Exponential models by Orlicz spaces and applications

Original
Exponential models by Orlicz spaces and applications / Santacroce, Marina; Siri, Paola; Trivellato, Barbara. - In: JOURNAL OF APPLIED PROBABILITY. - ISSN 0021-9002. - ELETTRONICO. - 55:3(2018), pp. 682-700.

Publisher:
Cambridge University Press

Published
DOI:10.1017/jpr.2018.45

Exponential Models by Orlicz Spaces and Applications

Marina Santacroce[1], Paola Siri *[1], and Barbara Trivellato[1]

[1]Dipartimento di Scienze Matematiche "G.L. Lagrange", Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy,
e-mails: marina.santacroce@polito.it, paola.siri@polito.it, barbara.trivellato@polito.it

Abstract

The geometric structure of the non-parametric statistical model of all positive densities connected by an open exponential arc and its intimate relation to Orlicz spaces give new insights to well known financial objects which arise in exponential utility maximization problems.

2000 Mathematics Subject Classification: 46E30, 46N30, 62B10, 91G80.
Key words and phrases: Orlicz spaces, maximal exponential model, exponential utility maximization, minimal entropy martingale measure, Reverse Hölder condition.

1 Introduction

Statistical exponential models built on Orlicz spaces arise in several fields, such as differential geometry, algebraic statistics and information theory. To our knowledge, their application to finance has not been investigated yet, although the use of Orlicz spaces in utility maximization and in risk measure theory is known (see, e.g., Cheridito and Li (2009), Biagini and Frittelli (2008)).

The aim of this paper is to provide a first investigation in this direction, particularly concerning maximal exponential models, by using some recent results contained in Santacroce, Siri and Trivellato (2016).

The theory of non-parametric maximal exponential models centered at a given positive density p starts with the work by Pistone and Sempi (1995). In that paper, and subsequently in Cena and Pistone (2007), by using the Orlicz space associated to an exponentially growing Young function, the set of positive densities is endowed with a structure of exponential Banach manifold. Such a manifold setting turns out to be well-suited for applications in physics as some recent papers show (see, e.g., Lods and Pistone (2015)).

One of the main result in Cena and Pistone (2007) states that any density belonging to the maximal exponential model centered at p is connected by an *open* exponential arc to p and viceversa, (by *open*, we essentially mean that the two densities are not the extremal points of the arc). In Santacroce, Siri and Trivellato (2016), the equivalence between the equality of the

*corresponding author

maximal exponential models centered at two (connected) densities p and q and the equality of the Orlicz spaces referred to the same densities is proved.

This work is a natural continuation of the previous one and, moreover, it includes applications to finance.

The paper is essentially composed of two parts. In the first part of the paper we give new theoretical results concerning exponential models which can be useful to understand their underlying geometrical structure. Specifically, after recalling in Sections 2 and 3 some preliminary results on Orlicz spaces and exponential models, in Subsection 3.1 we show that the equality of Orlicz spaces referred to connected densities is equivalent to the existence of a transport mapping between the corresponding coniugate spaces. Furthermore, in Subsection 3.2, we deal with densities projections on sub-σ-algebras and relate them to exponential sub-models. We show that exponential connection by arc is stable with respect to projections and that projected densities belong to suitable sub-models.

The second part of the work addresses the classical problem of exponential utility maximization in incomplete markets. In the literature, the study of the optimal solution of the corresponding dual problem is often related to the so-called Reverse Hölder condition. In Section 4, assuming this condition, we show that the minimal entropy martingale density measure belongs to a maximal exponential model. This reflects on the solution of the primal problem, which translates into a smoothness condition on the optimal wealth process. We use the exponential connection by arcs to slightly improve some well-known duality results and we do it by exploiting the equivalent conditions proved in Santacroce, Siri and Trivellato (2016). We conclude with Subsection 4.1, where our results are illustrated in some classical examples of financial markets taken from the literature.

2 Preliminaries on Orlicz Spaces

In this section we recall some known results from the theory of Orlicz spaces, which will be useful in the sequel. For further details on Orlicz spaces, the reader is referred to Rao and Ren (1991, 2002).

Let $(\mathcal{X}, \mathcal{F}, \mu)$ be a fixed measure space. Young functions can be seen as generalizations of the functions $f(x) = \frac{|x|^a}{a}$, with $a > 1$, and consequently, Orlicz spaces are generalizations of the Lebesgue spaces $L^a(\mu)$. The definition of Young function and of the related Orlicz space are given in the following.

Definition 2.1. *A Young function Φ is an even, convex function $\Phi : \mathbb{R} \to [0, +\infty]$ such that*

i) $\Phi(0) = 0$,

ii) $\lim_{x\to\infty} \Phi(x) = +\infty$,

iii) $\Phi(x) < +\infty$ *in a neighborhood of 0.*

The *conjugate function* Ψ of Φ, is defined as $\Psi(y) = \sup_{x\in\mathbb{R}}\{xy - \Phi(x)\}$, $\forall y \in \mathbb{R}$ and it is itself a Young function.

Let L^0 denote the set of all measurable functions $u : X \to \mathbb{R}$ defined on $(\mathcal{X}, \mathcal{F}, \mu)$.

