mathcal{B}([0,1])$. As Z_t is a.s. càdlàg, its trajectories have countably many points of discontinuity (a.s.). By $d\Phi_t^+ \ll Leb$, this implies

$$Y_n^+ := \int_0^1 Z_t d\Phi_t^+(n) \to \int_0^1 Z_t d\Phi_t^+ =: Y^+,$$

almost surely. Furthermore,

$$|\int_0^1 Z_t d\Phi_t^+(n)| = |\sum_{k=1}^n Z_{k/n}[\Phi_{k/n}^+(n) - \Phi_{(k-1)/n}^+(n)]| \leq \sup_t |Z_t|\Phi_1^+ \tag{42}$$

where $\sup_{t\in[0,T]}|Z_t| \in L^{\gamma'}$ by assumption and $\Phi_1^+ \in L^\beta$ by (41). It follows by Hölder's inequality that the sequence Y_n^+ is Q-uniformly integrable, so $E_Q Y_n^+ \to E_Q Y^+$, $n \to \infty$. From (42) we get, noting that $\Phi_0^+(n) = 0$,

$$E_Q Y_n^+ = E_Q \left[\sum_{l=0}^{n-1} (Z_{l/n} - Z_{(l+1)/n})\Phi_{l/n}^+(n)\right] + E_Q Z_1 \Phi_1^+(n) = E_Q Z_1 \Phi_1^+(n), \tag{43}$$

by the Q-martingale property of Z. Analogously, as $n \to \infty$,

$$E_Q Y_n^- = E_Q Z_1 \Phi_1^-(n) \to E_Q Y^-,$$

where Y_n^- is defined analogously to Y_n^+ using $d\Phi_t^-$ instead of $d\Phi_t^+$ and

$$Y^- := \int_0^1 Z_t d\Phi_t^-.$$

Since $\Phi_1(n) = \Phi_1 = 0$, (43) implies that $E_Q(Y_n^+ - Y_n^-) = 0$ for all n, whence also

$$E_Q(Y^+ - Y^-) = E_Q \int_0^T \phi_t Z_t dt = 0,$$

concluding the proof. □

References

[1] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40, 2001.

[2] F. Astic and N. Touzi. No arbitrage conditions and liquidity. *J. Math. Econom.*, 43(6):692–708, 2007.

[3] E. J. Balder. Infinite-dimensional extension of a theorem of Komlós. *Probab. Theory Related Fields*, 81(2):185–188, 1989.

[4] D. Bertsimas and A. Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1:1–50, 1998.

[5] Fischer Black. Noise. *The Journal of Finance*, 41(3):529–543, 1986.

[6] Bruno Bouchard and Adrien Nguyen Huu. No marginal arbitrage of the second kind for high production regimes in discrete time production-investment models with proportional transaction costs. *Math. Finance*, 23(2):366–386, 2013.

[7] Bruno Bouchard and Erik Taflin. No-arbitrage of second kind in countable markets with proportional transaction costs. *Ann. Appl. Probab.*, 23(2):427–454, 2013.

[8] Umut Çetin, Robert A. Jarrow, and Philip Protter. Liquidity risk and arbitrage pricing theory. *Finance Stoch.*, 8(3):311–341, 2004.

[9] Umut Çetin and L. C. G. Rogers. Modeling liquidity effects in discrete time. *Math. Finance*, 17(1):15–29, 2007.

[10] Umut Çetin, H. Mete Soner, and Nizar Touzi. Option hedging for small investors under liquidity costs. *Finance Stoch.*, 14(3):317–341, 2010.

[11] Rama Cont, Arseniy Kukanov, and Sasha Stoikov. The price impact of order book events. *Journal of Financial Econometrics*, 2013.

[12] F. Delbaen and W. Schachermayer. *The mathematics of arbitrage*. Springer Finance. Springer-Verlag, Berlin, 2006.

