ADMISSIBLE STRATEGIES IN SEMIMARTINGALE PORTFOLIO SELECTION

SARA BIAGINI* AND ALEŠ ČERNÝ†

Dedicated to Walter Schachermayer on the occasion of his 60th birthday.

Abstract. The choice of admissible trading strategies in mathematical modelling of financial markets is a delicate issue, going back to Harrison and Kreps [HK79]. In the context of optimal portfolio selection with expected utility preferences this question has been the focus of considerable attention over the last twenty years.

We propose a novel notion of admissibility that has many pleasant features – admissibility is characterized purely under the objective measure P; each admissible strategy can be approximated by simple strategies using finite number of trading dates; the wealth of any admissible strategy is a supermartingale under all pricing measures; local boundedness of the price process is not required; neither strict monotonicity, strict concavity nor differentiability of the utility function are necessary; the definition encompasses both the classical mean-variance preferences and the monotone expected utility.

For utility functions finite on $\mathbb{R}$, our class represents a minimal set containing simple strategies which also contains the optimizer, under conditions that are milder than the celebrated *reasonable asymptotic elasticity* condition on the utility function.

Key words. utility maximization, non locally bounded semimartingale, incomplete market, σ-localization and $\mathcal{I}$-localization, σ-martingale measure, Orlicz space, convex duality

AMS subject classifications. primary 60G48, 60G44, 49N15, 91B16; secondary 46E30, 46N30

JEL subject classifications. G11, G12, G13

1. Introduction. A central concept of financial theory is the notion of a self-financing investment strategy H, whose discounted wealth is expressed mathematically by the stochastic integral

$$x + H \cdot S_t := x + \int_{(0,t]} H_s dS_s,$$

where S is a semimartingale process on a stochastic basis $(\Omega, (\mathcal{F}_t)_{0\leq t\leq T}, P)$, representing discounted prices of d traded assets, and x is the initial wealth.

Stochastic integration theory formulates minimal requirements for the integral above to exist, see Protter [Pr05]. The class of predictable processes H for which the integral exists is denoted by $L(S;P)$ or simply $L(S)$. However, the *whole* of $L(S)$ is *not appropriate* for financial applications. Specifically, Harrison and Kreps [HK79] noted that when all processes in $L(S)$ are allowed as trading strategies, arbitrage opportunities arise even in the standard Black-Scholes model. This is not a problem of the model S – the reason is that the theory of stochastic integration operates with a set of integrands far too rich for such applications. The solution proposed by the subsequent no-arbitrage literature, see [Sch94, DS98], is to restrict attention to a subset $\mathcal{H}^b \subseteq L(S)$ of strategies whose wealth is bounded uniformly from below by a constant.

*University of Pisa (sara.biagini@ec.unipi.it). Part of this research was conducted while visiting Collegio Carlo Alberto in Moncalieri, Turin, Italy in Spring 2009. Warm hospitality and financial support of the Collegio are gratefully acknowledged.

†Cass Business School, City University London (Ales.Cerny.1@city.ac.uk).

Now consider a concave non decreasing utility function U and an agent who wishes to maximize the expected utility of her terminal wealth, $E[U(x + H \cdot S_T)]$. In this context, $\mathcal{A} \subseteq L(S)$ will be a good set of trading strategies if the utility maximization over $H \in \mathcal{A}$ is well posed and if $\mathcal{A}$ contains the optimizer,

$$U(+\infty) > \sup_{H\in\mathcal{A}} E[U(x + H \cdot S_T)] = \max_{H\in\mathcal{A}} E[U(x + H \cdot S_T)].$$

Historically, the search for a good definition of admissibility has proved to be a difficult task and it has evolved in two streams. For utility functions finite on a half-line, for example a logarithmic utility, there is a natural definition: admissible strategies are again those in $\mathcal{H}^b$, see [KS99, CSW01, KS03]. Remarkably, this theoretical framework is valid for *any* arbitrage-free S.

For utility functions finite on the whole $\mathbb{R}$, the situation is more complicated. The definition of admissibility via $\mathcal{H}^b$ works only to a certain extent. Here S has to be locally bounded (or σ-bounded) to ensure that $\mathcal{H}^b$ is sufficiently rich for a duality framework to work, cf. [Sch01]. Moreover, the class $\mathcal{H}^b$ will typically fail to contain the optimizer – this happens, for example, in the classical Black-Scholes model under exponential utility.

A possible choice in this situation is to consider all strategies whose wealth is a martingale under all suitably defined pricing measures (see §3.1). This approach works well for exponential utility, cf. [Dal02, KSt02]. However, the seminal work of Schachermayer [Sch03] shows that, for general utilities, the martingale class is too narrow to catch the optimizer. The optimal strategy only exists among strategies whose wealth is a *supermartingale* under all pricing measures. For this reason, the *supermartingale class* is now considered the best notion of admissibility.

It is evident from our discussion that admissibility is currently defined in a primal way for utility functions finite on $\mathbb{R}_+$ but for utilities finite on $\mathbb{R}$ *the definition is dual, via pricing measures*. A connection between the two approaches is foreshadowed in Schachermayer [Sch01] who defines a set of *admissible terminal wealths* as those positions whose utility can be approximated in $L^1(P)$ by strategies with wealth bounded from below. Under suitable technical assumptions, the optimal wealth exists and there is a trading strategy in the supermartingale class which leads to the optimal wealth, see also Owen [O02] and Bouchard et al. [BTZ04].

All of the papers above dealing with utility finite on $\mathbb{R}$ use *locally bounded price processes*. Biagini and Frittelli [BF05] employ a wider class of well-behaved price processes compatible with the utility U. In [BF07] they show that for this class of price processes there is always an optimizer in Schachermayer's set of supermartingale strategies. In a subsequent paper [BF08], they propose a unified treatment for utility functions finite on a half-line as well as those finite on the whole $\mathbb{R}$, for an even wider class of semimartingales S. As we show in §3.1 their hypotheses on S amount to our Assumption 3.1. In contrast to the present paper, [BF08] use admissible strategies $\mathcal{H}^W$ whose wealth is controlled from below by (a multiple of) an exogenously given, fixed random variable $W > 0$. When W is constant, one recovers the usual set $\mathcal{H}^b$ of strategies with wealth bounded uniformly from below. Here, too, the optimal strategy may fail to be in $\mathcal{H}^W$, there is no approximation result for the optimizer, and when S is not particularly well behaved the optimizer may in principle depend on the choice of the loss control W.

The philosophy of the present paper is to make the definition of admissibility general enough to provide a "unified treatment" of utility functions in the spirit of [BF08], while keeping the definition as natural and intuitive as possible by *not* resorting to duality. We use a bottom-up approach whereby we first define a class of well-behaved simple trading strategies $\mathcal{H}$ which can be interpreted as buy-and-hold strategies over finitely many dates (see Definition 3.2 for details). In the locally bounded case $\mathcal{H}$ corresponds to buy-and-hold strategies whose wealth is uniformly bounded in absolute value. We then define admissible strategies $\overline{\mathcal{H}}$ as suitable limits of strategies in $\mathcal{H}$.

Definition 1.1. *$H \in L(S)$ is an* admissible integrand *if $U(H \cdot S_T) \in L^1(P)$ and if there exists an approximating sequence $(H^n)_n$ in $\mathcal{H}$ such that:*

(i) *$H^n \cdot S_t \to H \cdot S_t$ in probability for all $t \in [0, T]$;*
(ii) *$U(H^n \cdot S_T) \to U(H \cdot S_T)$ in $L^1(P)$.*

The set of all admissible integrands is denoted by $\overline{\mathcal{H}}$.

The two requirements above are natural assumptions if considered *separately.* Item (i) is in the spirit of the construction of the stochastic integral itself, while item (ii) ensures that utility of an admissible strategy can be approximated by the utility from simple strategies. Definition 1.1 combines these *two desirable approximation features* together.

The key point of the present paper is that we do not ask for approximation of terminal utility *only*, as is done in [Sch01, O02, BTZ04], but we also require an approximation of the wealth process at intermediate times, as in Černý and Kallsen [ČK07, Definition 2.2]. What is more, our definition does not rely on regularity properties of U, such as strict concavity, strict monotonicity or differentiability.

Our results then follow rather smoothly: $\overline{\mathcal{H}}$ is a subset of the supermartingale class (Proposition 3.8) and the optimizer belongs to $\overline{\mathcal{H}}$ under very mild conditions, as shown in the main Theorem 4.10. Therefore, as a byproduct, we also obtain an extremely compact proof of the supermartingale property of the optimal solution.

The paper is organized as follows. In §2.1-§2.3 there are basic definitions from convex analysis, theory of Orlicz spaces and stochastic integration. Section 2.4 contains a new result on σ-localization. In §3.1 and §3.2 we discuss conditions imposed on the price process S and the corresponding definitions of simple strategies. In §3.3 we prove the martingale property of simple strategies. In §3.4 we define the admissible strategies and prove their supermartingale property. In §4.1 and §4.2 we discuss the customary conditions of *reasonable asymptotic elasticity* and other related conditions used in the literature and we contrast them with a weaker Inada condition at $+\infty$ employed in this paper. The main result (Theorem 4.10) is stated and proved in §4.3. Section 5 provides more details on the main assumptions and on the advantages of our framework compared to the existing literature. Section 6 contains technical lemmata.

2. Mathematical preliminaries.

2.1. Utility functions.

A utility function U is a proper, concave, non-decreasing, upper semicontinuous function. Its effective domain is the non-empty set

$$\operatorname{dom} U := \{x \mid U(x) > -\infty\}. \tag{2.1}$$

The infimum of the effective domain of U is denoted by

$$\underline{x} := \inf(\operatorname{dom} U). \tag{2.2}$$

Let $U(+\infty) := \lim_{x \to +\infty} U(x)$ and define

$$\overline{x} := \inf\{x \mid U(x) = U(+\infty)\}. \tag{2.3}$$

In the economic literature $\overline{x}$ is known as the *satiation point* or *bliss point*. For strictly increasing utility functions $\overline{x} = +\infty$, while for truncated utility functions, which feature for example in shortfall risk minimization, $\overline{x} < +\infty$ represents a point where further increase in wealth does not produce additional enjoyment in terms of utility. In economics this is interpreted as the point of maximum satisfaction, or bliss.

By construction $\underline{x} \leq \overline{x}$ and the equality arises only when U is constant on its entire effective domain in which case the utility maximization problem is trivial since "doing nothing" is always optimal. Therefore, modulo a translation, the following assumption entails no loss of generality.

ASSUMPTION 2.1. $\underline{x} < 0 < \overline{x}$ *and* $U(0) = 0$.

The convex conjugate of U is defined by

$$V(y) := \sup_{x\in\mathbb{R}}\{U(x) - xy\}.$$

Our assumptions on U imply that V is a proper, convex, lower semi-continuous function, equal to $+\infty$ on $(-\infty, 0)$, and it verifies $V(0) = U(+\infty)$. For example, with exponential utility one obtains the following conjugate pair of functions U, V:

$$U(x) = 1 - e^{-x}; \qquad V(y) = y\ln y - y + 1. \tag{2.4}$$

In the sequel we will often exploit the following form of the Fenchel inequality, obtained as a simple consequence of the definition of V:

$$U(x) \leq xy + V(y). \tag{2.5}$$

2.2. Young functions, Orlicz spaces and the Orlicz space induced by U. We recall basic facts on Young functions and induced Orlicz spaces. The interested reader is referred to the monographs by Rao and Ren [RR91] and Krasnosel'skii and Rutickii [KR61] for proofs.

A Young function $\Psi : \mathbb{R} \to [0, +\infty]$ is an even, convex and lower semicontinuous function with the properties:

(i) $\Psi(0) = 0$; (ii) $\Psi(+\infty) = +\infty$; (iii) $\Psi < +\infty$ on an open neighborhood of 0.

Note that Ψ may jump to $+\infty$ outside a bounded neighborhood of 0, but when Ψ is finite valued, it is also continuous by convexity. In either case, Ψ is nondecreasing over $\mathbb{R}_+$ and *countably convex* (see Lemma 6.1).

The Orlicz space L^Ψ induced by Ψ on $(\Omega, \mathcal{F}_T, P)$ is defined as

$$L^\Psi = \{X \in L^0 \mid E[\Psi(cX)] < +\infty \text{ for some } c > 0\}.$$

It is a Banach space when endowed with the Luxemburg (gauge) norm

$$N_\Psi(X) = \inf\left\{k > 0 \mid E\left[\Psi\left(\frac{X}{k}\right)\right] \leq 1\right\}.$$

Orlicz spaces are generalizations of L^p spaces whereby $\Psi(x) = |x|^p, p \geq 1$ yields $L^\Psi \equiv L^p$, while $\Psi(x) = I_{\{|x|\leq 1\}}$ induces the space L^∞ with the supremum norm. Intuitively, the faster Ψ increases to $+\infty$ the smaller the space L^Ψ and the stronger its topology. It is also clear that two *distinct* choices of the Young function may give rise to *isomorphic* Orlicz spaces, the Luxemburg norms being equivalent. These statements are made precise by the following definition and theorem.

DEFINITION 2.2 (Krasnosel'skii and Rutickii). *Let Ψ_1 and Ψ_2 be two Young functions. We write $\Psi_1 \succeq \Psi_2$, if there are constants $\lambda > 0$ and x_0 such that for $x \geq x_0$,*

$$\Psi_1(\lambda x) \geq \Psi_2(x).$$

We say that Ψ_1 and Ψ_2 are equivalent if $\Psi_1 \succeq \Psi_2$ and $\Psi_1 \preceq \Psi_2$.

THEOREM 2.3 (Krasnosel'skii and Rutickii). *The following statements are equivalent:*

(i) $\Psi_1 \succeq \Psi_2$;
(ii) $L^{\Psi_1} \hookrightarrow L^{\Psi_2}$;
(iii) *there is $\lambda > 0$ such that*

$$N_{\Psi_2}(X) \leq \lambda N_{\Psi_1}(X) \text{ for all } X \in L^{\Psi_1}.$$

Consequently, any Orlicz space L^Ψ satisfies the embeddings

$$L^\infty \hookrightarrow L^\Psi \hookrightarrow L^1,$$

and two Orlicz spaces are isomorphic if and only if their Young functions are equivalent.

The *Morse subspace* of L^Ψ, also called the "Orlicz heart", is given by

$$M^\Psi = \{X \in L^0 \mid E[\Psi(cX)] < \infty \text{ for all } c > 0\}.$$

The inclusion of M^Ψ in L^Ψ may be strict and in particular $M^\Psi = \{0\}$ when $L^\Psi = L^\infty$. On the other hand, $M^p = L^p$ for any $1 \leq p < +\infty$. More generally, when Ψ is finite on $\mathbb{R}$ then

$$L^\infty \hookrightarrow M^\Psi \hookrightarrow L^\Psi. \tag{2.6}$$

We end these considerations with a classic example of strict inclusion of M^Ψ in L^Ψ.