Definition 2.2. *The Orlicz space $L^{\Phi}(\mu)$ associated to the Young function Φ is defined as*

$$L^{\Phi}(\mu) = \left\{ u \in L^0 \ : \ \exists \ \alpha > 0 \ s.t. \int_{\mathcal{X}} \Phi(\alpha u) d\mu < +\infty \right\}. \tag{2.1}$$

The Orlicz space $L^{\Phi}(\mu)$ is a vector space. Moreover, one can show that it is a Banach space when endowed with the *Luxembourg norm*

$$\|u\|_{\Phi,\mu} = \inf \left\{ k > 0 \ : \ \int_{\mathcal{X}} \Phi\left(\frac{u}{k}\right) d\mu \leq 1 \right\}. \tag{2.2}$$

Consider the Orlicz space $L^{\Phi}(\mu)$ with the Luxembourg norm $\|\cdot\|_{\Phi,\mu}$ and denote by $B(0,1)$ the open unit ball and by $\overline{B(0,1)}$ the closed one. Let us observe that,

$$\begin{aligned} u \in B(0,1) &\iff \exists \ \alpha > 1 \text{ s.t. } \int_{\mathcal{X}} \Phi(\alpha u) d\mu \leq 1, \\ u \in \overline{B(0,1)} &\iff \int_{\mathcal{X}} \Phi(u) d\mu \leq 1. \end{aligned}$$

Moreover, the Luxembourg norm is equivalent to the *Orlicz norm*

$$N_{\Phi,\mu}(u) = \sup_{v \in L^{\Psi}(\mu): \int_{\mathcal{X}} \Psi(v) d\mu \leq 1} \left\{ \int_{\mathcal{X}} |uv| d\mu \right\}, \tag{2.3}$$

where Ψ is the conjugate function of Φ.

It is worth to recall that the same Orlicz space can be related to different *equivalent* Young functions.

Definition 2.3. *Two Young functions Φ and Φ' are said to be equivalent if there exists $x_0 > 0$, and two positive constants $c_1 < c_2$ such that , $\forall x \geq x_0$,*

$$\Phi(c_1 x) \leq \Phi'(x) \leq \Phi(c_2 x).$$

In such a case the Orlicz spaces $L^{\Phi}(\mu)$ and $L^{\Phi'}(\mu)$ are equal as sets and have equivalent norms as Banach spaces.

From now on, we consider a probability space $(\mathcal{X}, \mathcal{F}, \mu)$ and we denote with $\mathcal{P}$ the set of all densities which are positive μ-a.s. Moreover, we use the notation $\mathbb{E}_p$ to indicate the expectation with respect to $p d\mu$, for each fixed $p \in \mathcal{P}$.

In the sequel, we use the Young function $\Phi_1(x) = \cosh(x) - 1$, which is equivalent to the more commonly used $\Phi_2(x) = e^{|x|} - |x| - 1$.
We recall that the conjugate function of $\Phi_1(x)$ is $\Psi_1(y) = \int_0^y \sinh^{-1}(t) dt$, which, in its turn, is equivalent to $\Psi_2(y) = (1 + |y|) \log(1 + |y|) - |y|$.

Furthermore, in order to stress that we are working with densities $p \in \mathcal{P}$, we denote by $L^{\Phi_1}(p)$ the Orlicz space associate to Φ_1, defined with respect to the measure induced by p, i.e.

$$L^{\Phi_1}(p) = \left\{ u \in L^0 \ : \ \exists \ \alpha > 0 \ s.t. \ \mathbb{E}_p(\Phi_1(\alpha u)) < +\infty \right\}. \tag{2.4}$$

It is worth to note that, in order to prove that a random variable u belongs to $L^{\Phi_1}(p)$, it is sufficient to check that $\mathbb{E}_p(e^{\alpha u}) < +\infty$, with α belonging to an open interval containing 0.
Finally, let us remark the following chain of inclusions:

$$L^{\infty}(p) \subseteq L^{\Phi_1}(p) \subseteq L^{a}(p) \subseteq L^{\psi_1}(p) \subseteq L^{1}(p),\ a > 1.$$

3 Exponential Models

We start by recalling the definitions of exponential arcs and some related results.

Definition 3.1. *Two densities $p, q \in \mathcal{P}$ are connected by an open exponential arc if there exists an open interval $I \supset [0,1]$ such that $p(\xi) \propto p^{(1-\xi)} q^{\xi}$ belongs to $\mathcal{P}$, for every $\xi \in I$.*

The following proposition gives an equivalent definition of exponential connection by arc. Its proof can be found in Santacroce, Siri and Trivellato (2016).

Proposition 3.2. *$p, q \in \mathcal{P}$ are connected by an open exponential arc if and only if there exist an open interval $I \supset [0,1]$ and a random variable $u \in L^{\Phi_1}(p)$, such that $p(\xi) \propto e^{\xi u} p$ belongs to $\mathcal{P}$, for every $\xi \in I$ and $p(0) = p$, $p(1) = q$.*

The connections by open exponential arcs is an equivalence relation (see Cena and Pistone (2007) for the proof).

In the following, we recall the definition of the cumulant generating functional and its properties, in order to introduce the notion of maximal exponential model. In the next section, the maximal exponential model at p is proved to coincide with the set of all densities $q \in \mathcal{P}$ which are connected to p by an open exponential arc.
Let us denote

$$L_0^{\Phi_1}(p) = \{u \in L^{\Phi_1}(p) : \mathbb{E}_p(u) = 0\}.$$

Definition 3.3. *The cumulant generating functional is the map*

$$\begin{array}{rccc} K_p : & L_0^{\Phi_1}(p) & \longrightarrow & [0, +\infty] \\ & u & \longmapsto & \log \mathbb{E}_p(e^{u}). \end{array} \tag{3.5}$$

Theorem 3.4. *The cumulant generating functional K_p satisfies the following properties:*

i) $K_p(0) = 0$*; for each* $u \neq 0$*,* $K_p(u) > 0$*.*

ii) K_p *is convex and lower semicontinuous, moreover its proper domain*

$$dom\, K_p = \{u \in L_0^{\Phi_1}(p) : K_p(u) < +\infty\}$$

is a convex set which contains the open unit ball of $L_0^{\Phi_1}(p)$. In particular, its interior $\overset{\circ}{dom}\, K_p$ is a non empty convex set.

For the proof one can see Pistone and Sempi (1995).

Definition 3.5. *For every density $p \in \mathcal{P}$, the maximal exponential model at p is*

$$\mathcal{E}(p) = \left\{ q = e^{u - K_p(u)} p : u \in \overset{\circ}{dom}\, K_p \right\} \subseteq \mathcal{P}.$$

Remark 3.6. We have defined K_p on the set $L_0^{\Phi_1}(p)$ because centering random variables guarantees the uniqueness of the representation of $q \in \mathcal{E}(p)$.