[13] C. Dellacherie and P.-A. Meyer. *Probabilities and potential*, volume 29 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1978.

[14] C. Dellacherie and P.-A. Meyer. *Probabilities and potential. B*, volume 72 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1982. Theory of martingales, Translated from the French by J. P. Wilson.

[15] E. Denis and Yu. Kabanov. Consistent price systems and arbitrage opportunities of the second kind in models with transaction costs. *Finance Stoch.*, 16(1):135–154, 2012.

[16] J. C. Dermody and R. T. Rockafellar. Cash stream valuation in the face of transaction costs and taxes. *Math. Finance*, 1(1):31–54, 1991.

[17] J. C. Dermody and R. T. Rockafellar. Tax basis and nonlinearity in cash stream valuation. *Math. Finance*, 5(2):97–119, 1995.

[18] Y. Dolinsky and H. M. Soner. Duality and convergence for binomial markets with friction. Forthcoming.

[19] Alfonso Dufour and Robert F Engle. Time and the price impact of a trade. *The Journal of Finance*, 55(6):2467–2498, 2000.

[20] N. Garleanu and L. Pedersen. Dynamic trading with predictable returns and transaction costs. *Journal of Finance*. Forthcoming.

[21] P. Guasoni, E. Lépinette, and M. Rásonyi. The fundamental theorem of asset pricing under transaction costs. *Finance Stoch.*, 16(4):741–777, 2012.

[22] P. Guasoni, M. Rásonyi, and W. Schachermayer. Consistent price systems and face-lifting pricing under transaction costs. *Ann. Appl. Probab.*, 18(2):491–520, 2008.

[23] J. E. Ingersoll. *Theory of Financial Decision Making*. Studies in Financial Economics. Rowman & Littlefield Publishers, Inc., Lanham, MD, 1987.

[24] Yu. Kabanov and M. Safarian. *Markets with transaction costs*. Springer Finance. Springer-Verlag, Berlin, 2009. Mathematical theory.

[25] Yu. M. Kabanov and D. O. Kramkov. Large financial markets: asymptotic arbitrage and contiguity. *Teor. Veroyatnost. i Primenen.*, 39(1):222–229, 1994.

[26] Jan Kallsen and Johannes Muhle-Karbe. On using shadow prices in portfolio optimization with transaction costs. *The Annals of Applied Probability*, 20(4):1341–1358, 2010.

[27] J. Komlós. A generalization of a problem of Steinhaus. *Acta Math. Acad. Sci. Hungar.*, 18:217–229, 1967.

[28] A. Kyle. Continuous auctions and insider trading. *Econometrica*, 29:1315–1335, 1985.

[29] F. Maris and H. Sayit. Consistent price systems in multiasset markets. *Int. J. Stoch. Anal.*, 2012, 2012.

[30] T. Pennanen. Indifference pricing in illiquid markets. Working Paper, King's College London.

[31] T. Pennanen. Superhedging in illiquid markets. *Math. Finance*, 21(3):519–540, 2011.

[32] T. Pennanen and I. Penner. Hedging of claims with physical delivery under convex transaction costs. *SIAM J. Financial Math.*, 1:158–178, 2010.

[33] M. Rásonyi. Arbitrage under transaction costs revisited. In *Optimality and risk—modern trends in mathematical finance*, pages 211–225. Springer, Berlin, 2009.

[34] L. C. G. Rogers. Arbitrage with fractional Brownian motion. *Math. Finance*, 7(1):95–105, 1997.

[35] L. C. G. Rogers and S. Singh. The cost of illiquidity and its effects on hedging. *Math. Finance*, 20(4):597–615, 2010.

[36] A. Schied and T. Schöneborn. Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets. *Finance Stoch.*, 13(2):181–204, 2009.

[37] H. v. Weizsäcker. Can one drop L^1-boundedness in Komlós's subsequence theorem? *Amer. Math. Monthly*, 111(10):900–903, 2004.