EXAMPLE 2.4. *Let $\Psi(x) = (\cosh x - 1)$. Simple calculations show that L^Ψ is the space of random variables X with* some *absolute exponential moment finite, $E[e^{c|X|}] < +\infty$ for some $c > 0$. M^Ψ is the proper subspace of those X with* all *absolute exponential moments finite. Therefore, as soon as Ω is infinite, $M^\Psi \subsetneqq L^\Psi$.*

From §3 onwards, the Young function will be

$$\widehat{U}(x) := -U(-|x|),$$

meaning that the Orlicz space in consideration is generated by the lower tail of the utility function. Then,

$$X \in L^{\widehat{U}} \text{ iff } E[U(-c|X|)] > -\infty \text{ for some } c > 0. \tag{2.7}$$

For utility functions with lower tail which is asymptotically a power, say $p > 1$, $L^{\widehat{U}}$ is isomorphic to L^p and $L^{\widehat{U}} \equiv M^{\widehat{U}}$. When U is exponential, say $U(x) = 1 - e^{-\gamma x}$, with $\gamma > 0$, $\widehat{U}(x) = e^{\gamma|x|} - 1$ and the induced space is isomorphic to that of Example 2.4, so that $L^{\widehat{U}} \supsetneqq M^{\widehat{U}}$ in the relevant case $|\Omega| = +\infty$.

For utility functions with half-line as their effective domain, such as $U(x) = \ln(1 + x)$, $L^{\widehat{U}}$ is isomorphic to L^∞ and $M^{\widehat{U}} = \{0\}$.

2.3. Semimartingale norms. There are two standard norms in stochastic calculus. Let S be an $\mathbb{R}^d$-valued semimartingale on the filtered space $(\Omega, (\mathcal{F}_t)_{0\leq t\leq T}, P)$ and let $S_t^* = \sum_{i=1}^d \sup_{0\leq s\leq t} |S_s^i|$ be the corresponding maximal process. For $p \in [1,\infty]$ let

$$\|S\|_{\mathscr{S}^p} := \|S_T^*\|_{L^p},$$

and denote the class of semimartingales with finite $\mathscr{S}^p$-norm also by $\mathscr{S}^p$. This definition is due to Meyer [M78]. We extend the definition slightly to allow for an arbitrary Orlicz space $L^\Psi(P)$ or its Morse subspace $M^\Psi(P)$,

$$\mathscr{S}^\Psi := \{\text{semimartingale } S \mid S_T^* \in L^\Psi\}, \tag{2.8}$$

$$\mathscr{S}^{M^\Psi} := \{\text{semimartingale } S \mid S_T^* \in M^\Psi\}. \tag{2.9}$$

REMARK 2.5. *Note for future use that $\mathscr{S}^\Psi$ and $\mathscr{S}^{M^\Psi}$ are stable under stopping, that is if S belongs to $\mathscr{S}^\Psi$ or $\mathscr{S}^{M^\Psi}$ and if τ is a stopping time, then the stopped process $S^\tau := (S_{\tau\wedge t})_t$ is in $\mathscr{S}^\Psi$ or $\mathscr{S}^{M^\Psi}$, respectively.*

Following Protter [Pr05], for any special semimartingale S with canonical decomposition into local martingale part M and predictable finite variation part A, $S = S_0 + M + A$, we define the following semimartingale norm,

$$\|S\|_{\mathscr{H}^p} = \|S_0\|_{L^p} + \|[M,M]_T^{1/2}\|_{L^p} + \|\mathrm{var}(A)_T\|_{L^p},$$

where $\mathrm{var}(A)$ denotes the absolute variation of process A. The class of processes with finite $\mathscr{H}^p$-norm is denoted by $\mathscr{H}^p$. As usual we let

$$\mathscr{M}^p := \mathscr{H}^p \cap \mathscr{M},$$

where $\mathscr{M}$ is the set of uniformly integrable P-martingales.

2.4. Localization and beyond: σ-localization and $\mathcal{I}$-localization. Recall that for a given semimartingale S on $(\Omega, (\mathcal{F}_t)_{0\leq t\leq T}, P)$, $L(S)$ indicates the class of predictable and $\mathbb{R}^d$-valued, S-integrable processes H under P, while $H\cdot S$ denotes the resulting scalar-valued integral process. In contrast, when φ is a *scalar* predictable process belonging to $\cap_{i=1}^d L(S^i)$ we follow [Pr05, §IV.9] and [DS05, Definition 8.3.2] in writing $\varphi\cdot S$ for the *vector-valued* process $(\varphi\cdot S^1, \dots, \varphi\cdot S^d)$.

Now, let $\mathscr{C}$ be some fixed class of semimartingales. The following methods of extending $\mathscr{C}$ appear in the literature:

(i) $S \in \mathscr{C}_{\mathrm{loc}}$, i.e. S is *locally* in $\mathscr{C}$, if there is a sequence of stopping times τ_n increasing to $+\infty$ (called *localizing sequence*) such that each of the stopped processes $S^{\tau_n} = I_{[0,\tau_n]}\cdot S$ is in $\mathscr{C}$.

(ii) $S \in \mathscr{C}_\sigma$, i.e. S is *σ-locally* in $\mathscr{C}$, if there is a sequence of predictable sets D_n increasing to $\Omega\times\mathbb{R}_+$ such that for every n the vector-valued process $I_{D_n}\cdot S$ is in $\mathscr{C}$.

(iii) $S \in \mathscr{C}_{\mathcal{I}}$, i.e. S is *$\mathcal{I}$-locally* in $\mathscr{C}$, if there is some scalar process $\varphi \in \cap_{i=1}^d L(S^i)$, $\varphi > 0$ such that $\varphi\cdot S$ is in $\mathscr{C}$.

The first two items are standard (cf. [JS03, I.1.33], [Ka04]) while the third item is an ad hoc definition. By construction, for an arbitrary semimartingale class $\mathscr{C}$ one has $\mathscr{C}_\sigma \supseteq \mathscr{C}_{\mathrm{loc}} \supseteq \mathscr{C}$. However it is not *a priori* clear what inclusions hold for $\mathscr{C}_{\mathcal{I}}$, apart from the obvious $\mathscr{C}_{\mathcal{I}} \supseteq \mathscr{C}$. Émery [E80, Proposition 2] has shown that when $\mathscr{C} = \mathscr{M}^p$ or $\mathscr{H}^p$, the following equalities hold

$$\mathscr{M}^p_\sigma = \mathscr{M}^p_{\mathcal{I}}, \qquad \mathscr{H}^p_\sigma = \mathscr{H}^p_{\mathcal{I}}, \qquad \text{for } p \in [1,+\infty). \tag{2.10}$$

To complicate matters, some authors use σ-localization to mean $\mathcal{I}$-localization, see [DS98, Pr05, KSî06]. In this paper we deliberately make a *clear distinction* between the two localization procedures.

The name $\mathcal{I}$-localization ($\mathcal{I}$ standing for integral) is probably a misnomer, since no localization procedure is involved. But we have chosen it because in Émery's result $\mathcal{I}$-localization coincides with σ-localization. In general, however, $\mathscr{C}_\mathcal{I} \neq \mathscr{C}_\sigma$. Intuition suggests that the two localizations coincide whenever the primary class $\mathscr{C}$ is defined via some sort of integrability property, as in the case above: martingale property and its generalizations, boundedness or more generally Orlicz integrability conditions on the maximal process. The next result in this direction appears to be new.

PROPOSITION 2.6. *For any Orlicz space L^Ψ, its Morse subspace M^Ψ and the corresponding semimartingale normed spaces $\mathscr{S}^\Psi, \mathscr{S}^{M^\Psi}$, the following identities hold: $\mathscr{S}^\Psi_\sigma = \mathscr{S}^\Psi_\mathcal{I}$ and $\mathscr{S}^{M^\Psi}_\sigma = \mathscr{S}^{M^\Psi}_\mathcal{I}$.*

Proof. We prove the statement only for $\mathscr{S}^\Psi$, since the proof for $\mathscr{S}^{M^\Psi}$ is analogous.

(i) Inclusion $\mathscr{S}^\Psi_\sigma \subseteq \mathscr{S}^\Psi_\mathcal{I}$. Fix $S \in \mathscr{S}^\Psi_\sigma$. Then, there are predictable sets D_n increasing to $\Omega \times \mathbb{R}_+$ such that $(I_{D_n} \cdot S)^*_T \in L^\Psi$, for all $n \geq 1$. Thus there exist constants $c_n > 0$ such that $0 \leq E[\Psi(c_n(I_{D_n} \cdot S)^*_T)] < +\infty$. Since Ψ is nondecreasing over $\mathbb{R}_+$, c_n can be assumed $(0,1]$-valued. Let

$$b_n := E[\Psi(c_n(I_{D_n} \cdot S)^*_T)], \qquad d_n := h\, 2^{-n}(1+b_n)^{-1},$$

where $h := 1/(\sum_{n\geq 1} 2^{-n}(1+b_n)^{-1})$ is a normalizing constant, and define the following strictly positive, finite valued process

$$\varphi := \sum_{n\geq 1} c_n d_n I_{D_n}.$$

Since $0 \leq \varphi_m := \sum_{n=1}^m c_n d_n I_{D_n} \uparrow \varphi \leq \sum_{n\geq 1} d_n = 1$, the Dominated Convergence Theorem for stochastic integrals ([Pr05, Theorem 32]) applies. Therefore, $\varphi \in L(S)$ and $(\varphi_m \cdot S - \varphi \cdot S)^*_T$ tends to 0 in probability. Passing to a subsequence if necessary, we can assume the convergence holds P-a.s. Now,

$$(\varphi\cdot S)^*_T \leq (\varphi\cdot S - \varphi_m\cdot S)^*_T + (\varphi_m\cdot S)^*_T \leq (\varphi\cdot S - \varphi_m \cdot S)^*_T + \sum_{n=1}^m c_n d_n (I_{D_n}\cdot S)^*_T$$

and taking the limit on m, $(\varphi \cdot S)^*_T \leq \sum_{n\geq 1} c_n d_n (I_{D_n} \cdot S)^*_T$. Monotonicity of Ψ then ensures

$$E[\Psi((\varphi \cdot S)^*_T)] \leq E[\Psi(\sum_{n\geq 1} c_n d_n (I_{D_n} \cdot S)^*_T)].$$

Countable convexity of Ψ (Lemma 6.1) implies the latter term is majorized by $\sum_{n\geq 1} d_n E[\Psi(c_n(I_{D_n} \cdot S)^*_T)]$ and thus

$$\begin{aligned} E[\Psi((\varphi \cdot S)^*_T)] &\leq \sum_{n\geq 1} d_n E[\Psi(c_n(I_{D_n} \cdot S)^*_T)] \\ &= \sum_{n\geq 1} d_n b_n = h \sum_{n\geq 1} 2^{-n} \frac{b_n}{1+b_n} \leq h \leq 2(1+b_1), \end{aligned}$$

i.e. $S \in \mathscr{S}^\Psi_\mathcal{I}$.

(ii) Inclusion $\mathscr{S}_{\mathcal{I}}^{\Psi} \subseteq \mathscr{S}_{\sigma}^{\Psi}$. The line of the proof is: (a) fix $S \in \mathscr{S}_{\mathcal{I}}^{\Psi}$ and show $S \in (\mathscr{S}_{\mathrm{loc}}^{\Psi})_{\sigma}$; (b) then, as $\mathscr{S}^{\Psi}$ is stable under stopping (see Remark 2.5), a result by Kallsen ([Ka04, Lemma 2.1]) ensures $(\mathscr{S}_{\mathrm{loc}}^{\Psi})_{\sigma} = \mathscr{S}_{\sigma}^{\Psi}$, whence the conclusion follows.

We only need to prove (a), so let us fix $S \in \mathscr{S}_{\mathcal{I}}^{\Psi}$ and pick $\varphi > 0$ such that $\varphi \cdot S \in \mathscr{S}^{\Psi}$. By construction $D_n := \{\frac{1}{n} < \varphi < n\}$ is a sequence of predictable sets increasing to $\Omega \times \mathbb{R}_+$. We now show $I_{D_n} \cdot S \in \mathscr{S}_{\mathrm{loc}}^{\Psi}$ for all n. To this end, let $\tau_k^n = \inf\{t \mid (I_{D_n} \cdot S)_t^* > k\}$. Then

$$\begin{aligned}(I_{D_n} \cdot S^{\tau_k^n})_T^* &\le (I_{D_n} \cdot S^{\tau_k^n})_{T-}^* + |(I_{D_n} \cdot S^{\tau_k^n})_T| \\ &\le k + |(I_{D_n} \cdot S^{\tau_k^n})_{T-}| + |\Delta(I_{D_n} \cdot S^{\tau_k^n})_T| \le 2k + |\Delta(I_{D_n} \cdot S^{\tau_k^n})_T|,\end{aligned}$$

and the last jump term verifies

$$|\Delta(I_{D_n} \cdot S^{\tau_k^n})_T| = |\Delta((I_{D_n}/\varphi) \cdot (\varphi \cdot S^{\tau_k^n}))_T| = \left(\frac{I_{D_n}}{\varphi}\right)_T |\Delta(\varphi \cdot S^{\tau_k^n})_T| \le n 2(\varphi \cdot S)_T^*,$$

so that

$$(I_{D_n} \cdot S^{\tau_k^n})_T^* \le 2k + 2n(\varphi \cdot S)_T^* \in L^{\Psi}.$$

Therefore, for any fixed n, $(I_{D_n} \cdot S^{\tau_k^n})_T^*$ is also in L^{Ψ} for all k, whence $I_{D_n} \cdot S \in \mathscr{S}_{\mathrm{loc}}^{\Psi}$. This precisely means $S \in (\mathscr{S}_{\mathrm{loc}}^{\Psi})_{\sigma}$, which completes the proof. □

3. The strategies.

3.1. Conditions on S and simple strategies. Let S be a d-dimensional semimartingale which models the discounted evolution of d underlyings. As hinted in the introduction, to accommodate popular models for S, including exponential Lévy processes, we do not assume that S is locally bounded. However, to make sure that there is a sufficient number of well-behaved simple strategies we impose the following condition on S:

Assumption 3.1. *$S \in \mathscr{S}_{\sigma}^{\widehat{U}}$.*

The class $\mathscr{S}_{\sigma}^{\widehat{U}}$ introduced here appears to be the most comprehensive class of price processes to have been systematically studied in the context of utility maximization to date. Most papers in the literature assume S locally bounded, in our notation $S \in \mathscr{S}_{\mathrm{loc}}^{\infty}$. Sigma-bounded semimartingales, that is processes in $\mathscr{S}_{\sigma}^{\infty}$, appear in Kramkov and Sîrbu [KSî06]. For $p \in (1, +\infty)$ it can be shown, cf. [ČK07, Lemma A.2], that the class of semimartingales which are *locally in L^p* coincides with $\mathscr{S}_{\mathrm{loc}}^{p}$. These processes feature in Delbaen and Schachermayer [DS96]. Biagini and Frittelli [BF05] require existence of a suitable and compatible loss control for process S which in our notation corresponds to $S \in \mathscr{S}_{\mathcal{I}}^{M^{\widehat{U}}}$. In [BF08] this requirement is weakened to $S \in \mathscr{S}_{\mathcal{I}}^{\widehat{U}}$ which by Proposition 2.6 is equivalent to Assumption 3.1.