3.1 Characterizations

From now on we use the notation $D(q\|p)$ to indicate the Kullback-Leibler divergence of $q \cdot \mu$ with respect to $p \cdot \mu$ and we simply refer to it as the divergence of q from p.
We first state two results related to Orlicz spaces, which will be used in the sequel. Their proofs can be found in Cena and Pistone (2007).

Proposition 3.7. *Let p and q belong to $\mathcal{P}$ and let Φ be a Young function.*
The Orlicz spaces $L^{\Phi}(p)$ and $L^{\Phi}(q)$ coincide if and only if their norms are equivalent.

Lemma 3.8. *Let $p, q \in \mathcal{P}$, then $D(q\|p) < +\infty \iff \frac{q}{p} \in L^{\Psi_1}(p) \iff \log\frac{q}{p} \in L^1(q)$.*

The following theorem is an important improvement of Theorem 21 of Cena and Pistone (2007). Its proof can be found in Santacroce, Siri and Trivellato (2016). In particular, the novel points are the equivalence between the equality of the exponential models $\mathcal{E}(p)$ and $\mathcal{E}(q)$ and the equality of the Orlicz spaces $L^{\Phi_1}(p)$ and $L^{\Phi_1}(q)$ (statement iv)), and the integrability conditions on the ratios $\frac{q}{p}$ and $\frac{p}{q}$ (statement vi)).

Theorem 3.9. *(Portmanteau Theorem) Let $p, q \in \mathcal{P}$. The following statements are equivalent.*

i) $q \in \mathcal{E}(p)$*;*

ii) *q is connected to p by an open exponential arc;*

iii) $\mathcal{E}(p) = \mathcal{E}(q)$*;*

iv) $L^{\Phi_1}(p) = L^{\Phi_1}(q)$*;*

v) $\log\frac{q}{p} \in L^{\Phi_1}(p) \cap L^{\Phi_1}(q)$;

vi) *$\frac{q}{p} \in L^{1+\varepsilon}(p)$ and $\frac{p}{q} \in L^{1+\varepsilon}(q)$, for some $\varepsilon > 0$.*

Corollary 3.10. *$u \in \overset{\circ}{dom} K_p$ if and only if $u \in L_0^{\Phi_1}(p)$ and $e^u \in L^{1+\varepsilon}(p)$ for some $\varepsilon > 0$.*

Proof. It immediately follows from the equivalence of i) and vi) in Portmanteau Theorem. □

Corollary 3.11. *If $q \in \mathcal{E}(p)$, then the divergences $D(q\|p) < +\infty$ and $D(p\|q) < +\infty$.*

The converse of this corollary does not hold. In Santacroce, Siri and Trivellato (2016) a counterexample is shown, here we provide a simpler one.

Example 3.12. *Let $\mathcal{X} = (2, \infty)$, endowed with the probability measure μ defined by $\mu(dx) \propto \frac{1}{x^2(\log x)^3}dx$. Consider $p, q \in \mathcal{P}$ where $p(x) = 1$ and $q(x) \propto x$. In the following, $C > 0$ denotes a constant which may vary from line to line.*
Let us observe that $q \notin L^{1+\varepsilon}(p) = L^{1+\varepsilon}(\mu)$, for any $\varepsilon > 0$. In fact, if $0 < \varepsilon < 1$, we have

$$\int_{\mathcal{X}} q^{1+\varepsilon}(x)d\mu(x) = C\int_2^\infty \frac{1}{x^{1-\varepsilon}(\log x)^3}dx > C\int_2^\infty \frac{1}{x}dx = \infty.$$

Then $q \notin \mathcal{E}(p)$. On the other hand

$$D(q\|p) \le C\left(\int_2^\infty \frac{1}{x(\log x)^3}dx + \int_2^\infty \frac{1}{x(\log x)^2}dx\right) < \infty$$

and

$$D(p\|q) \le C\left(\int_2^\infty \frac{1}{x^2(\log x)^3}dx - \int_2^\infty \frac{1}{x^2(\log x)^2}dx\right) < \infty.$$

It is worth noting that, among all conditions of Portmanteau Theorem, *iv*) and *vi*) are the most useful from a practical point of view. As we will see later, the first one allows to switch from one Orlicz space to the other at one's convenience, while the second one permits to work with Lebesgue spaces.
The equality $L^{\Phi_1}(p) = L^{\Phi_1}(q)$ is important also from a geometric point of view. On the one hand, it implies that the *exponential transport mapping*, or *e-transport*, ${}^e\mathbb{U}_p^q : u \to u - \mathbb{E}_q(u)$ from $L_0^{\Phi_1}(p)$ to $L_0^{\Phi_1}(q)$ is well defined. On the other hand, it also implies that $L^{\Psi_1}(q) = \frac{p}{q}L^{\Psi_1}(p)$. As a consequence, the *mixture transport mapping*, or *m-transport*, ${}^m\mathbb{U}_p^q : v \to \frac{p}{q}v$ from $L^{\Psi_1}(p)$ to $L^{\Psi_1}(q)$ is well defined and is a Banach isomorphism (see Proposition 22 of Cena and Pistone (2007)). In the following we prove the converse statement, thereby obtaining an additional equivalent condition in Portmanteau Theorem.

Theorem 3.13. *$L^{\Phi_1}(p) = L^{\Phi_1}(q)$ if and only if the mapping*

$$\begin{array}{rccc} {}^m\mathbb{U}_p^q: & L^{\Psi_1}(p) & \longrightarrow & L^{\Psi_1}(q) \\ & v & \longmapsto & \dfrac{p}{q}v \end{array} \tag{3.6}$$

is an isomorphism of Banach spaces.

Proof. One implication is due to Proposition 22 of Cena and Pistone (2007). In order to show the converse we prove that if the mapping ${}^m\mathbb{U}_p^q$ is an isomorphism of Banach spaces then $L^{\Phi_1}(p) \subseteq L^{\Phi_1}(q)$. Let us choose $u \in L^{\Phi_1}(p)$, i.e. such that

$$N_{\Phi_1,p}(u) = \sup_{v \in L^{\Psi_1}(p), \mathbb{E}_p(\psi_1(v)) \leq 1} \mathbb{E}_p(uv) < +\infty.$$

We show that

$$N_{\Phi_1,q}(u) = \sup_{w \in L^{\Psi_1}(q), \mathbb{E}_q(\psi_1(w)) \leq 1} \mathbb{E}_q(uw) < +\infty.$$

In fact, since by hypothesis $L^{\Psi_1}(q) = \frac{p}{q}L^{\Psi_1}(p)$, we can write $w = \frac{p}{q}v$, with $v \in L^{\Psi_1}(p)$, so that

$$N_{\Phi_1,q}(u) = \sup_{v \in L^{\Psi_1}(p), \mathbb{E}_q(\psi_1(\frac{p}{q}v)) \leq 1} \mathbb{E}_p(uv).$$

From the continuity of the mapping $({}^m\mathbb{U}_p^q)^{-1} = {}^m\mathbb{U}_q^p$, we get $\overline{B_q^{\psi_1}(0,1)} \subseteq \frac{p}{q}\,\overline{B_p^{\psi_1}(0,\alpha)}$ for some $\alpha > 0$. Therefore $\mathbb{E}_q(\psi_1(\frac{p}{q}v)) \leq 1$ implies $\mathbb{E}_p(\psi_1(v) \leq \alpha$ for some $\alpha > 0$. We deduce that $N_{\Phi_1,q}(u) \leq C\, N_{\Phi_1,p}(u) < +\infty$ for a suitable constant C. □

Mixture and exponential transport mappings turn out to be useful tools in physics applications of exponential models, as one can see from the recent research production on the subject (see, e.g. Pistone (2013), Lods and Pistone (2015), Brigo and Pistone (2016)).

3.2 Densities projections and exponential sub-models

In this paragraph we give some results concerning the projection on sub-sigma-algebras, induced by conditional expectation.
Let us consider the probability space $(\mathcal{X}, \mathcal{F}, \mu)$ and a sub-sigma-algebra $\mathcal{G} \subseteq \mathcal{F}$. Let $p \in \mathcal{P}$ and denote by $p_{\mathcal{G}} = \mathbb{E}_\mu(p|\mathcal{G})$.
The following proposition states that exponential connections by arc are stable with respect to projections on $\mathcal{G}$. From a geometrical point of view, this result implies that divergence finiteness is preserved.

Proposition 3.14. *Let $p, q \in \mathcal{P}$. If $q \in \mathcal{E}(p)$ then $q_\mathcal{G} \in \mathcal{E}(p_\mathcal{G})$.*

Proof. By hypothesis, using condition $vi)$ of Portmanteau Theorem, we can find $\varepsilon > 0$ such that $\frac{q}{p} \in L^{1+\varepsilon}(p)$. Moreover $q_\mathcal{G} = \mathbb{E}_\mu(q|\mathcal{G}) = p_\mathcal{G}\mathbb{E}_p\left(\frac{q}{p}|\mathcal{G}\right)$. Then, using Jensen's inequality, we get

$$\mathbb{E}_p\left(\left(\frac{q_\mathcal{G}}{p_\mathcal{G}}\right)^{1+\varepsilon}\right) = \mathbb{E}_p\left(\left(\mathbb{E}_p\left(\frac{q}{p}\Big|\mathcal{G}\right)\right)^{1+\varepsilon}\right) \leq \mathbb{E}_p\left(\mathbb{E}_p\left(\left(\frac{q}{p}\right)^{1+\varepsilon}\Big|\mathcal{G}\right)\right) = \mathbb{E}_p\left(\left(\frac{q}{p}\right)^{1+\varepsilon}\right) < +\infty.$$

Since $\frac{q_\mathcal{G}}{p_\mathcal{G}}$ is $\mathcal{G}$-measurable, we get $\frac{q_\mathcal{G}}{p_\mathcal{G}} \in L^{1+\varepsilon}(p_\mathcal{G})$. In the same way we can prove $\frac{p_\mathcal{G}}{q_\mathcal{G}} \in L^{1+\varepsilon}(q_\mathcal{G})$ and conclude. □

Remark 3.15. The connection by exponential arcs between p and q implies that there is an exponential arc between the projections $p_\mathcal{G}$ and $q_\mathcal{G}$, but we point out that it does not follow that the arc connecting $p_\mathcal{G}$ and $q_\mathcal{G}$ is the projection of the arc connecting p and q.

Counterexample 3.16. Let us show an example where $p(\xi)$ belongs to the arc connecting p and q, but its projection $\mathbb{E}_\mu(p(\xi)|\mathcal{G})$ does not belong to the arc connecting $p_\mathcal{G}$ and $q_\mathcal{G}$.
Let us denote by $\mathcal{X} = [-1, 1]$, $\mathcal{F} = \mathcal{B}([-1, 1])$ and μ the corresponding normalized Lebesgue measure. Consider the densities $p(x) = \frac{1+x}{2}$ and $q(x) = \frac{1-x}{2}$ and the sigma-algebra $\mathcal{G}$ generated by the symmetric intervals.
It is easy to prove that $q \in \mathcal{E}(p)$, exploiting condition $vi)$ of Portmanteau Theorem. In particular, if we fix $\xi \in I \supset [0, 1]$ then $p(\xi) \propto \left(\frac{1+x}{2}\right)^{(1-\xi)}\left(\frac{1-x}{2}\right)^{\xi}$ belongs to $\mathcal{P}$.
Since p is the simmetric function of q (and viceversa), we find $p_\mathcal{G} = q_\mathcal{G} = \frac{p+q}{2} = \frac{1}{2}$, that is the uniform measure. In this case the arc between $p_\mathcal{G}$ and $q_\mathcal{G}$ reduces to a single point.
On the other hand, the projection of $p(\xi)$ is

$$\mathbb{E}_\mu(p(\xi)|\mathcal{G}) \propto \frac{1}{2}\left[\left(\frac{1+x}{2}\right)^{(1-\xi)}\left(\frac{1-x}{2}\right)^{\xi} + \left(\frac{1-x}{2}\right)^{(1-\xi)}\left(\frac{1+x}{2}\right)^{\xi}\right] \neq \frac{1}{2},$$

which means that it does not belong to the (degenerate) exponential arc between $p_\mathcal{G}$ and $q_\mathcal{G}$.