As has already been pointed out in [BF08], the σ-localization in Assumption 3.1 provides a substantial amount of flexibility since there are many interesting cases with $S \notin \mathscr{S}_{\mathrm{loc}}^{\infty}$ which fit in this setup. However, the cost of considering price processes of increasing generality is reflected in progressively less attractive interpretations of simple trading strategies:

Definition 3.2. *Define $\varphi \in \cap_{i=1}^{d} L(S^i; P)$, $\varphi > 0$, and a sequence of stopping times $(\tau_n)_n$ as follows:*

(i) *For $S \in \mathscr{S}^{\widehat{U}}$ let $\varphi \equiv 1$, $\tau_n \equiv T$ for all n;*

(ii) *For* $S \in \mathscr{S}^{\widehat{U}}_{\mathrm{loc}} \setminus \mathscr{S}^{\widehat{U}}$ *let* $\varphi \equiv 1$ *and let* $(\tau_n)_n$ *be a localizing sequence for* S *from the definition of* $\mathscr{S}^{\widehat{U}}_{\mathrm{loc}}$*;*

(iii) *For* $S \in \mathscr{S}^{\widehat{U}}_{\sigma} \setminus \mathscr{S}^{\widehat{U}}_{\mathrm{loc}}$ *let* $\tau_n \equiv T$ *and let* φ *be a fixed* $\mathcal{I}$*-localizing integrand for* S *such that* $\varphi \cdot S \in \mathscr{S}^{\widehat{U}}$*, which is possible by virtue of Proposition 2.6.*

We say H *is a* simple integrand *if it is of the form* $H = \sum_{k=1}^{N} H_k I_{]T_{k-1},T_k]}\,\varphi$ *where* $T_1 \leq \cdots \leq T_N$ *is a finite sequence of stopping times with* T_N *dominated by* τ_n *for some* n*, and each* H_k *is an* $\mathbb{R}^d$*-valued random variable,* $\mathcal{F}_{T_{k-1}}$*-measurable and bounded. The vector space of all simple integrands is denoted by* $\mathcal{H}$*.*

As can be seen from the definition, when $S \in \mathscr{S}^{\widehat{U}}$ no localization is needed. Every simple integrand is *simple* also in the sense of integration theory and it represents a buy-and-hold strategy on S over finitely many trading dates. Vice versa, every buy-and-hold strategy implemented over a finite set of dates is simple. One may thus wonder which models fall in this category. Some common examples are:

(a) discrete time models satisfying $|S_t| \in L^{\widehat{U}}$ for $t = 1, 2, \ldots, T$;

(b) Lévy processes, when (i) the utility U is exponential and the Lévy measure ν satisfies

$$\int e^{\lambda|x|} I_{\{|x|>1\}} d\nu(x) < +\infty \text{ for some } \lambda > 0;$$

or (ii) the utility $U(x)$ behaves asymptotically like $-|x|^p, p > 1$ when $x \to -\infty$ and the Lévy measure ν satisfies

$$\int |x|^p I_{\{|x|>1\}} d\nu(x) < +\infty.$$

Such conditions on ν are equivalent to integrability conditions on the maximal functional S^*, i.e. $S \in \mathscr{S}^{\widehat{U}}$, which in turn are equivalent to $\widehat{U}$-integrability of S_t at *some* $t > 0$. This follows from general results on g-moments of Lévy processes, when g is a submultiplicative function (see [Sat99, Theorems 25.3 and 25.18]). Explicit examples of utility maximization in this case can be found in Biagini and Frittelli [BF05, §3.2], [BF08, Example 35]. Here, U is exponential utility and S is a compound Poisson process with Gaussian or doubly exponentially distributed jumps;

(c) exponential Lévy processes belong to $\mathscr{S}^{\widehat{U}}$ whenever $\widehat{U}$ behaves asymptotically like a power function with exponent $p \in (1, +\infty)$ and the Lévy measure of $\ln S$, ν, satisfies

$$\int e^{px} I_{\{x>1\}} d\nu(x) < +\infty.$$

This is derived similarly as in (b) once $\ln S$ has been decomposed into a sum of two independent Lévy processes, one of which represents large jumps of $\ln S$.

For $S \in \mathscr{S}^{\widehat{U}}_{\mathrm{loc}} \setminus \mathscr{S}^{\widehat{U}}$ it is still true that all simple strategies are of the buy-and-hold type but one can no longer pick the trading dates arbitrarily. From a practical point of view most commonly used price processes fall into this category. For example, in the Black-Scholes model the risky asset is represented by a geometric Brownian motion which does not belong to $\mathscr{S}^{\widehat{U}}$ when U stands for the exponential utility. On the other hand S is continuous and therefore locally bounded which means $S \in \mathscr{S}^{\infty}_{\mathrm{loc}} \subseteq \mathscr{S}^{\widehat{U}}_{\mathrm{loc}} \subseteq \mathscr{S}^{\widehat{U}}_{\sigma}$ for any utility function satisfying our assumptions, including the exponential. The same line of reasoning applies to diffusions and more generally to all semimartingales with bounded jumps which therefore automatically belong to $\mathscr{S}_\sigma^{\widehat{U}}$ for *any* utility function U. In the special case $\mathscr{S}_{\rm loc}^{\widehat{U}} = \mathscr{S}_{\rm loc}^{p}$ our definition of simple strategies mirrors the definition in Delbaen and Schachermayer [DS96].

Finally, the price paid for allowing $S \in \mathscr{S}_\sigma^{\widehat{U}} \setminus \mathscr{S}_{\rm loc}^{\widehat{U}}$ is that simple strategies can no longer be interpreted as buy-and-hold with respect to the original price process S but only with respect to the better-behaved process $S' := \varphi \cdot S$. This case is interesting mainly theoretically since the $\mathcal{I}$-localizing strategy φ has already appeared in the literature on utility maximization. It plays an important role in the work of Biagini [Bia04] where the maximal process $(\varphi \cdot S)^*$ is taken as a dynamic loss control for the strategies in the utility maximization problem. Within setups of increasing generality in Biagini and Frittelli [BF05, BF08] φ gives rise to so-called *suitable* and *(weakly) compatible* loss control variables $W := (\varphi \cdot S)^*_T$.

3.2. σ-martingale measures. To motivate the definition of simple strategies mathematically we now define dual asset pricing measures.

DEFINITION 3.3. *$Q \ll P$ is a σ-martingale measure for S iff S is a σ-martingale under Q. The set of all σ-martingales measures for S is denoted by $\mathcal{M}$ and the subset of equivalent measures by $\mathcal{M}^e$.*

The concept of σ-martingale measure was introduced to Mathematical Finance by Delbaen and Schachermayer [DS98]. When S is (locally) bounded, it can be shown that $\mathcal{M}$ coincides with the absolutely continuous (local) martingale measures for S (see e.g. Protter [Pr05, Theorem 91]). Therefore, σ-martingales are a natural generalization of local martingales in the case when S is not locally bounded and the elements of $\mathcal{M}$ which are equivalent to P can be used as arbitrage-free pricing measures for the derivative securities whose payoff depends on S. The recent book [DS05] contains an extensive treatment of the financial applications of this mathematical concept.

When $S \in \mathscr{S}_\sigma^{\widehat{U}} \setminus \mathscr{S}^{\widehat{U}}$, one may wonder to what extent the utility maximization problem depends on the particular choice of φ (or of the localizing sequence $(\tau_n)_n$). Thanks to Émery's equality (2.10) the set of absolutely continuous σ-martingale measures for S is the same as the set of σ-martingale measures for $S' = \varphi \cdot S$. Specifically, $Q \ll P$ is a σ-martingale measure for S by (2.10) if and only if there exists a Q-positive, predictable process $\psi_Q \in \cap_{i=1}^d L(S^i; Q)$ such that $\psi_Q \cdot S$ is a Q-martingale. And this happens if and only if $\psi'_Q \cdot (\varphi \cdot S)$ is a Q-martingale, where $\psi'_Q = \frac{\psi_Q}{\varphi}$.

Since the sets of σ-martingale measures for S and S' are the same, the *dual* problem to the utility maximization also remains the same. Under suitable conditions (see the statement of the main Theorem 4.10), we thus end up with the same optimizer, regardless of a specific choice of the $\mathcal{I}$-localizing strategy φ.

3.3. Generalized relative entropy and properties of simple integrals.
DEFINITION 3.4. *A probability Q has finite generalized relative entropy with respect to P, notation: $Q \in P_V$, if there is $y_Q > 0$ such that*

$$v_Q(y_Q) := E\left[V\left(y_Q \frac{dQ}{dP}\right)\right] < \infty. \tag{3.1}$$

For exponential utility $U(x) = 1 - e^{-x}$ we have seen in (2.4) that $V(y) = y \ln y - y + 1$, and in this case a probability Q verifies (3.1) if and only if its probability density has finite Kullback-Leibler [KL51] divergence:

$$H(Q\|P) := E\left[\frac{dQ}{dP}\ln\frac{dQ}{dP}\right] < +\infty.$$

The Kullback-Leibler divergence is also known in Information Theory as *relative entropy* of Q with respect to P. Intuitively speaking, $H(Q\|P)$ is a non-symmetric measure of the distance between probabilities Q and P. In Financial Economics it measures the extra amount of wealth an agent with exponential utility perceives to have if she invests optimally in a complete market with pricing measure Q, as opposed to investing all her wealth in the risk-free asset.

In the 1960-ies, Csiszár treated a wide class of statistical distances replacing the weighting function $y\ln y$ by a convex function V verifying $V(1) = 0$. In his terminology, Q has finite V-divergence with respect to P if

$$E\left[V\left(\frac{dQ}{dP}\right)\right] < +\infty. \tag{3.2}$$

The interested reader can also consult Liese and Vajda [LV87].

In Mathematical Finance applications the function V is typically the convex conjugate of a utility function, see Kramkov and Schachermayer [KS99], Bellini and Frittelli [BeF02], Goll and Rüschendorf [GR01] and basically all the contemporary literature on utility maximization. Here, a $Q \in P_V$ is said to have *finite generalized relative entropy.* Our definition pushes the generalization one step further, since we do not require $y_Q = 1$ in (3.1).

The proof of the following simple Lemma is omitted.

Lemma 3.5. *Consider $Q_i \ll P$, $i = 1,2$, such that $v_{Q_i}(y_i) < +\infty$ for some $y_i > 0$. Then for $0 \le \lambda \le 1$*

$$v_{\lambda Q_1+(1-\lambda)Q_2}\left(\frac{1}{\lambda/y_1 + (1-\lambda)/y_2}\right) < \infty.$$

Corollary 3.6. *P_V is convex.*

Simple integrals have good mathematical properties with respect to σ-martingale measures with finite generalized relative entropy.

Lemma 3.7. *The wealth process $X = H\cdot S$ of every $H \in \mathcal{H}$ is a uniformly integrable martingale under all $Q \in \mathcal{M}\cap P_V$.*

Proof. (i) $S \in \mathscr{S}_\sigma^{\widehat{U}} \setminus \mathscr{S}_{\mathrm{loc}}^{\widehat{U}}$. Since $H \in \mathcal{H}$, the maximal functional X^* verifies $X_T^* \le c(\varphi\cdot S)_T^*$ for some constant $c > 0$ and some $\mathcal{I}$-localizing integrand φ which exists by Proposition 2.6. By (2.7) then $E[U(-\alpha(\varphi\cdot S)_T^*)] \in \mathbb{R}$ for some constant $\alpha > 0$ and, as a consequence,

$$0 \ge E\bigg[U\bigg(-\frac{\alpha}{c}X_T^*\bigg)\bigg] > -\infty.$$

For any fixed $Q \in \mathcal{M}\cap P_V$, the Fenchel inequality $U(x) - xy \le V(y)$ applied with $x = -\frac{\alpha}{c}X_T^*, y = y_Q\frac{dQ}{dP}$ gives

$$U\bigg(-\frac{\alpha}{c}X_T^*\bigg) + \frac{\alpha}{c}X_T^* y_Q\frac{dQ}{dP} \le V\bigg(y_Q\frac{dQ}{dP}\bigg),$$

whence $0 \leq \frac{\alpha}{c} y_Q X_T^* \frac{dQ}{dP} \leq V(y_Q \frac{dQ}{dP}) - U(-\frac{\alpha}{c} X_T^*)$, and therefore X_T^* is in $L^1(Q)$. As Q is a σ-martingale probability for S, X is also a Q-σ-martingale. Since its maximal process is integrable, X is in fact a Q-uniformly integrable martingale (see Protter [Pr05, Chapter IV-9]).

(ii) $S \in \mathscr{S}^{\widehat{U}}_{\mathrm{loc}}$. Proceed as in (i), replacing φ with $I_{[0,\tau_n]}$. $\square$

In financial terms, the message of the above Lemma is that each $Q \in \mathcal{M} \cap P_V$ represents a pricing rule that assigns a correct price to every simple self-financing strategy.

3.4. Admissible integrands and integrals. As anticipated in the introduction, simple integrands are unlikely to contain the solution of the utility maximization problem. The appropriate class of admissible integrands is an extension given in terms of suitable limits of strategies in $\mathcal{H}$. We recall the definition of admissibility here for convenience.

DEFINITION 1.1. *$H \in L(S)$ is an* admissible integrand *if $U(H \cdot S_T) \in L^1(P)$ and if there exists an approximating sequence $(H^n)_n$ in $\mathcal{H}$ such that:*

(i) *$H^n \cdot S_t \to H \cdot S_t$ in probability for all $t \in [0,T]$;*

(ii) *$U(H^n \cdot S_T) \to U(H \cdot S_T)$ in $L^1(P)$.*

The set of all admissible integrands is denoted by $\overline{\mathcal{H}}$.

While for $H \in \mathcal{H}$ the wealth process $H \cdot S$ is always a martingale under $Q \in \mathcal{M} \cap P_V$ due to Lemma 3.7, the following result shows that $\overline{\mathcal{H}}$ is a subset of the supermartingale class of strategies $\mathcal{H}^{\mathrm{s}}$ introduced by [Sch03],

$$\mathcal{H}^{\mathrm{s}} := \{H \in L(S) \mid H \cdot S \text{ is a local martingale} \\ \text{and a supermartingale under any } Q \in \mathcal{M} \cap P_V\}. \tag{3.3}$$

PROPOSITION 3.8. *$\overline{\mathcal{H}} \subseteq \mathcal{H}^{\mathrm{s}}$.*

Proof. Let $X = H \cdot S$ for some $H \in \overline{\mathcal{H}}$ and let $(X^n := H^n \cdot S)_n$ with $H^n \in \mathcal{H}$ be an approximating sequence. Fix a $Q \in \mathcal{M} \cap P_V$ and a corresponding scaling y_Q as in Definition 3.4. Item (i) of Definition 1.1 applied at time T implies $(X_T^n)^-$ converges in P-probability to X_T^-. Moreover, Fenchel inequality gives

$$U(X_T^n) - V(y_Q \frac{dQ}{dP}) \leq X_T^n \, y_Q \frac{dQ}{dP}.$$

From Definition 1.1, item (ii), the left hand side above converges in $L^1(P)$, whence the family $(Y^n)_n, Y^n := (X_T^n)^- \frac{dQ}{dP}$ is P-uniformly integrable, so $((X_T^n)^-)_n$ is Q-uniformly integrable (see Lemma 6.2). Uniform integrability plus convergence in probability ensures $(X_T^n)^- \to X_T^-$ in $L^1(Q)$. By passing to a subsequence if necessary, the next is an integrable lower bound for $(X_T^n)_n$,

$$W^Q := \sum_n |(X_T^{n+1})^- - (X_T^n)^-| \in L^1(Q)$$

Denote by Z^Q the associated Q-martingale, $Z_t^Q := E_Q[W^Q \mid \mathcal{F}_t]$. Note that when $\operatorname{dom} U$ is a half-line we could also have chosen trivially $W^Q := -\inf \operatorname{dom} U$.