We now introduce the notion of exponential sub-model. For this purpose, it is essential to first define the concept of splitting. For the classical definition and further details see Abraham et al. (1988). In the following we give an equivalent definition suitable for our aims.

Definition 3.17. *Let V be a closed subspace of a Banach space E. We say that V splits in E if there exists a closed subspace $W \subseteq E$ such that E is the algebraic direct sum $V \oplus_a W$, i.e. $E = V + W$ and $V \cap W = \{0\}$.*

Splitting and projections are closely related as the next proposition shows.

Proposition 3.18. (see Abraham et al. (1988), Corollary 2.2.18)
V splits in E if and only if there exists a continuous linear projection Π from E to itself such that $V = Im\,\Pi$. In such case, $E = V \oplus_a Ker\,\Pi$.

Let us now introduce the notion of exponential sub-models related to a subspace V, as in Pistone and Rogantin (1999).

Definition 3.19. *Let V be a closed subspace of $L_0^{\Phi_1}(p)$. The exponential sub-model of $\mathcal{E}(p)$ related to V is the set*

$$\mathcal{E}_V(p) = \left\{q = e^{u-K_p(u)}p : u \in \overset{\circ}{dom}\, K_p \cap V\right\}. \tag{3.7}$$

In the literature, V is usually chosen to split in $L_0^{\Phi_1}(p)$ so that $\mathcal{E}_V(p)$ preserve the structure of manifold inherited by $\mathcal{E}(p)$.
Many statistical models can be seen as exponential sub-models (see, for example, Imparato and Trivellato (2009)). Here we only focus on the conditional expectation model, treated also in Pistone and Rogantin (1999).

Let $V_{\mathcal{G}}$ denote the (closed) subset of $L_0^{\Phi_1}(p)$ given by the $\mathcal{G}$-measurable random variables. The map given by the conditional expectation

$$\mathbb{E}_p[\cdot|\mathcal{G}] : L_0^{\Phi_1}(p) \to V_{\mathcal{G}}$$

is well defined. In fact, for any $u \in L_0^{\Phi_1}(p)$, $\mathbb{E}_p[u|\mathcal{G}] \in V_{\mathcal{G}}$ since it has zero expectation, is $\mathcal{G}$-measurable and, by Jensen inequality,

$$\mathbb{E}_p[\Phi_1(\alpha\mathbb{E}_p[u|\mathcal{G}])] \leq \mathbb{E}_p[\mathbb{E}_p[\Phi_1(\alpha u)|\mathcal{G}]] = \mathbb{E}_p[\Phi_1(\alpha u)] < +\infty.$$

Moreover, it is surjective, since $V_{\mathcal{G}}$ is mapped in itself, and continuous. As a consequence $W_{\mathcal{G}} = \{u \in L_0^{\Phi_1}(p) : \mathbb{E}_p[u|\mathcal{G}] = 0\}$ is closed, being the kernel of a continuous and linear map.
Finally, it is not difficult to see that any element u in $L_0^{\Phi_1}(p)$ can be uniquely written as the sum of two elements belonging respectively to $V_{\mathcal{G}}$ and $W_{\mathcal{G}}$

$$u = \mathbb{E}_p[u|\mathcal{G}] + (u - \mathbb{E}_p[u|\mathcal{G}]).$$

With this choice of $V_{\mathcal{G}}$ and $W_{\mathcal{G}}$, $\mathcal{E}_{V_{\mathcal{G}}}(p)$ as defined by (3.7) is an exponential sub-model of $\mathcal{E}(p)$.

The following results are used to show that the projection of $\mathcal{E}(p)$, induced by the conditional expectation, is the whole set $\mathcal{E}_{V_{\mathcal{G}}}(p_{\mathcal{G}})$.

Lemma 3.20. *If p is $\mathcal{G}$-measurable then $\mathcal{E}_{V_{\mathcal{G}}}(p) = \mathcal{E}(p) \cap L^0(\mathcal{X}, \mathcal{G}, \mu)$.*

Proof. If $q \in \mathcal{E}_{V_{\mathcal{G}}}(p)$, by definition $q = e^{u-K_p(u)}p \in \mathcal{E}(p)$, with $u \in V_{\mathcal{G}}$ hence $\mathcal{G}$-measurable. It immediately follows, by the assumption, that $q \in L^0(\mathcal{X}, \mathcal{G}, \mu)$. Conversely, if $q = e^{u-K_p(u)}p \in \mathcal{E}(p)$ is $\mathcal{G}$-measurable, then trivially $u \in V_{\mathcal{G}}$ and therefore $q \in \mathcal{E}_{V_{\mathcal{G}}}(p)$. □

Lemma 3.21. *Let $u \in L^0(\mathcal{X}, \mathcal{G}, \mu)$. Then*

i) $K_p(u) = K_{p_{\mathcal{G}}}(u)$,

ii) $u \in \overset{\circ}{\mathrm{dom}}\, K_p$ *if and only if* $u \in \overset{\circ}{\mathrm{dom}}\, K_{p_{\mathcal{G}}}$.

Proof. Condition $i)$ trivially follows by expectation properties. Let us prove $ii)$. Fix $u \in \overset{\circ}{\mathrm{dom}}\, K_p$. Since $\overset{\circ}{\mathrm{dom}}\, K_p$ is an open convex set containing 0, there exists $\alpha > 1$ such that αu still belongs to $\overset{\circ}{\mathrm{dom}}\, K_p$. By $i)$ we immediately get that $\alpha u \in \mathrm{dom}\, K_{p_{\mathcal{G}}}$. Observing that $0 \in \overset{\circ}{\mathrm{dom}}\, K_{p_{\mathcal{G}}}$, and since u is a convex combination of αu and 0, we deduce that u belongs to $\overset{\circ}{\mathrm{dom}}\, K_{p_{\mathcal{G}}}$. The converse can be proved in the same way and the thesis follows. □

Proposition 3.22. *The map*

$$\begin{array}{rrcl} \mathbb{E}_\mu(\cdot|\mathcal{G}) : & \mathcal{E}(p) & \longrightarrow & \mathcal{E}_{V_{\mathcal{G}}}(p_{\mathcal{G}}) \\ & q & \longmapsto & q_{\mathcal{G}} \end{array} \tag{3.8}$$

is surjective.

Proof. We first prove that the map is well defined. If $q \in \mathcal{E}(p)$ then, by Proposition 3.14, $q_{\mathcal{G}} \in \mathcal{E}(p_{\mathcal{G}})$. Since $q_{\mathcal{G}}$ is obviously $\mathcal{G}$-measurable, by Lemma 3.20, $q_{\mathcal{G}} \in \mathcal{E}_{V_{\mathcal{G}}}(p_{\mathcal{G}})$. In order to show the surjectivity, let us fix $r \in \mathcal{E}_{V_{\mathcal{G}}}(p_{\mathcal{G}})$. Then $r = e^{u-K_{p_{\mathcal{G}}}(u)}p_{\mathcal{G}}$, with $u \in \operatorname{dom}^{\circ} K_{p_{\mathcal{G}}}$. Moreover $u \in L^0(\mathcal{X}, \mathcal{G}, \mu)$, so that, by Lemma 3.21, $u \in \operatorname{dom}^{\circ} K_p$ and $K_p(u) = K_{p_{\mathcal{G}}}(u)$. As a consequence, $q := e^{u-K_p(u)}p \in \mathcal{E}(p)$. By measurability, we deduce that $q_{\mathcal{G}} = \mathbb{E}_\mu(e^{u-K_p(u)}p|\mathcal{G}) = e^{u-K_p(u)}p_{\mathcal{G}} = r$ and the thesis follows. $\square$

Remark 3.23. The surjectivity in Lemma 3.22 can be alternatively proved using condition $vi)$ of Portmanteau Theorem.
In fact, if $r \in \mathcal{E}_{V_{\mathcal{G}}}(p_{\mathcal{G}})$ then $\frac{r}{p_{\mathcal{G}}} \in L^{1+\varepsilon}(p_{\mathcal{G}})$ and $\frac{p_{\mathcal{G}}}{r} \in L^{1+\varepsilon}(r)$, i.e. $\frac{p_{\mathcal{G}}}{r} \in L^{\varepsilon}(p_{\mathcal{G}})$. Since r is $\mathcal{G}$-measurable (see Lemma 3.20), $\frac{r}{p_{\mathcal{G}}} \in L^{1+\varepsilon}(p)$ and $\frac{p_{\mathcal{G}}}{r} \in L^{\varepsilon}(p)$. Now, choosing $q = \frac{r}{p_{\mathcal{G}}}p$, we immediately get $\frac{q}{p} \in L^{1+\varepsilon}(p)$ and $\frac{p}{q} \in L^{\varepsilon}(p)$, i.e. $q \in \mathcal{E}(p)$, and $q_{\mathcal{G}} = r$.

4 Applications to finance

If applications of exponential models to physics, statistical geometry, information theory are well known in the literature, the same can not be said for applications to finance. For the first time to our knowledge, in this section we investigate some connections between martingale measures theory in finance and maximal exponential models. Besides, we see that the results illustrated in the previous sections turn out to be useful tools. In fact, in many well known works on relative entropy minimization, the minimal entropy martingale (density) measure $q^\star$ satisfies condition $vi)$ of Portmanteau Theorem. Furthermore, thanks to condition $iv)$, the equality $L^{\Phi_1}(p) = L^{\Phi_1}(q^*)$ helps us to improve some duality results. Some explicit examples in which $q^\star$ belongs to $\mathcal{E}(p)$ are also provided in the end of the section.
We endow the probability space $(\mathcal{X}, \mathcal{F}, \mu)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ satisfying the usual conditions, and $\mathcal{F} = \mathcal{F}_T$, where $T \in (0, \infty]$ is a fixed time horizon. We fix $p \in \mathcal{P}$ and consider $\mathbb{P} = \int p\, d\mu$. Let $X = (X)_{0\leq t\leq T}$ be a real-valued $(\mathbb{F}, \mathbb{P})$-locally bounded semimartingale, which represents the discounted price of a risky asset in a financial market.
We denote by $\mathcal{M}$ the set of all probability densities $q = \frac{d\mathbb{Q}}{d\mu}$, where $\mathbb{Q}$ is a $\mathbb{P}$-absolutely continuous *local martingale measure* for X, that is a probability measure absolutely continuous with respect to $\mathbb{P}$ such that X is a local $(\mathbb{F}, \mathbb{Q})$-martingale. Without the risk of misunderstanding, when saying that X is a q-local martingale, with $q \in \mathcal{M}$, we will intend that X is a local martingale with respect to $\mathbb{Q}$. Moreover, let $\mathcal{M}^e$ be the subset of $\mathcal{M}$ consisting of those densities q which are strictly positive μ-a.s. and define

$$\mathcal{M}_f = \{q \in \mathcal{M} \,:\, D(q\|p) < \infty\}, \qquad \mathcal{M}_f^e = \mathcal{M}_f \cap \mathcal{M}^e. \tag{4.9}$$