Since $X_T^n \geq -W^Q$ and process X^n is a Q-martingale for all n by Lemma 3.7, we obtain

$$X_t^n = E_Q[X_T^n \mid \mathcal{F}_t] \geq -E_Q[W^Q \mid \mathcal{F}_t] = -Z_t^Q, \tag{3.4}$$

so that the sequence X^n is controlled from below by the Q-martingale Z^Q. Therefore by Delbaen and Schachermayer compactness result [DS99, Theorem D] (in the version stated in §5, [DS98]) there exists a limit càdlàg supermartingale $\widetilde{V}$ to which a sequence $K^n \cdot S$, where K^n is a suitable convex combination of tails $K^n \in \operatorname{conv}(H^n, H^{n+1}, \ldots)$, converges Q-almost surely for every rational time $0 \le q \le T$. By item (i), $(X_t^n)_n$ converges in P-probability to X_t for every t, thus $K^n \cdot S_t$ converges to X_t for every t as well. Therefore $\widetilde{V}$ coincides Q-a.s. with X on rational times, and since X is also càdlàg as it is an integral, X and $\widetilde{V}$ are indistinguishable, so that X is a Q-supermartingale. By assumption Q is a σ-martingale measure, so $X = H \cdot S = (\frac{1}{\varphi} H) \cdot (\varphi \cdot S)$ where $\varphi > 0$ and $\varphi \cdot S$ is a Q-martingale. As X also satisfies $X \ge -Z^Q$, Ansel and Stricker lemma [AS94, Corollaire 3.5] implies that X is a local Q-martingale. ∎

REMARK 3.9. *Proposition 3.8 would go through if one replaced our class $\mathcal{H}$ with the set of integrands with wealth bounded from below*

$$\mathcal{H}^b = \{H \in L(S) \mid H \cdot S \ge c \text{ for some } c \in \mathbb{R}\}, \tag{3.5}$$

as in Schachermayer [Sch01] when $S \in \mathscr{S}^\infty_{\mathrm{loc}}$, or more generally with the larger set of strategies whose losses are in some sense well controlled as in Biagini and Frittelli [BF05, BF08],

$$\mathcal{H}^{\hat{U}} = \{H \in L(S) \mid \exists W \ge 0, E[U(-W)] > -\infty, H \cdot S \ge -W\}, \tag{3.6}$$

see also Biagini and Sîrbu [BS09]. An application of the Ansel and Stricker lemma [AS94, Corollaire 3.5] shows that wealth processes for strategies in $\mathcal{H}^{\hat{U}} \supseteq \mathcal{H}^b$ are local martingales and supermartingales under any $Q \in \mathcal{M} \cap P_V$ – but not martingales in general. In contrast, our smaller class $\mathcal{H}$ has the stronger martingale property as shown in Proposition 3.8. Mathematically, however, it is the supermartingale property of approximating strategies that really matters. This is also true in the proof of the main Theorem 4.10 where one can replace arguments relying on the martingale property of approximating strategies [Yor78, Corollaire 2.5.2] with supermartingale compactness results of [DS99].

The list below summarizes the advantages of $\overline{\mathcal{H}}$ over current definitions of admissibility:

(a) Definition 1.1 is *primal.* No pricing measures come into play, and admissibility can thus be checked under P.
(b) The present definition is *dynamic*, that is the whole wealth process, rather than just its terminal value, is involved in the definition of $\overline{\mathcal{H}}$. As a result *all* admissible strategies are in the supermartingale class.
(c) The loss controls required in the proof of the supermartingale property are generated endogenously, via approximating sequences. This provides a great deal of flexibility and ensures that for U finite on $\mathbb{R}$ the optimizer is in $\overline{\mathcal{H}}$ under very mild conditions, milder than the conditions assumed to obtain the supermartingale property of the optimizer in [Sch03, BF07]. Since under our assumptions the optimal utilities over $\overline{\mathcal{H}}$ and $\mathcal{H}^{\mathrm{s}}$ coincide, see (4.17), the smaller class $\overline{\mathcal{H}}$ seems to be more appropriate than $\mathcal{H}^{\mathrm{s}}$ not only economically but also mathematically.
(d) Approximation by strategies in $\mathcal{H}$ is built into the definition of admissibility, it does not have to be deduced separately (cf. [St03]).
(e) The desirable properties above hold without any technical assumptions on U. It can be finite on $\mathbb{R}$ or only on a half-line; bounded from above or not, or even truncated; neither strict monotonicity, strict convexity nor differentiability are required.

(f) Our definition is compatible with the existing definition of admissibility for non-monotone quadratic preferences, see Remark 3.10 below. We have therefore found a good notion of admissibility which encompasses both the classical mean-variance preferences *and* monotone expected utility.

REMARK 3.10. *For the purpose of this remark only, we admit non-monotone* U*. Specifically, let* $U(x) := x - x^2/2$*, which represents a normalized quadratic utility. In such case,* $H \in \overline{\mathcal{H}}$ *if and only if there is a sequence of* $H^n \in \mathcal{H}$ *such that:*

(a) $H^n \cdot S_t \to H \cdot S_t$ *in probability for all* $t \in [0,T]$*, and*

(b) $H^n \cdot S_T \to H \cdot S_T$ *in* $L^2(P)$*.*

In other words, when U *is quadratic the admissibility criterion in Definition 1.1 coincides with the notion of admissibility pioneered by Jan Kallsen in [ČK07, Definition 2.2], which inspired our work.*

Proof. Since (a) above and (i) in Definition 1.1 coincide, the only thing to prove is that (ii) in our definition is equivalent to (b) above:

$\Rightarrow$ Suppose first $H \in \overline{\mathcal{H}}$. The $L^1(P)$ convergence of utilities implies $E[U(X^n_T)] \to E[U(X_T)]$ so that X^n_T are uniformly bounded in $L^2(P)$. Since $L^2(P)$ is a reflexive space there is a sequence of convex combinations of tails $(X^k_T)_{k\geq n}$, say $\widetilde{X}^n_T$, which converges in $L^2(P)$ to a square integrable random variable which necessarily is $X_T = H \cdot S_T$ thanks to Definition 1.1, item (i). By considering the corresponding convex combinations of strategies, which are again simple, we obtain the existence of an approximating sequence à la Kallsen for H.

$\Leftarrow$ Conversely, let $X = H \cdot S$ be an integral approximated à la Kallsen by simple integrals $(X^n)_n$. $L^1(P)$ convergence of the utilities $U(X^n_T)$ to $U(X_T)$ is then a consequence of the Cauchy-Schwartz inequality. □

4. Optimal trading strategy is in $\overline{\mathcal{H}}$. The optimal investment problem can be formulated over $\mathcal{H}$, $\overline{\mathcal{H}}$ or over $\mathcal{H}^{\mathrm{s}}$, respectively,

$$u_{\mathcal{H}}(x) := \sup_{H\in\mathcal{H}} E[U(x + H \cdot S_T)], \tag{4.1}$$

$$u_{\overline{\mathcal{H}}}(x) := \sup_{H\in\overline{\mathcal{H}}} E[U(x + H \cdot S_T)], \tag{4.2}$$

$$u_{\mathcal{H}^{\mathrm{s}}}(x) := \sup_{H\in\mathcal{H}^{\mathrm{s}}} E[U(x + H \cdot S_T)]. \tag{4.3}$$

Alongside, we consider auxiliary complete market utility maximization problems, each obtained by fixing an arbitrary $Q \in \mathcal{M} \cap P_V$:

$$u_Q(x) := \sup_{X\in L^1(Q), E_Q[X]\leq x} E[U(X)]. \tag{4.4}$$

The value functions $u_{\mathcal{H}}(x), u_{\overline{\mathcal{H}}}(x), u_{\mathcal{H}^{\mathrm{s}}}(x), u_Q(x)$ are also known as indirect utilities (from the respective domains of maximization). The next lemma is an easy consequence of the definition of $\overline{\mathcal{H}}$ and of the supermartingale property of the strategies in $\overline{\mathcal{H}}$ and $\mathcal{H}^{\mathrm{s}}$. The proof is omitted.

LEMMA 4.1. *For any* $x > \underline{x}$ *and for any* $Q \in \mathcal{M} \cap P_V$

$$u_{\mathcal{H}}(x) = u_{\overline{\mathcal{H}}}(x) \leq u_{\mathcal{H}^{\mathrm{s}}}(x) \leq u_Q(x). \tag{4.5}$$

4.1. Reasonable Asymptotic Elasticity and Inada conditions. It is well known in the literature that the existence of an optimizer is not guaranteed yet, neither in $\overline{\mathcal{H}}$ nor in the larger supermartingale class $\mathcal{H}^s \supseteq \overline{\mathcal{H}}$. An additional condition has to be imposed, essentially to ensure that the expected utility functional $k \mapsto E[U(k)]$ is upper semicontinuous with respect to some weak topology on terminal wealths.

Kramkov and Schachermayer were the first to address this issue in [KS99, Sch01] for regular U, that is utilities that are strictly increasing, strictly concave and differentiable in the interior of their effective domain. To the end of recovering an optimizer they introduced the celebrated Reasonable Asymptotic Elasticity condition on U (RAE(U)),

$$\limsup_{x\to+\infty} \frac{xU'(x)}{U(x)} < 1, \tag{4.6}$$

$$\text{and also} \quad \liminf_{x\to-\infty} \frac{xU'(x)}{U(x)} > 1, \ \text{when } U \text{ is finite on } \mathbb{R}, \tag{4.7}$$

as a necessary and sufficient condition to be imposed on the utility U *only*, regardless of the probabilistic model. This condition is now very popular, see [OŽ09, RS05, Sch03, B02] to mention just a few contributions.

In subsequent work, in the context of utilities finite on $\mathbb{R}_+$, Kramkov and Schachermayer [KS03] put forward less restrictive conditions[1], *imposed jointly on the model and on the preferences*, in order to recover the optimal terminal wealth. Here they work under assumptions which are equivalent to the existence of $Q \in \mathcal{M}^e \cap P_V$ and the following Inada condition on the indirect utility $u_{\mathcal{H}^b}$, where the class $\mathcal{H}^b$ is defined in (3.5):

$$\lim_{x\to+\infty} u_{\mathcal{H}^b}(x)/x = 0. \tag{4.8}$$

It is important to note that for utility functions finite on a half-line the modulus of the conjugate function $V(y)$ grows only linearly for large y and therefore the following implication holds automatically:

$$Q \in \mathcal{M} \cap P_V \Rightarrow v_Q(y) < +\infty \text{ for all } y \text{ sufficiently high.} \tag{4.9}$$

On the other hand, for utilities finite on $\mathbb{R}$ condition (4.9) has to be imposed explicitly, together with an appropriate generalization of condition (4.8).

Assumption 4.2. *Condition* (4.9) *is satisfied and*

$$\textit{there exists } Q \in \mathcal{M} \cap P_V \textit{ such that } \lim_{x\to+\infty} u_Q(x)/x = 0. \tag{4.10}$$

Note first that the requirement (4.10) automatically holds, and for all $Q \in \mathcal{M} \cap P_V$, if $U(+\infty) < +\infty$. Since for any $Q \in \mathcal{M} \cap P_V$ one has $u_Q(x) \geq u_{\mathcal{H}}(x) \geq U(x)$, and U is monotone, condition (4.10) implies an identical Inada condition both on the indirect utility $u_{\mathcal{H}}$ and also on the original utility function U at $+\infty$. An identical chain of inequalities for the indirect utilities holds if we replace $\mathcal{H}$ with $\mathcal{H}^b$ and for this reason condition (4.10) is slightly stronger than the condition (4.8) imposed in

[1] The interested reader is referred also to the recent Biagini and Guasoni [BG09] for counterexamples and a different, *relaxed* framework that allows optimal terminal wealth to be a measure and not only a random variable.

[KS03] when U is finite on a half-line. It is an open question whether condition (4.10) can be weakened to

$$\mathcal{M} \cap P_V \neq \emptyset \text{ and } \lim_{x\to+\infty} u_{\mathcal{H}}(x)/x = 0. \tag{4.11}$$

Further discussion of Assumption 4.2 and its relation to RAE(U) and the Inada condition (4.8) can be found in §5.1. The results of the next section go in that direction.

4.2. Complete market duality. Here we study a complete market $Q \in P_V$ and hence no specific model for S is required. Among other results, we provide an alternative characterization of the Inada condition (4.10) in terms of the generalized relative entropy of Q.

LEMMA 4.3. *Fix $Q \in P_V$ and consider the function v_Q defined in* (3.1). *For any $x > \underline{x}$,*

$$u_Q(x) = \min_{y\geq 0}\{xy + v_Q(y)\} < +\infty. \tag{4.12}$$

An Orlicz duality based proof of the above lemma is given in §6. Here we only remark that the minimizer may not be unique. This is due to lack of strict convexity of V, which in turn is due to lack of strict concavity of U.

COROLLARY 4.4. *Fix $Q \in P_V$. The following statements are equivalent:*

(i) *u_Q verifies the Inada condition at $+\infty$: $\lim_{x\to+\infty} u_Q(x)/x = 0$;*

(ii) *there is $y_Q > 0$ such that*

$$v_Q(y) = E\left[V\left(y\frac{dQ}{dP}\right)\right] < +\infty \textit{ for all } y \in (0, y_Q]. \tag{4.13}$$

Proof. (ii)⇒(i) Suppose that $v_Q(y)$ is finite in a right neighborhood of 0. By Fenchel inequality, $E[U(X)] - E[y\frac{dQ}{dP}X] \leq E[V(y\frac{dQ}{dP})]$ for all $X \in L^1(Q)$ so that $u_Q(x) \leq xy + v_Q(y)$ for all $y > 0$. Fixing y one obtains $\lim_{x\to+\infty} u_Q(x)/x \leq y$ and on letting $y \to 0$ the Inada condition on u_Q follows.