Note that, by Lemma 3.8, $\mathcal{M}_f = \mathcal{M} \cap pL^{\psi_1}(p)$ and, by Corollary 3.11,

$$\mathcal{M} \cap \mathcal{E}(p) = \mathcal{M}_f \cap \mathcal{E}(p) = \mathcal{M}_f^e \cap \mathcal{E}(p).$$

A *self-financing trading strategy* is denoted by $\theta = (\theta_t)_{0\leq t\leq T}$, where θ_t represents the number of shares invested in the asset. We assume that θ is in $L(X)$, that is an $\mathbb{F}$-predictable and X-integrable process. The stochastic integral process $W(\theta) = \theta \cdot X = \int \theta\, dX$ is then well defined and, assuming an initial capital equals to zero, $W_t(\theta)$ represents the portfolio wealth at time t. Let $U(x) = -e^{-\gamma x}$ be the exponential utility function with risk aversion parameter $\gamma \in (0, +\infty)$

(without loss of generality we will set $\gamma = 1$). Consider the related problem of maximizing the expected utility of the final wealth

$$\sup_{\theta \in \Theta} \mathbb{E}_p \left[U(W_T(\theta)) \right] \tag{4.10}$$

over a set Θ of *admissible strategies* in $L(X)$ to be specified in order to obtain a *duality result* of the form:

$$\sup_{\theta \in \Theta} \mathbb{E}_p \left[U(W_T(\theta)) \right] = U(\inf_{q \in \mathcal{M}_f} D(q \| p)). \tag{4.11}$$

It is well known that if $\mathcal{M}_f \neq \emptyset$ then there exists a unique $q^* \in \mathcal{M}_f$ that minimizes $D(q\|p)$ over all $q \in \mathcal{M}_f$ (see Theorem 2.1 in Frittelli (2000)). This q^* is called the minimal entropy martingale (density) measure. Moreover, if in addition $\mathcal{M}^e_f \neq \emptyset$ then $q^* > 0$ μ-a.s..

In the literature duality problems with general utilities have been widely explored with different classes of strategies and under various model assumptions (for general results dealing with Orlicz spaces, see, e.g. Biagini and Frittelli (2008)).

In this work, fixed a probability measure μ, we study expected utility maximization and the related dual problem and connect martingale measures to the maximal exponential model centered at p.

In the spirit of the recent literature on models uncertainty, the investigation when p ranges on a certain set of densities without any a priori choice of reference measure is an ongoing research. Nevertheless in the last section we include an example of a complete market model which shows the basic ideas in a nutshell.

In the case of an exponential utility, like in (4.11), the dual problem of finding the minimal entropy martingale measure has been treated by several authors assuming that a *Reverse Hölder condition* is satisfied; see e.g. Grandits and Rheinländer (2002), Delbaen et al. (2002), Mania et al. (2003).

In the following we introduce all the inequalities we need in the paper using a notation consistent with exponential models.

Let $q \in \mathcal{P}$ and denote the two densities projections $q_t = \mathbb{E}_\mu(q|\mathcal{F}_t)$ and $p_t = \mathbb{E}_\mu(p|\mathcal{F}_t)$.

Definition 4.1. $(\mathbf{R_{LlogL}(p)})$ *We say that q satisfies the Logarithmic Reverse Hölder inequality with respect to p, if there exists a constant $C > 0$ such that*

$$\mathbb{E}_p \left[\frac{q/p}{q_\tau/p_\tau} \log \left(\frac{q/p}{q_\tau/p_\tau} \right) \Big| \mathcal{F}_\tau \right] \leq C \ \textit{ for all stopping times } \tau \leq T. \tag{4.12}$$

Definition 4.2. $(\mathbf{R_{1+\varepsilon}(p)})$ *We say that q satisfies the $(1+\varepsilon)$-Power Reverse Hölder inequality with respect to p, for $\varepsilon > 0$, if there exists a constant $C > 0$ such that*

$$\mathbb{E}_p \left[\left(\frac{q/p}{q_\tau/p_\tau} \right)^{1+\varepsilon} \Big| \mathcal{F}_\tau \right] \leq C \ \textit{ for all stopping times } \tau \leq T. \tag{4.13}$$

Definition 4.3. $(\mathbf{A_\varepsilon(p)})$ *We say that q satisfies the ε-Muckenhoupt inequality with respect to p, for $\varepsilon > 0$, if there exists a constant $C > 0$ such that*

$$\mathbb{E}_p \left[\left(\frac{q/p}{q_\tau/p_\tau} \right)^{-\varepsilon} \Big| \mathcal{F}_\tau \right] \leq C \ \textit{ for all stopping times } \tau \leq T. \tag{4.14}$$

It is well known that if there exists $q \in \mathcal{M}_f^e$ which satisfies $R_{L\log L}(p)$, then the minimal entropy martingale measure q^* also satisfies $R_{L\log L}(p)$ (see e.g. Lemma 3.1 in Delbaen et al. (2002)). When the process X is continuous, this fact then implies that $q^* \in \mathcal{E}(p)$, as the following proposition shows.