(i)⇒(ii) For a given $x > \underline{x}$, select one dual minimizer in (4.12) and denote it by y_x. Now, $u_Q(x) = xy_x + v_Q(y_x)$, $v_Q(y_x)$ is finite, and the chain of inequalities

$$u_Q(x) = xy_x + v_Q(y_x) \overset{Jensen}{\geq} xy_x + V(y_x) \geq xy_x \geq 0$$

holds for any x as V is nonnegative. Dividing by $x > 0$ and sending x to $+\infty$, (4.10) implies $\lim_{x\to+\infty} y_x = 0$. Finiteness of v_Q over the set $\{y_x\}_x$, whose closure contains 0, and convexity of v_Q finally imply v_Q is finite in the interval $(0, y_Q]$, with y_Q from (3.1). □

COROLLARY 4.5. *If $\mathcal{M}^e \cap P_V \neq \emptyset$ then the measure Q in* (4.10) *can be chosen equivalent to P.*

Proof. Take $Q^e \in \mathcal{M}^e \cap P_V$ and assume Q satisfies (4.10). By Corollary 4.4 $v_Q(y)$ is finite for all y near zero. Define

$$Q^* := \frac{1}{2}Q + \frac{1}{2}Q^e.$$

Thus, $Q^* \sim P$ and by Lemma 3.5 $v_{Q^*}(y)$ is finite for all y near zero. Therefore u_{Q^*} satisfies the Inada conditon (4.10). □

The next proposition contains a novel characterization of the condition

$$u_Q(x) < U(+\infty),$$

which is a kind of "no utility-based arbitrage" condition, when Q has finite generalized relative entropy. Agents cannot reach satiation utility $U(+\infty)$ if the initial capital x is below the satiation point $\overline{x}$, and vice versa.

PROPOSITION 4.6. *For $Q \in P_V$ and $x > \underline{x}$ the following statements are equivalent:*

(i) $x < \overline{x}$;

(ii)

$$u_Q(x) = \min_{y>0}\{xy + v_Q(y)\} < U(+\infty). \tag{4.14}$$

Proof. (ii)⇒(i) $U(x) \le u_Q(x) < U(+\infty)$ implies $x < \overline{x}$.
(i)⇒(ii) Let $Z := y_Q dQ/dP$, with y_Q from (3.1). When $U(+\infty) = V(0) = +\infty$ there is nothing to prove in view of (4.12). Consider therefore the remaining case $0 < U(+\infty) = V(0) < +\infty$. Function $f(y) := V(y) + xy$ is convex and by Rockafellar [R70, Theorem 23.5] it attains its minimum at $\widehat{y} := U'_-(x) > 0$ with $f(\widehat{y}) = V(\widehat{y}) + x\widehat{y} = U(x)$. Convexity then gives

$$f(y) \le f(0) - y\frac{f(0) - f(\widehat{y})}{\widehat{y}} = U(+\infty) - y\frac{U(+\infty) - U(x)}{\widehat{y}} \qquad \text{for } y \in [0, \widehat{y}],$$

$$f(y/k) \le f(0) + \frac{f(y) - f(0)}{k} \le U(+\infty) + \frac{f(y)}{k} \qquad \text{for } k \ge 1, y \ge 0.$$

For $k \ge 1$ these estimates imply

$$\begin{aligned} E[f(Z/k)] &= E[f(Z/k)1_{\{Z\le k\widehat{y}\}}] + E[f(Z/k)1_{\{Z>k\widehat{y}\}}] \\ &\le U(+\infty) - \frac{1}{k}\left(\frac{U(+\infty) - U(x)}{\widehat{y}}E[Z1_{\{Z\le k\widehat{y}\}}] - E[f(Z)1_{\{Z>k\widehat{y}\}}]\right), \end{aligned}$$

and, as $x < \overline{x}$ implies $U(x) < U(+\infty)$, for sufficiently large k $E[f(Z/k)] < U(+\infty) = V(0)$, which completes the proof. □

REMARK 4.7. *Corollary 4.4 and Proposition 4.6 should be contrasted with an example by Schachermayer [Sch01, Lemma 3.8], where the author constructs an arbitrage-free complete market with unique pricing measure Q for which $u_Q(x) \equiv U(+\infty)$, while U is strictly increasing and bounded from above (and therefore it satisfies the Inada condition at $+\infty$). This is possible because the measure Q in question does not belong to P_V.*

COROLLARY 4.8. *If $x \in (\underline{x}, \overline{x})$ then*

$$u_{\mathcal{H}}(x) \le u_{\overline{\mathcal{H}}}(x) \le u_{\mathcal{H}^s}(x) \le \inf_{Q\in\mathcal{M}\cap P_V} u_Q(x) = \inf_{y>0, Q\in\mathcal{M}\cap P_V}\{xy + v_Q(y)\}. \tag{4.15}$$

Proof. The inequalities follow from Lemma 4.1 and Proposition 4.6. □

4.3. The main result. The minimization problem on the right-hand side of (4.15) is a natural candidate as a dual problem to the utility maximization on the left-hand side. However, the general theory of [BF08] shows that in order to catch the minimizer the dual domain must be extended beyond probability densities. Rephrased in our terminology, whenever $S \in \mathscr{S}_\sigma^{\widehat{U}} \setminus \mathscr{S}_\sigma^{M^{\widehat{U}}}$ the dual problem may have a minimizer which has a non zero singular part, but for $S \in \mathscr{S}_\sigma^{M^{\widehat{U}}}$ the singular parts in the dual problem disappear and there is no duality gap in (4.15) under Assumption 4.2. We make these statements precise in Theorem 5.1 and Corollary 5.2.

Our main result hinges on the absence of singularities in the dual problem, which is what we now assume. Within the confines of Assumption 4.9, which can be imposed also when $S \in \mathscr{S}_\sigma^{\widehat{U}} \setminus \mathscr{S}_\sigma^{M^{\widehat{U}}}$, we provide a unified treatment for utility functions finite on $\mathbb{R}$ or only on a half-line.

Assumption 4.9. *For any $x \in (\underline{x}, \overline{x})$, the following dual relation holds:*

$$u_{\mathcal{H}}(x) = \min_{Q \in \mathcal{M} \cap P_V} u_Q(x) = \min_{y \geq 0, Q \in \mathcal{M} \cap P_V} \{xy + v_Q(y)\}. \tag{4.16}$$

As indicated above, this assumption represents no loss of generality for $S \in \mathscr{S}_\sigma^{M^{\widehat{U}}}$, including situations where U is finite on $\mathbb{R}$ and

(a) S is "sufficiently integrable". Some commonly found examples are locally bounded processes, such as diffusions or jump diffusions with bounded relative jumps, regardless of the specification of U; jump diffusions with relative jumps in $M^{\widehat{U}}$; Lévy processes with large jumps in $M^{\widehat{U}}$;

(b) $L^{\widehat{U}} = M^{\widehat{U}}$, under the standing Assumption 3.1. This happens when e.g. U has left tail that behaves asymptotically like a power, x^p, with $p > 1$.

When $S \in \mathscr{S}_\sigma^{\widehat{U}} \setminus \mathscr{S}_\sigma^{M^{\widehat{U}}}$, which includes all cases where U is finite only on a half-line, unfortunately there is no known sufficient condition for the strong duality (4.16) to hold. The appropriate modification of Theorem 4.10 which would work without Assumption 4.9 remains an interesting area for future research.

Assumption 4.9 together with (4.15) immediately yields the following, apparently stronger, statement for $x \in (\underline{x}, \overline{x})$

$$u_{\mathcal{H}}(x) = u_{\overline{\mathcal{H}}}(x) = u_{\mathcal{H}^s}(x) = \min_{y > 0, Q \in \mathcal{M} \cap P_V} \left\{ xy + E\left[V\left(y \frac{dQ}{dP} \right) \right] \right\}. \tag{4.17}$$

Any optimal dual pair in (4.17) is denoted by $(\widehat{y}, \widehat{Q})$, dependence on x is understood. The lack of uniqueness of the optimal dual pair is again due to the lack of strict convexity of V, stemming from the lack of strict concavity of U.

Most results in the literature are obtained under the assumption $\widehat{Q} \sim P$. This condition is satisfied automatically for utility functions unbounded from above since $V(0) = U(+\infty) = +\infty$ while $E[V(\widehat{y}\frac{d\widehat{Q}}{dP})]$ must be finite. When U is strictly monotone but bounded, a well-known sufficient condition for $\widehat{Q} \sim P$ is the existence of an equivalent σ-martingale measure with finite generalized relative entropy. This can be gleaned from (a.i) and (a.iii) in Theorem 4.10, on observing that $\overline{x} = +\infty$.

As a general comment, Theorem 4.10 provides a desirable approximation result for the optimal strategy $\widehat{H} \in \overline{\mathcal{H}}$. The approximation holds under very mild conditions: U may lack strict monotonicity and strict concavity; $S \in \mathscr{S}_\sigma^{\widehat{U}}$; and $\widehat{Q}$ may be only absolutely continuous with respect to P. These results are novel not only for utility finite on $\mathbb{R}$ but also for utility functions finite on a half-line.

For U finite on $\mathbb{R}$ our framework is a further improvement over the current literature: [Sch01], [KSt02], [St03], [OŽ09], [BTZ04] all assume S locally bounded. Approximation by simple strategies has so far been shown only for exponential utility, for locally bounded S and for expected utility only cf. [St03, Theorem 5] – not in the stronger sense of $L^1(P)$ convergence of the utilities given by item (ii) in Definition 1.1.

For comparison, Schachermayer [Sch01] proves an approximation similar to (4.19) for the terminal wealth of the optimal solution $\widehat{f} = \widehat{H} \cdot S_T$ via integrals bounded from below. This work is extended further by Bouchard et al. [BTZ04] who allow for non-differentiable and non-monotone utility functions. Moreover, in [Sch03] $\widehat{H}$ is shown to be in the supermartingale class of strategies through a (hard) contradiction argument, which is later extended by [BF07] to $S \in \mathscr{S}_{\mathcal{I}}^{M^{\widehat{U}}}$ with a proof along the same lines.

In the present paper the supermartingale property of $\widehat{H}$ is shown in a general setup and in a very natural way, as a consequence of $\overline{\mathcal{H}} \subseteq \mathcal{H}^{\mathrm{s}}$. We also extend results of Bouchard et al. [BTZ04] beyond $S \in \mathscr{S}^{\infty}_{\mathrm{loc}}$ under the weaker condition from Assumption 4.2 instead of the RAE(U) condition (5.1), while considerably simplifying the required proofs thanks to the Orlicz duality approach.

When U is not strictly monotone, that is when U attains its global maximum at a satiation point $\overline{x} < +\infty$, the sufficient conditions for $\widehat{Q} \sim P$ known in the monotone case do not work; here typically $\widehat{Q}$ is *not* equivalent to P even when there are equivalent probabilities in $\mathcal{M} \cap P_V$. We nonetheless recover an integral representation *under* P, and thus existence of an optimal trading strategy, provided the budget constraint is binding, $E_Q[\widehat{f}] = x$, for *some* $Q \in \mathcal{M}^e \cap P_V$. This mild sufficient condition appears to be new in the literature. Our contribution in the case where U is strictly monotone but $\widehat{Q}$ is not equivalent to P is discussed in detail in §5.3.

THEOREM 4.10. *Under Assumptions 3.1, 4.2 and 4.9, for any initial wealth $x \in (\underline{x}, \overline{x})$ the following statements hold:*

(a) *There exists a $(-\infty, +\infty]$-valued claim $\widehat{f}$, not unique in general, with the following properties*
   (i) *$\widehat{f} < +\infty$ whenever $\mathcal{M}^e \cap P_V \neq \emptyset$;*
   (ii) *$\widehat{f}$ realizes the optimal expected utility, in the sense that*
$$E[U(\widehat{f})] = u_{\mathcal{H}}(x);$$
   (iii) *$E_{\widehat{Q}}[\widehat{f}] = x$, and the following equalities hold P-a.s. for any dual optimizers $\widehat{y}, \widehat{Q}$:*
$$V\left(\widehat{y}\frac{d\widehat{Q}}{dP}\right) = U(\widehat{f}) - \widehat{f}\widehat{y}\frac{d\widehat{Q}}{dP},$$
$$\{\widehat{f} \geq \overline{x}\} = \left\{\frac{d\widehat{Q}}{dP} = 0\right\};$$
   (iv) *$\widehat{f} \in L^1(Q)$ and $E_Q[\widehat{f}] \leq x$ for all $Q \in \mathcal{M} \cap P_V$;*
   (v) *In case U is strictly concave, V is strictly convex and the solutions of primal and dual problem $\widehat{f}, \widehat{y}, \widehat{Q}$ are unique. If in addition U is differentiable, these unique solutions satisfy $\widehat{y}\frac{d\widehat{Q}}{dP} = U'(\widehat{f})$;*

(b) *There is an approximating sequence of strategies $H^n \in \mathcal{H}$ with terminal values $f^n := x + H^n \cdot S_T$ such that:*
   (i)
$$f_n \overset{P\text{-}a.s.}{\to} \widehat{f}, \tag{4.18}$$
   provided $\mathcal{M}^e \cap P_V \neq \emptyset$ or $\overline{x} = +\infty$;
   (ii)
$$U(f_n) \overset{L^1(P)}{\to} U(\widehat{f}); \tag{4.19}$$

(iii)

$$f_n \overset{L^1(\widehat{Q})}{\to} \widehat{f}, \tag{4.20}$$

and, provided (4.18) *holds, for any* $Q \in \mathcal{M} \cap P_V$ *such that* $E_Q[\widehat{f}] = x$

$$f_n \overset{L^1(Q)}{\to} \widehat{f}; \tag{4.21}$$

(iv) *There exists an integral representation* $\widehat{f} = x + \widehat{H} \cdot S_T$ *with* $\widehat{H} \in L(S; \widehat{Q})$, *and* $\widehat{H} \cdot S$ *is a* $\widehat{Q}$*-martingale.*

(v) *When there is* $\widetilde{Q} \in \mathcal{M}^e \cap P_V$ *such that* $E_{\widetilde{Q}}[\widehat{f}] = x$ *then* $\widehat{H}$ *in* (iv) *can be chosen in* $\overline{\mathcal{H}}$ *and consequently* $\widehat{H}$ *is a utility maximizer over both* $\overline{\mathcal{H}}$ *and* $\mathcal{H}^{\mathrm{s}}$,

$$u_{\mathcal{H}}(x) = u_{\mathcal{H}^{\mathrm{s}}}(x) = \max_{H \in \overline{\mathcal{H}}} E[U(x + H \cdot S_T)]. \tag{4.22}$$

In particular, by virtue of (a.iii), (4.22) *holds whenever* $\widehat{Q} \sim P$.

Proof.

(a) Let us fix a pair $\widehat{y}, \widehat{Q}$ of dual minimizers. For ease of notation and without loss of generality we let $x = 0$ throughout.

(i.1) Select a maximizing sequence $(k_n)_n, k_n = K^n \cdot S_T, K^n \in \mathcal{H}$ so that $E[U(k_n)] \uparrow u_{\mathcal{H}}(0)$. Fix $Q^* \in \mathcal{M} \cap P_V$ as follows:

- in case $\mathcal{M}^e \cap P_V \neq \emptyset$, select Q^* as an equivalent measure satisfying (4.10). This is possible by Corollary 4.5;
- in case $\mathcal{M}^e \cap P_V = \emptyset$, take $Q^* = \widehat{Q}$. Here necessarily $V(0) = U(+\infty) < +\infty$, so $\widehat{Q}$ as well as any other measure in $\mathcal{M} \cap P_V$ satisfies (4.10).

Let

$$\overline{Q} := \frac{1}{2}\widehat{Q} + \frac{1}{2}Q^*.$$

Then, $\overline{Q} \in \mathcal{M} \cap P_V$; $\overline{Q} \sim P$ if $\mathcal{M}^e \cap P_V \neq \emptyset$; $\overline{Q}$ satisfies (4.10); $L^1(\overline{Q}) = L^1(\widehat{Q}) \cap L^1(Q^*)$; and $L^1(\overline{Q})$-convergence is equivalent to convergence in $L^1(\widehat{Q})$ *and* $L^1(Q^*)$ by construction.

(i.2) The sequence $(k_n)_n$ is bounded in $L^1(\overline{Q})$. In a general case this follows from the auxiliary Proposition 6.3, which in turn is a consequence of the Inada condition (4.10). In a special case when $\operatorname{dom} U$ is a half-line, $L^1(\overline{Q})$-boundedness also follows trivially from $k_n \geq \underline{x}$ and $E_Q[k_n] = 0$, which is a consequence of Lemma 3.7. In a second special case where U is bounded from above the claim can be alternatively deduced from the boundedness of $U^-(f_n)$ and the Fenchel inequality (2.5).

(i.3) $L^1(\overline{Q})$ boundedness of $(k_n)_n$ enables the application of the Komlós theorem, so that there exists a sequence of convex combinations $(f_n)_n$ with $f_n \in \operatorname{conv}(k_n, k_{n+1}, \ldots)$, that converges $\overline{Q}$-a.s. to a certain random variable $f \in L^1(\overline{Q}) \subseteq L^1(\widehat{Q})$. As $\mathcal{H}$ is a vector space, these f_n are terminal values of simple integrals $f_n = H^n \cdot S_T$, $H^n \in \mathcal{H}$. By concavity, the f_n are still maximizers, i.e. $E[U(f_n)] \uparrow u_{\mathcal{H}}(0)$.

(i.4) Define $\widehat{f}$ as follows:

- in case $\mathcal{M}^e \cap P_V \neq \emptyset$, $\widehat{f} := f$. Here, $\overline{Q} \sim P$ and f is a well-defined element of $L^0(\Omega, \mathcal{F}_T, P)$ with $f_n \overset{P\text{-a.s.}}{\to} f = \widehat{f}$;
- in case $\mathcal{M}^e \cap P_V = \emptyset$, and thus $\overline{Q} = \widehat{Q}$,

$$\widehat{f} := f I_{\{\frac{d\widehat{Q}}{dP} > 0\}} + \overline{x} I_{\{\frac{d\widehat{Q}}{dP} = 0\}}$$

By construction, $\widehat{f} \in L^1(\widehat{Q})$ in both cases.

(ii) It is easily seen that for $y > 0$ and $Q \in \{\widehat{Q}, \overline{Q}\}$

$$\limsup_n \left(U(f_n) - f_n y \frac{dQ}{dP} \right) \leq U(\widehat{f}) - \widehat{f} y \frac{dQ}{dP} \leq V\left(y \frac{dQ}{dP} \right), \tag{4.23}$$

using the convention $+\infty \cdot 0 = 0$. The Fatou lemma applied to (4.23) for any y sufficiently large yields

$$\begin{aligned} u_{\mathcal{H}}(0) &= \limsup_n E\left[U(f_n) - f_n y \frac{dQ}{dP} \right] \leq E\left[\limsup_n \left(U(f_n) - f_n y \frac{dQ}{dP} \right) \right] \\ &\leq E\left[U(\widehat{f}) - \widehat{f} y \frac{dQ}{dP} \right] \leq E\left[V\left(y \frac{dQ}{dP} \right) \right]. \end{aligned} \tag{4.24}$$

In particular, we derive $U(\widehat{f}) \in L^1(P)$. On taking $Q = \overline{Q}$, in virtue of (4.10) and Corollary 4.4 we can let $y \to 0$ to obtain

$$u_{\mathcal{H}}(0) \leq E[U(\widehat{f})]. \tag{4.25}$$

Also, on taking $Q = \widehat{Q}$ and sending $y \to +\infty$ we get

$$E_{\widehat{Q}}[\widehat{f}] \leq 0. \tag{4.26}$$

Equation (4.24) with the choice of the optimizers $Q = \widehat{Q}, y = \widehat{y}$ yields

$$u_{\mathcal{H}}(0) \leq E[U(\widehat{f})] - \widehat{y} E_{\widehat{Q}}[\widehat{f}] \leq E\left[V\left(\widehat{y} \frac{d\widehat{Q}}{dP} \right) \right] = u_{\mathcal{H}}(0), \tag{4.27}$$

which implies $E_{\widehat{Q}}[\widehat{f}] = 0$ and $u_{\mathcal{H}}(0) = E[U(\widehat{f})]$, in view of (4.25), (4.26) and $\widehat{y} > 0$ from (4.17).

(iii) The Fenchel optimal relation $U(\widehat{f}) - \widehat{f} \widehat{y} \frac{d\widehat{Q}}{dP} \overset{P\text{-a.s.}}{=} V(\widehat{y} \frac{d\widehat{Q}}{dP})$ now follows from (4.25-4.27). From here we conclude

$$\frac{d\widehat{Q}}{dP} = 0 \Leftrightarrow U(\widehat{f}) = U(\overline{x}) = U(+\infty).$$

The forward implication follows from $V(0) = U(+\infty)$ and the converse from $\widehat{y} > 0$. The equality $E_{\widehat{Q}}[\widehat{f}] = 0$ has just been shown in (a.ii.2).

(iv) Since $\limsup_n (U(f_n) - f_n y \frac{d\widehat{Q}}{dP}) \leq U(\widehat{f}) - \widehat{f} y \frac{d\widehat{Q}}{dP}$ and the inequalities in (4.24) are equalities for $y = \widehat{y}$ and $Q = \widehat{Q}$, one has $\limsup_n U(f_n) = U(\widehat{f}) = U(+\infty)$ on $A := \{\frac{d\widehat{Q}}{dP} = 0\}$. Therefore, by passing to a subsequence that converges to the limsup we can assume $U(f_n) I_A \to U(\widehat{f}) I_A$, whence globally

$$U(f_n) \overset{P\text{-a.s.}}{\to} U(\widehat{f}). \tag{4.28}$$

Consider now an arbitrary $Q \in \mathcal{M} \cap P_V$. Given (4.28) necessarily $\liminf_n f_n I_A \geq \overline{x} I_A$ and therefore

$$\liminf_n |f_n| \geq |\widehat{f}|, \text{ and } \liminf_n f_n \geq \widehat{f}.$$

Additionally, $(f_n)_n$ is $L^1(Q)$ bounded: $E_Q[f_n] = 0$ and $(E[U(f_n)])_n$ is bounded from below, so Proposition 6.3 applies again. Therefore, Fatou Lemma yields $\widehat{f} \in L^1(Q)$ and

$$E_Q[\widehat{f}] \leq E_Q[\liminf_n f_n] \leq \liminf_n E_Q[f_n] = 0.$$

(v) Finally, the results when U is strictly concave and differentiable follow now from the pointwise identity $U(x) - xU'(x) = V(U'(x))$.

(b) (i) This follows by construction when $\mathcal{M}^e \cap P_V \neq \emptyset$, cf. item (a.i.3) above, and otherwise from $U(f_n) \to U(\widehat{f})$ when $\overline{x} = +\infty$, cf. equation (4.28).

(ii) Since $U(f_n) \overset{P\text{-a.s.}}{\to} U(\widehat{f})$, the L^1 convergence of the utilities is equivalent to showing uniform integrability of $(U(f_n))_n$. Given the convergence of the expected utility, $E[U(f_n)] \uparrow E[U(\widehat{f})]$, an argument "à la Scheffé" shows that the uniform integrability of $(U(f_n))_n$ is equivalent to uniform integrability of any of the two families $(U^-(f_n))_n$, $(U^+(f_n))_n$. $U(0) = 0$ and monotonicity of U imply $U^-(f_n) = -U(-f_n^-)$ and $U^+(f_n) = U(f_n^+)$.
Suppose by contradiction that the family $(U^+(f_n))_n \equiv (U(f_n^+))_n$ is *not* uniformly integrable, and proceed as in [KS03, Lemma 1]. Given the supposed lack of uniform integrability, there exist disjoint measurable sets $(A_n)_n$ and a constant $\alpha > 0$ such that

$$E[U(f_n^+) I_{A_n}] \geq \alpha.$$

Set $g_n = \sum_{i=1}^n f_i^+ I_{A_i}$ and fix a $Q \in \mathcal{M} \cap P_V$ satisfying the Inada condition (4.10). $(f_n)_n$ is $L^1(Q)$ bounded by Proposition 6.3 and clearly $E_Q[g_n] \leq nC$ where C is a positive bound on the $L^1(Q)$ norms of the sequence $(f_n)_n$. In addition, $E[U(g_n)] \geq n\alpha$ because the $(A_n)_n$ are disjoint. Therefore,

$$\frac{u_Q(nC)}{nC} \geq \frac{E[U(g_n)]}{nC} \geq \frac{\alpha}{C} > 0$$

and passing to the limit when $n \uparrow \infty$ the conclusion contradicts (4.10). So the family $(U^+(f_n))_n$ is uniformly integrable, and $(U(f_n))_n$ as well, which means $U(f_n)$ tends in $L^1(P)$ to $U(\widehat{f})$.

(iii) To see that $f_n \to \widehat{f}$ in $L^1(\widehat{Q})$, from $U(\widehat{f}) - \widehat{f}\widehat{y}\frac{d\widehat{Q}}{dP} = V(\widehat{y}\frac{d\widehat{Q}}{dP}) \geq U(f_n) - f_n\widehat{y}\frac{d\widehat{Q}}{dP}$ the difference $U(\widehat{f}) - U(f_n) - (\widehat{f} - f_n)\widehat{y}\frac{d\widehat{Q}}{dP}$ is nonnegative and has P-expectation which tends to zero. Henceforth such difference is $L^1(P)$ convergent to 0, which, thanks to $L^1(P)$ convergence of $U(f) - U(f_n)$, yields $L^1(P)$ convergence to 0 of $(\widehat{f} - f_n)\frac{d\widehat{Q}}{dP}$.

From Fenchel inequality,

$$f_n^- y_Q \frac{dQ}{dP} \leq V\left(y_Q \frac{dQ}{dP}\right) - U(-f_n^-) \leq V\left(y_Q \frac{dQ}{dP}\right) + |U(f_n)|$$

and given the P-uniform integrability of $(U(f_n))_n$, proved in (b.ii), the Q-uniform integrability of $(f_n^-)_n$ follows (see Lemma 6.2). Admitting $f_n \overset{P\text{-a.s.}}{\to} \widehat{f}$ and $E_Q[\widehat{f}] = 0$, and in view of $0 = \lim_n E_Q[f_n]$, an application of the Scheffé lemma again yields $f_n \overset{L^1(Q)}{\to} \widehat{f}$.

(iv) Recall that $X^n := H^n \cdot S$ are all $\widehat{Q}$ *uniformly integrable* martingales by Lemma 3.7. Moreover, $\widehat{Q}$ is a σ-martingale measure for S, so $X^n = (H^n \frac{1}{\varphi_{\widehat{Q}}}) \cdot (\varphi_{\widehat{Q}} \cdot S)$, where $M = \varphi_{\widehat{Q}} \cdot S$ is a $\widehat{Q}$ martingale and $\varphi_{\widehat{Q}} > 0$ holds $\widehat{Q}$-a.s. The convergence (4.21) permits a straightforward application of a celebrated result by Yor [Yor78] on the closure of stochastic integrals, which gives an integral representation with respect to M of the limit $\widehat{f}$ under $\widehat{Q}$, $\widehat{f} = H^* \cdot M_T = \widehat{H} \cdot S_T$, with $\widehat{H} = H^* \varphi_{\widehat{Q}}$, and the optimal process $\widehat{X} := \widehat{H} \cdot S$ is also a $\widehat{Q}$-uniformly integrable martingale.

(v) When there is $\widetilde{Q} \in \mathcal{M}^e \cap P_V$ with $E_{\widetilde{Q}}[\widehat{f}] = 0$ convergence (4.18) applies and by virtue of (b.iii) the construction of $\widehat{H}$ can be performed under $\widetilde{Q}$ instead of $\widehat{Q}$ and therefore $\widehat{H} \in L(S, P)$. To show $\widehat{H} \in \overline{\mathcal{H}}$, note we have already proved (4.19) so we only need convergence in P-probability of the wealth process at intermediate times. The convergence in (4.21) and the martingale property of the X^n and of $\widehat{H} \cdot S$ under $\widetilde{Q}$ imply

$$E_{\widetilde{Q}}[|X_t^n - \widehat{H} \cdot S_t|] = E_{\widetilde{Q}}[|E_{\widetilde{Q}}[X_T^n - \widehat{H} \cdot S_T \mid \mathcal{F}_t]|] \overset{\text{Jensen}}{\leq} E_{\widetilde{Q}}[|X_T^n - \widehat{H} \cdot S_T|].$$

Therefore, for any t, $X_t^n \to \widehat{H} \cdot S_t$ in $L^1(\widetilde{Q})$ and therefore in $\widetilde{Q}$-probability, which is equivalent to convergence in P-probability. Thus, $\widehat{H} \in \overline{\mathcal{H}}$ follows. $\square$

5. On the main assumptions and connections to literature.

5.1. More details on Assumption 4.2. Condition (4.9) is automatically satisfied for utilities finite on a half-line. For utilities finite on $\mathbb{R}$ it makes sure that the claim $\widehat{f}$ constructed via the Komlós theorem satisfies the budget constraint $E_Q[\widehat{f}] \leq x$ for every $Q \in \mathcal{M} \cap P_V$.

To the best of our knowledge Assumption 4.2 is strictly weaker than any other assumption used in the current literature for U finite on $\mathbb{R}$. In current references, the typical assumption is RAE(U), which implies $v_Q(y) < +\infty$ for all $y > 0$ and for all $Q \in P_V$ by [Sch01, Corollary 4.2], whence Assumption 4.2 necessarily holds. In the non-smooth utility case studied by Bouchard et al. [BTZ04], equivalent asymptotic elasticity conditions are imposed on the Fenchel conjugate V,

$$\lim_{y \to 0_+} \frac{|V'_-(y)|y}{V(y)} < +\infty, \qquad \lim_{y \to +\infty} \frac{|V'_+(y)|y}{V(y)} < +\infty. \tag{5.1}$$

These again imply $v_Q(y) < +\infty$ for all $y > 0$ and for all $Q \in P_V$, see [BTZ04, Lemma 2.3].

On the other hand, Biagini and Frittelli [BF05, BF08] do not require RAE(U), but instead assume that $v_Q(y)$ is finite for all $Q \in \mathcal{M} \cap P_V$ and all $y > 0$, which is weaker than RAE(U) but clearly stronger than Assumption 4.2 by virtue of Corollary 4.4. Since condition (4.10) is only slightly stronger than the truly necessary condition (4.8) for utility functions finite on a half-line, Assumption 4.2 seems to be a very good choice for a unified treatment of utility maximization problems, regardless of the domain of U.

5.2. A general duality formula and more details on Assumption 4.9. Duality theory applied in the Orlicz spaces context shows that the dual problem associated with the utility maximization over a general Orlicz space may contain singular parts, see [BF08]. We have tried to make this section as self-contained as possible, but the reader can find more details on the structure of the dual of a general Orlicz space in [RR91]. The dual variables $z \in (L^{\widehat{U}})^*$ have, in general, a two-way decomposition $z = z_r + z_s$ in regular and singular part, where z_r only can be identified with a measure absolutely continuous with respect to P. Let $\langle \cdot, \cdot \rangle$ denote the bilinear form for the dual system $(L^{\widehat{U}}, (L^{\widehat{U}})^*)$. The convex conjugate $(I_U)^* : (L^{\widehat{U}})^* \to (-\infty, +\infty]$ of the expected utility functional $L^{\widehat{U}} \ni k \mapsto E[U(k)] := I_U(k)$ is then defined as:

$$(I_U)^*(z) := \sup_{k \in L^{\widehat{U}}} \{ I_U(k) - \langle z, k \rangle \}.$$

Recall that the polar set of a cone $C \subset L^{\widehat{U}}$ is the subset of $(L^{\widehat{U}})^*$ defined as $C^0 := \{ z \in (L^{\widehat{U}})^* \mid \langle z, k \rangle \leq 0 \text{ for all } k \in C \}$. The set of normalized elements in C^0, i.e. those z which verify $\langle z, I_\Omega \rangle = 1$, is denoted by C^0_1. Thus, when $z \in C^0_1$ is regular it is an absolutely continuous normalized measure (with sign). The following Theorem is the key to understanding the precise implications of Assumption 4.9. Its proof is basically identical to [BF08, Theorem 21], but with our strategies $\mathcal{H}$.

THEOREM 5.1. *Under Assumption 3.1 and 4.2, for any $x \in (\underline{x}, \overline{x})$ the following dual relation holds:*

$$u_{\mathcal{H}}(x) = \min_{z \in C^0} (I_U)^*(z) = \min_{y>0, z \in C^0_1} \left\{ y(x + \|z_s\|) + E\left[V\left(y \frac{dz_r}{dP} \right) \right] \right\}, \tag{5.2}$$

where $C := \{ k \in L^{\widehat{U}} \mid k \leq H \cdot S_T \text{ for some } H \in \mathcal{H} \}$. When there is a regular dual minimizer, the above formula simplifies to

$$u_{\mathcal{H}}(x) = \min_{y>0, Q \in \mathcal{M} \cap P_V} \left\{ yx + E\left[V\left(y \frac{dQ}{dP} \right) \right] \right\}. \tag{5.3}$$

Proof. The first part of the proof goes along the same lines of the proof of Lemma 4.3 and thus we give only a sketch. Suppose for simplicity $x = 0$. As in Lemma 4.3, $u_{\mathcal{H}}(0) = \sup_{k \in C} E[U(k)]$ and the concave expected utility functional I_U is proper and has a continuity point which belongs to C. Then, Fenchel Duality Theorem applies and

$$u_{\mathcal{H}}(0) = \sup_{k \in C} E[U(k)] = \min_{z \in C^0} (I_U)^*(z) = \min_{z \in C^0} \left\{ E\left[V\left(\frac{dz_r}{dP} \right) \right] + \|z_s\| \right\}, \tag{5.4}$$

where the second equality follows from the explicit expression of the convex conjugate $(I_U)^*(z) = E[V(\frac{dz_r}{dP})] + \|z_s\|$ found by Kozek [Ko79]. Note that $C \supseteq -L^{\widehat{U}}_+$, so C^0_1 consists of positive normalized functionals. Assumption 4.2 implies in particular $\mathcal{M} \cap P_V \neq \emptyset$ and since $0 \in (\underline{x}, \overline{x})$ Proposition 4.6 implies $u_{\mathcal{H}}(0) \leq u_Q(0) < U(+\infty)$ for any $Q \in \mathcal{M} \cap P_V$. Thus $u_{\mathcal{H}}(0) < U(+\infty)$, so the dual minimizers are non null and the dual problem can be re-written as

$$\min_{y>0, z \in C^0_1} \left\{ E\left[V\left(y \frac{dz_r}{dP} \right) \right] + y\|z_s\| \right\},$$

via the normalized dual variables in C^0_1, which proves (5.2). Any dual minimizer $\hat{z} \in C^0_1$ clearly satisfies the integrability condition $E[V(y\frac{d\hat{z}_r}{dP})] < +\infty$ for some y. Since $\langle \hat{z}, I_\Omega \rangle = E[\frac{d\hat{z}_r}{dP} I_\Omega] + \langle \hat{z}_s, I_\Omega \rangle = 1$, when $\hat{z}_s = 0$ this exactly means $\hat{z} = \hat{z}_r \in P_V$. Suppose there exists a regular dual minimizer. Then, the optimal dual value is reached upon $C^0_1 \cap P_V$. Therefore,

$$u_{\mathcal{H}}(0) = \min_{y>0, Q \in C^0_1 \cap P_V} E\left[V\left(y\frac{dQ}{dP}\right)\right].$$

The Lemmata 3.7 and 6.4 rely on Assumption 3.1 to give $\mathcal{M} \cap P_V = C^0_1 \cap P_V$, whence the conclusion (5.3) follows. □

The above Theorem shows that the additional Assumption 4.9 amounts to requiring $\hat{z}_s = 0$ for some dual optimizer $\hat{z}$ in (5.2). The next Corollary provides a simple sufficient condition which ensures that *any* dual optimizer is regular.

COROLLARY 5.2. *Let U be finite on the whole $\mathbb{R}$ and let $S \in \mathscr{S}^{M^{\widehat{U}}}_\sigma$. Under Assumption 4.2, for any $x \in (\underline{x}, \overline{x})$ the simpler dual relation* (5.3) *holds. In other words, Assumption 4.9 is automatically satisfied if Assumption 4.2 holds and $S \in \mathscr{S}^{M^{\widehat{U}}}_\sigma$.*

Proof. Note first that the condition $S \in \mathscr{S}^{M^{\widehat{U}}}_\sigma$ may coincide with the generally weaker Assumption 3.1. This happens when $L^{\widehat{U}} = M^{\widehat{U}}$, that is when U has left tail which goes to $-\infty$ at a "moderate speed". In such case, the dual space $(L^{\widehat{U}})^*$ is free of singular parts—exactly as in the dual system (L^p, L^q) when $1 \leq p < +\infty$—and Theorem 5.1 immediately yields the strong dual relation (5.3).

So, suppose $S \in \mathscr{S}^{M^{\widehat{U}}}_\sigma$ but $M^{\widehat{U}} \subsetneq L^{\widehat{U}}$. The most intuitive way to show (5.3) is to note that terminal values $H \cdot S_T, H \in \mathcal{H}$, are in $M^{\widehat{U}}$, to set $\check{C} := \{k \in M^{\widehat{U}} \mid k \leq H \cdot S_T \text{ for some } H \in \mathcal{H}\}$ and to work with the dual system $(M^{\widehat{U}}, (M^{\widehat{U}})^*)$ instead of the full $(L^{\widehat{U}}, (L^{\widehat{U}})^*)$. The advantage is that the elements of $(M^{\widehat{U}})^*$ are regular. Then, an application of the duality arguments of Theorem 5.1 with C replaced by $\check{C}$ gives

$$u_{\mathcal{H}}(x) = \min_{y>0, Q \in (\check{C})^0_1 \cap P_V} \left\{xy + E\left[V\left(y\frac{dQ}{dP}\right)\right]\right\}.$$

Now, $(\check{C})^0_1$ consists of probabilities and as in the final part of the Theorem $(\check{C})^0_1 \cap P_V = \mathcal{M} \cap P_V$, whence (5.3).

For the interested reader we provide an alternative proof which is less intuitive as it requires an analysis of the behavior of singular elements of $(L^{\widehat{U}})^*$, but this proof makes direct use of the general dual formula (5.2). When $S \in \mathscr{S}^{M^{\widehat{U}}}_\sigma$ the set C^0_1 has a special structure:

$$C^0_1 \ni z = z_r + z_s \Leftrightarrow z_r \in C^0_1.$$

This can be seen through the following steps: (i) C^0 coincides in fact with $\{z \in (L^{\widehat{U}})^*_+ \mid \langle z, H \cdot S_T \rangle = 0, \ \forall H \in \mathcal{H}\}$, where the equality holds as $\mathcal{H}$ is a vector space, and here $H \cdot S_T \in M^{\widehat{U}}$; (ii) when U is finite on $\mathbb{R}$, $\widehat{U}$ is also finite everywhere and with such Young functions singular elements in the dual space are null over the Orlicz heart: if $z = z_s$ then z is null over $M^{\widehat{U}}$; (iii) the Orlicz heart contains L^∞; 4) consequently

$z \in C_1^0$ iff z is a positive functional and

$$\langle z, I_\Omega \rangle = E\left[\frac{dz_r}{dP}\right] + \langle z, I_\Omega \rangle = E\left[\frac{dz_r}{dP}\right] = 1,$$
$$\langle z, H\cdot S_T \rangle = E\left[\frac{dz_r}{dP} H\cdot S_T\right] + \langle z_s, H\cdot S_T \rangle = E\left[\frac{dz_r}{dP} H\cdot S_T\right] = 0 \text{ for all } H \in \mathcal{H},$$

that is, iff $z_r \in C_1^0$. Now, a simple inspection of the dual problem in (5.2) shows that, for any $z \in C_1^0$, setting z_s to zero makes the dual function to be minimized smaller. Hence, any minimizer is regular, i.e. we have shown (5.3). □

5.3. Characterization of the optimal solution: $\overline{x} = +\infty$, $\widehat{Q}$ not equivalent to P. When U is strictly monotone (a typical example is the exponential utility) but $\widehat{Q}$ is not equivalent to P one can express the optimal terminal wealth $\widehat{f}$ using integrands in $L(S, \widehat{Q})$ but no longer using the more natural strategies in $L(S, P)$. An *approximation result* for $\widehat{f}$ via integrands in $L(S, P)$ was first shown by Acciaio [A05], under the following technical conditions:

(i) U is differentiable, monotone, strictly concave and it satisfies RAE(U) (4.6, 4.7);
(ii) S is locally bounded;
(iii) the stopping times of the filtration are predictable.

Acciaio builds a sequence of integrals $\widetilde{H}_n \cdot S_T$, whose expected utility tends to the optimum, and which satisfies $(x + \widetilde{H}_n \cdot S_T) \to \widehat{f}$ P-a.s.

Our setup allows us to remove the technical conditions above while proving P-a.s. convergence of terminal wealths in item (b.i) of Theorem 4.10 and a stronger $L^1(P)$ convergence of utilities in item (b.ii), which implies convergence of expected utility.

6. Auxiliary results. LEMMA 6.1. *Let $\Psi : \mathbb{R} \to (-\infty, +\infty]$ be a convex, lower semicontinuous function. For a given sequence $(x_n)_n$, if $d_n \in \mathbb{R}_+$, $\sum_{n\geq 1} d_n = 1$ and $\sum_{n\geq 1} d_n x_n$ converges, then*

$$\Psi(\sum_{n\geq 1} d_n x_n) \leq \liminf_N \sum_{n=1}^{N} d_n \Psi(x_n).$$

When Ψ is bounded from below, the above inequality simplifies to

$$\Psi(\sum_{n\geq 1} d_n x_n) \leq \sum_{n\geq 1} d_n \Psi(x_n).$$

Proof. From convexity of Ψ,

$$\Psi(\sum_{n=1}^{N} d_n x_n) \leq (1 - \sum_{n=1}^{N} d_n)\Psi(0) + \sum_{n=1}^{N} d_n \Psi(x_n).$$

When $N \uparrow +\infty$, $\sum_{n=1}^{N} d_n x_n \to \sum_{n\geq 1} d_n x_n$ so that lower semicontinuity of Ψ implies $\Psi(\sum_{n\geq 1} d_n x_n) \leq \liminf_{N\to+\infty} \Psi(\sum_{n=1}^{N} d_n x_n)$. The above displayed chain shows that such lim inf is dominated by $\liminf_N \sum_{n=1}^{N} d_n \Psi(x_n)$. Finally, when Ψ is bounded from below, the latter series admits a limit (finite or $+\infty$). □

LEMMA 6.2. *Let $Q \ll P$. If $(Z^n \frac{dQ}{dP})_n$ is P-uniformly integrable, then $(Z^n)_n$ is Q-uniformly integrable, and vice versa.*

Proof. This intuitive Lemma is a consequence of the Dunford-Pettis criterion: A subset $K \subset L^1$ is uniformly integrable if and only if it is relatively compact for the weak topology. However, here is an elementary proof. Recall $(X_\alpha)_\alpha$ is uniformly integrable when

$$\lim_{r\to+\infty}\sup_\alpha E[|X_\alpha|\, I_{\{X_\alpha\geq r\}}]=0.$$

There is a well-known equivalent characterization of uniform integrability for random variables: $(X_\alpha)_\alpha$ is uniformly integrable if and only if *i)* the family is uniformly bounded in $L^1(P)$ and *ii)* for every $\varepsilon>0$ there exists $\delta>0$ such that whenever $P(A)<\delta$, $\sup_\alpha E[I_A|X_\alpha|]<\varepsilon$ (see e.g. the book [Sh96, Chapter 2.6]). So, suppose $(Z^n\frac{dQ}{dP})_n$ is P-uniformly integrable. Then, for every $r>0$

$$\begin{aligned}E_Q\left[I_{\{|Z^n|>r\}}|Z^n|\right] &\leq E_Q\left[I_{\{|Z^n|>r,\,\frac{dQ}{dP}>\frac{1}{\sqrt{r}}\}}|Z^n|\right]+E_Q\left[I_{\{\frac{dQ}{dP}\leq\frac{1}{\sqrt{r}}\}}|Z^n|\right]\\ &\leq E\left[I_{\{|Z^n|\frac{dQ}{dP}>\sqrt{r}\}}|Z^n|\frac{dQ}{dP}\right]+E\left[I_{\{0<\frac{dQ}{dP}\leq\frac{1}{\sqrt{r}}\}}|Z^n|\frac{dQ}{dP}\right],\end{aligned}$$

whence

$$\begin{aligned}&\lim_{r\to+\infty}\sup_n E_Q\left[I_{\{|Z^n|>r\}}|Z^n|\right]\\ &\quad\leq\lim_{r\to+\infty}\sup_n\left(E\left[I_{\{|Z^n|\frac{dQ}{dP}>\sqrt{r}\}}|Z^n|\frac{dQ}{dP}\right]+E\left[I_{\{0<\frac{dQ}{dP}\leq\frac{1}{\sqrt{r}}\}}|Z^n|\frac{dQ}{dP}\right]\right)=0,\end{aligned}$$

where the last equality follows from P-uniform integrability of $(Z^n\frac{dQ}{dP})_n$ and from the fact that $\{0<\frac{dQ}{dP}\leq\frac{1}{\sqrt{r}}\}$ has P-probability which tends to 0 when r goes to $+\infty$.

The converse implication follows directly from $Q\ll P$:

$$\begin{aligned}\lim_{P(A)\to0}\sup_n E\left[I_A|Z^n|\frac{dQ}{dP}\right] &=\lim_{P(A)\to0}\sup_n E_Q[I_A|Z^n|]\\ &\leq\lim_{Q(A)\to0}\sup_n E_Q[I_A|Z^n|]=0.\qquad\square\end{aligned}$$

Proof of Lemma 4.3. $u_Q(x)<+\infty$ follows from Fenchel inequality and from finite generalized relative entropy of Q: if X satisfies $E_Q[X]\leq x$, $E[U(X)]\leq xy_Q+v_Q(y_Q)$, with y_Q from Definition 3.4. The dual formula to be proved is actually a straightforward consequence of the Fenchel duality formula and of the results obtained by Rockafellar in the 1970-ies on conjugates of functionals in integral form (here, expected utility). However, we give a different proof based on Orlicz duality, since it is useful for Theorem 5.1 where the Orlicz setup is necessary.

The utility maximization problem $\sup_{E_Q[X]\leq x}E[U(X)]$ can be rewritten over the utility-induced Orlicz space $L^{\widehat{U}}(P)$ defined in (2.2). This can be done because: i) the supremum will be reached over those X such that $E[U(X)]$ is finite, so that $-X^-\in L^{\widehat{U}}(P)$; ii) if $E[U(-X^-)]>-\infty$ then the truncated sequence $X_n=X\wedge n$ is also in the Orlicz space and by Fatou Lemma in the limit it delivers the same expected utility from X; iii) $L^{\widehat{U}}(P)\subseteq L^1(Q)$, which follows from $Q\in P_V$, from (2.7) and Fenchel inequality (this also implies Q is in the topological dual of $L^{\widehat{U}}$). Therefore, $u_Q(x)=\sup_{X\in L^{\widehat{U}},E_Q[X]\leq x}E[U(X)]$. On $L^{\widehat{U}}$, the concave functional $I_U(X):=E[U(X)]$ is proper:

$$X\in L^{\widehat{U}}\Rightarrow X\in L^1(P)\text{ so that }E[U(X)]\overset{\text{Jensen}}{\leq}U(E[X])<+\infty.$$

Moreover, I_U has a continuity point which belongs to the maximization domain $D = \{X \in L^{\widehat{U}} \mid E_Q[X] \le x\}$. This is more subtle to check, but it can be proved that the set

$$\mathcal{B} := \{X \in L^{\widehat{U}} \mid E[U(-(1+\epsilon)X^-)] > -\infty \text{ for some } \epsilon > 0\},$$

coincides with the interior of the proper domain of I_U (see [BFG08, Lemma 4.1] modulo a sign change), where I_U is automatically continuous by the Extended Namioka Theorem (see e.g. [BF09]). Then, as $x > \underline{x}$, the constant x is in $\mathcal{B} \cap D$.

The dual formula (4.12) is thus a consequence of Fenchel Duality Theorem [Bre83, Chapter 1], of the fact that the polar set of the constraint $C := \{X \mid E_Q[X] \le x\} \supseteq -L^\infty_+$, i.e. the set $\{\mu \in (L^{\widehat{U}})^* \mid \mu(X) \le x \,\forall X \in C\}$, by the Bipolar Theorem is the positive ray $\{yQ \mid y \ge 0\}$, and of the expression of the convex conjugate $(I_U)^*$ of I_U over the variables $y\frac{dQ}{dP}$: $(I_U)^*(y\frac{dQ}{dP}) = E[V(y\frac{dQ}{dP})] = v_Q(y)$. □

Proposition 6.3. *Suppose $(k_n)_n$ is a sequence of random variables such that $(E[U(k_n)])_n$ is bounded from below and assume $(E_{\widetilde{Q}}[k_n])_n$ is bounded from above for some $\widetilde{Q} \in P_V$ satisfying the Inada condition* (4.10). *Then the following statements hold:*

(i) *$U(k_n)$ is $L^1(P)$-bounded;*

(ii) *k_n is $L^1(Q)$-bounded for* any *$Q \in P_V$ for which $E_Q[k_n]$ is bounded from above. The indirect utility u_Q need not satisfy the Inada condition* (4.10).

Proof. In this proof c refers to a constant, not necessarily the same on each line.

(i) By hypothesis there is $0 < y_1 < y_2$ such that $v_{\widetilde{Q}}(y_i) < +\infty$ for $i = 1, 2$. The Fenchel inequality implies

$$E[U(-k_n^-)] \le v_{\widetilde{Q}}(y_2) - y_2 E_{\widetilde{Q}}[k_n^-], \tag{6.1}$$

$$E[U(k_n^+)] \le v_{\widetilde{Q}}(y_1) + y_1 E_{\widetilde{Q}}[k_n^+], \tag{6.2}$$

which yields

$$E[U(k_n)] \le c + y_1 E_{\widetilde{Q}}[k_n] - (y_2 - y_1) E_{\widetilde{Q}}[k_n^-]. \tag{6.3}$$

By assumption, $(E_{\widetilde{Q}}[k_n])_n$ is bounded from above and $(E[U(k_n)])_n$ is bounded from below, whereby one concludes from (6.3) and from $y_2 - y_1 > 0$ that $(E_{\widetilde{Q}}[k_n^-])_n$ is bounded and consequently $(E_{\widetilde{Q}}[k_n^+])_n$ is also bounded. Finally, by (6.2) the sequence $(E[U(k_n^+)])_n$ is bounded. Since $U(k_n^+) \ge 0, U(-k_n^-) \le 0$ and $(E[U(k_n)])_n$ is bounded from below the $L^1(P)$-boundedness of $U(k_n)$ follows.

(ii) The inequality (6.1) applies for any $Q \in P_V$, i.e. there is $y_Q > 0$ such that

$$E[U(-k_n^-)] \le c - y_Q E_Q[k_n^-]. \tag{6.4}$$

By (i) the sequence $(E[U(-k_n^-)])_n$ is bounded from below whereby $(E_Q[k_n^-])_n$ must be bounded. As in (i), this and boundedness from above of the expectations $(E_Q[k_n])_n$ ensure $(E_Q[k_n^+])_n$ is also bounded. □

Lemma 6.4. *Let $Q \in P_V$ verify $E_Q[X_T] = 0$ for all $X = H \cdot S, H \in \mathcal{H}$. Then $Q \in \mathcal{M} \cap P_V$.*

Proof. We just need to show $Q \in \mathcal{M}$. Consider $S \in \mathscr{S}^{\widehat{U}}_\sigma \setminus \mathscr{S}^{\widehat{U}}_{\mathrm{loc}}$, fix any $\mathcal{I}$-localizing φ from Proposition 2.6 and let $S' = \varphi \cdot S$. For any $A \in \mathcal{F}_s, s \in [0, T[, t > s$ let $H = I_A I_{]s,t]}\varphi$, which is in $\mathcal{H}$. Since $H \cdot S = (I_A I_{]s,t]}) \cdot S'$ and $E_Q[H \cdot S_T] = E_Q[I_A(S'_t - S'_s)] = 0$, for all $A \in \mathcal{F}_s, s < t$, S' is a then Q-martingale, and hence $Q \in \mathcal{M}$. For $S \in \mathscr{S}^{\widehat{U}}_{\mathrm{loc}}$ we proceed as above, replacing φ with $I_{[0,\tau_n]}$. □

Acknowledgements. We would like to thank Jan Kallsen, Fabio Maccheroni, Mark Owen, Antonis Papapantoleon, Maurizio Pratelli, Mihai Sîrbu and two anonymous referees for helpful comments.

REFERENCES

[A05] B. Acciaio, *Absolutely continuous optimal martingale measures*, Statist. Decisions, 23 (2005), pp. 81–100.

[AS94] J.-P. Ansel and Ch. Stricker, *Couverture des actifs contingents et prix maximum*, Ann. Inst. H. Poincaré Probab. Statist., 30 (1994), pp.303–315.

[BeF02] F. Bellini and M. Frittelli, *On the existence of minimax martingale measures*, Math. Finance, 12 (2002), pp. 1–21.

[Bia04] S. Biagini, *A new class of strategies and application to utility maximization for unbounded processes*, Technical report of the University of Perugia, 2004, available at http://sites.google.com/site/sarabiagini/.

[BF05] S. Biagini and M. Frittelli, *Utility maximization in incomplete markets for unbounded processes*, Finance Stoch., 9 (2005), pp. 493–517.

[BF07] ———, *The supermartingale property of the optimal wealth process for general semimartingales*, Finance Stoch., 11 (2007), pp. 253–266.

[BF08] ———, *A unified framework for utility maximization problems: an Orlicz space approach*, Ann. Appl. Probab., 18 (2008), pp. 929–966.

[BF09] ———, *On the extension of the Namioka-Klee Theorem and on the Fatou property for risk measures*, in Optimality and Risk – Modern Trends in Mathematical Finance: Youri Kabanov's Festschrift, F. Delbaen, M. Rasonyi and Ch. Stricker, eds., Springer-Verlag, Berlin, 2009, pp. 1–23.

[BFG08] S. Biagini, M. Frittelli and M. Grasselli, *Indifference price with general semimartingales*. Math. Finance, to appear. Preprint version available at http://sites.google.com/site/sarabiagini/.

[BG09] S. Biagini and P. Guasoni, *Relaxed utility maximization in complete markets*, Math. Finance, to appear. Preprint version available at http://sites.google.com/site/sarabiagini/.

[BS09] S. Biagini and M. Sîrbu, *A note on investment opportunities when the credit line is infinite*. Preprint available at http://sites.google.com/site/sarabiagini/.

[B02] B. Bouchard, *Utility maximization on the real line under proportional transaction costs*. Finance Stoch., 6 (2002), pp. 495–516.

[BTZ04] B. Bouchard, N. Touzi and A. Zeghal, *Dual formulation of the utility maximization problem: The case of nonsmooth utility*, Ann. Appl. Probab., 14 (2004), pp. 678–717.

[Bre83] H. Brezis, *Analyse Fonctionnelle*, Masson, Paris, 1983.

[ČK07] A. Černý and J. Kallsen, *On the structure of general mean–variance hedging strategies*, Ann. Probab., 35 (2007), pp. 1479–1531.

[CSW01] J. Cvitanić, W. Schachermayer and H. Wang, *Utility maximization in incomplete markets with random endowment*, Finance Stoch., 5 (2001), pp. 237–259.

[Dal02] F. Delbaen, P. Grandits, T. Rheinländer, D. Samperi, M. Schweizer and Ch. Stricker, *Exponential hedging and entropic penalties*, Math. Finance, 12 (2002), pp. 99–123.

[DS96] F. Delbaen and W. Schachermayer, *Attainable claims with p-th moments*, Ann. Inst. H. Poincaré, 32 (1996), pp. 743–763.

[DS98] ———, *The fundamental theorem of asset pricing for unbounded stochastic processes*, Math. Ann., 312 (1998), pp. 215–250.

[DS99] ———, *A compactness principle for bounded sequences of martingales with applications*, in Proceedings of the Seminar on Stochastic Analysis, Random Fields and Applications, vol. 45 of Progr. Probab., Birkhäuser, Basel, 1999, pp. 137–173.

[DS05] ———, *The Mathematics of Arbitrage*, Springer Finance, Springer-Verlag, Berlin, 2005.

[E80] M. Émery, *Compensation de processus à variation finie non localement intégrables*, in Séminaire de Probabilités XIV, 1978/79, vol. 784 of Lecture Notes in Math., Springer-Verlag, Berlin, 1980, pp. 152–160.

[GR01] T. Goll and L. Rüschendorf, *Minimax and minimal distance martingale measures and their relationship to portfolio optimization*, Finance Stoch., 5 (2001), pp. 557–581.

[HK79] J. M. Harrison and D. M. Kreps, *Martingales and arbitrage in multiperiod securities* *market*, J. Econom. Theory, 20 (1979), pp. 381–408.

[JS03] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, vol. 288 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, second ed., 2003.

[KSt02] Y. Kabanov and C. Stricker, *On the optimal portfolio for the exponential utility maximization: Remarks to the six-authors paper.* Math. Finance, 12 (2002), pp. 125–134.

[Ka04] J. Kallsen, *σ-localization and σ-martingales*, Theory Probab. Appl., 48 (2004), pp. 152–163.

[Ko79] A. Kozek, *Convex integral functionals on Orlicz spaces*, Comment. Math., 21 (1980), pp. 109–135.

[KS99] D. Kramkov and W. Schachermayer, *The asymptotic elasticity of utility functions and optimal investment in incomplete markets.* Ann. Appl. Probab., 9 (1999), pp. 904–950.

[KS03] ———, *Necessary and sufficient conditions in the problem of optimal investment in incomplete markets*, Ann. Appl. Probab., 13 (2003), pp. 1504–1516.

[KSî06] D. Kramkov and M. Sîrbu, *On the two-times differentiability of the value functions in the problem of optimal investment in incomplete markets*, Ann. Appl. Probab., 16 (2006), 1352–1384.

[KR61] M. A. Krasnosel'skii and Ya. B. Rutickii, *Convex Functions and Orlicz Spaces*, P. Noordhoff Ltd., Groningen, 1961.

[KL51] S. Kullback and R. Leibler, *On information and sufficiency*, Ann. Math. Statistics, 22 (1951), pp. 79–86.

[LV87] F. Liese and I. Vajda, *Convex Statistical Distances*, Teubner, Leipzig, 1987.

[M78] P.-A. Meyer, *Inégalités de normes pour les intégrales stochastiques*, in Séminaire de Probabilités XII, 1976/77, vol. 649 of Lecture Notes in Math., Springer-Verlag, Berlin, 1978, pp. 757–762.

[O02] M. P. Owen, *Utility-based optimal hedging in incomplete markets*, Ann. Appl. Probab., 12 (2002), pp. 691-709.

[OŽ09] M. P. Owen and G. Žitković, *Optimal investment with an unbounded random endowment and utility-based pricing*, Math. Finance, 19 (2009), pp. 129–159.

[Pr05] P. Protter, *Stochastic Integration and Differential Equations*, Springer-Verlag, Berlin, second ed., 2005.

[RR91] M. M. Rao and Z. D. Ren, *Theory of Orlicz Spaces*, Marcel Dekker Inc., New York, 1991.

[RS05] M. Rásonyi and Ł. Stettner, *On utility maximization in discrete-time financial market models*, Ann. Appl. Probab., 15 (2005), pp. 1367–1395.

[R70] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[Sat99] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, vol. 68 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1999.

[Sch94] W. Schachermayer, *Martingale measures for discrete time processes with infinite horizon*, Math. Finance, 4 (1994), pp. 25–55.

[Sch01] ———, *Optimal investment in incomplete markets when wealth may become negative*, Ann. Appl. Probab., 11 (2001), pp. 694-734.

[Sch03] ———, *A super-martingale property of the optimal portfolio process.* Finance Stoch., 4 (2003), pp. 433-457.

[Sh96] A. N. Shiryaev, *Probability*, vol. 95 of Graduate Texts in Mathematics, Springer-Verlag, New York, second ed., 1996.

[St03] Ch. Stricker, *Simple strategies in exponential utility maximization*, in Séminaire de Probabilités, XXXVI, vol. 1801 of Lecture Notes in Math., Springer-Verlag, Berlin, 2003, pp. 415–418.

[Yor78] M. Yor, *Sous-espaces denses dans L^1 ou H^1 et représentation des martingales*, in Séminaire de Probabilités XII, 1976/1977, vol. 649 of Lecture Notes in Math., Springer-Verlag, Berlin, 1978, pp. 265–309.