Proposition 4.4. *Let X be a continuous semimartingale and assume there exists $q \in \mathcal{M}_f^e$ which satisfies $R_{L\log L}(p)$. Then $q^* \in \mathcal{E}(p)$.*

Proof. As observed above, q^* satisfies $R_{L\log L}(p)$. By Lemma 2.2 and Lemma 4.6 of Grandits and Rheinländer (2002), this implies that q^* also satisfies $R_{1+\varepsilon}(p)$ for some $\varepsilon > 0$. As a consequence, we get that $\frac{q^*}{p} \in L^{1+\varepsilon}(p)$ and the first integrability condition in $vi)$ of Portmanteau Theorem is satisfied. The validity of the second integrability condition follows from Proposition 5 of Doléans-Dade and Meyer (1979). In fact, if q^* satisfies $R_{1+\varepsilon}(p)$ for some $\varepsilon > 0$, then in turn it satisfies the Muckenhoupt condition $A_\varepsilon(p)$, which in particular implies $\frac{p}{q^*} \in L^\varepsilon(p)$ for some $\varepsilon > 0$. □

Remark 4.5. *Due to Proposition 3.22, if $q^* \in \mathcal{E}(p)$ then $q_t^* \in \mathcal{E}_{V_{\mathcal{F}_t}}(p_t)$ for all $t \leq T$.*

In the next subsection we provide some examples which show that the optimal solution to the dual problem q^* belongs to $\mathcal{E}(p)$. Now we investigate how this fact is reflected on the solution of the primal problem.
We start by recalling that if $\mathcal{M}_f^e \neq \emptyset$ then q^* has the form

$$q^* = c^* e^{-W_T(\theta^*)} p, \tag{4.15}$$

where $c^* = e^{D(q^* \| p)} > 0$ and $\theta^* \in L(X)$ is such that the wealth process $W(\theta^*)$ is a q^*-martingale (see Corollary 2.1 of Frittelli (2000) and Proposition 3.2 of Grandits and Rheinländer (2002)).

Theorem 4.6. *If $q^* \in \mathcal{E}(p)$ then*

i) $e^{-W_T(\theta^*)} \in L^{1+\varepsilon}(p)$ for some $\varepsilon > 0$;

ii) $W_t(\theta^*) \in L^{\Phi_1}(p)$ for all $t \in [0, T]$.

Proof. If $q^* \in \mathcal{E}(p)$ then q^* can be written in the form of

$$q^* = e^{u^* - K_p(u^*)} p, \tag{4.16}$$

where $u^* \in \overset{\circ}{\operatorname{dom}} K_p$ and $K_p(u^*) = D(p \| q^*)$. Comparing (4.16) with (4.15) we deduce that

$$W_T(\theta^*) = -u^* + D(q^* \| p) + D(p \| q^*). \tag{4.17}$$

From Corollary 3.10 we obtain $i)$ and $W_T(\theta^*) \in L^{\Phi_1}(p)$. We are left with the task of proving that $W_t(\theta^*) \in L^{\Phi_1}(p)$ for any $t < T$. In order to do this, we exploit condition $iv)$ of Portmanteau Theorem, that is the equality $L^{\Phi_1}(p) = L^{\Phi_1}(q^*)$. In fact, since the process $W(\theta^*)$ is a q^*-martingale, taking into account (4.17), we get

$$W_t(\theta^*) = -\mathbb{E}_{q^*}\left[u^* \,|\mathcal{F}_t\right] + D(q^* \| p) + D(p \| q^*), \tag{4.18}$$

so that $W_t(\theta^*) \in L^{\Phi_1}(p) = L^{\Phi_1}(q^*)$ if and only if $\mathbb{E}_{q^*}\left[u^* \,|\mathcal{F}_t\right] \in L^{\Phi_1}(q^*)$. Since u^* belongs to $L^{\Phi_1}(q^*)$, we have $\mathbb{E}_{q^*}(e^{\alpha u^*}) < +\infty$ for α varying in an open interval containing 0. Then, by Jensen inequality, it follows

$$\mathbb{E}_{q^*}\left(e^{\alpha \mathbb{E}_{q^*}[u^* \,|\mathcal{F}_t]}\right) \leq \mathbb{E}_{q^*}\left(\mathbb{E}_{q^*}\left[e^{\alpha u^*} \,|\mathcal{F}_t\right]\right) = \mathbb{E}_{q^*}(e^{\alpha u^*}) < +\infty,$$

which concludes the proof. □

Remark 4.7. Condition i) in Theorem 4.6 guarantees the integrability of the exponential optimal utility $U(W_T(\theta^*))$, and it is even stronger. On the other hand, condition ii) is a smoothness condition on the optimal wealth process, which turns out to have all moments at any time t.

Remark 4.8. *From (4.17) we deduce*

$$\mathbb{E}_p\left[W_T(\theta^*)\right] = \mathbb{E}_{q^*}[u^*] = D(q^*\|p) + D(p\|q^*), \tag{4.19}$$

by observing that u^* is centered with respect to p and $W(\theta^*)$ is a q^*-martingale.

We now go back to deal with the duality problem (4.11), assuming $R_{L\log L}(p)$ condition and $\mathcal{M}_f^e \neq \emptyset$. In Delbaen et al. (2002), under such conditions, the authors solved the problem with the two classes of strategies Θ_2 and Θ_3 defined respectively by

$$\Theta_2 = \{\theta \in L(X)\,|\, e^{-W_T(\theta)} \in L^1(p) \text{ and } W(\theta) \text{ is a } q\text{-martingale for all } q \in \mathcal{M}_f\},$$
$$\Theta_3 = \{\theta \in L(X)\,|\; W(\theta) \text{ is bounded (uniformly in time and over } \mathcal{X})\}.$$

Specifically, they show that the supremum in (4.11) is attained as maximum when computed over the class Θ_2 (the same can not be stated for Θ_3).
Furthermore, we remind that the infimum is attained and the minimal entropy martingale measure q^* is in $\mathcal{M}_f^e$.
Assuming the price process X is continuous, Proposition 4.4 shows, in fact, that q^* belongs to the maximal exponential model $\mathcal{E}(p)$. As a consequence, Theorem 4.6 proves that the maximum of the utility problem is reached on the smaller class $\widehat{\Theta}_2 \subseteq \Theta_2$, defined as

$$\begin{aligned}\widehat{\Theta}_2 = \{\theta \in L(X)\,|\, e^{-W_T(\theta)} \in L^{1+\varepsilon}(p), W_t(\theta) \in L^{\Phi_1}(p), \forall t \in [0,T] \\ \text{and } W(\theta) \text{ is a } q\text{-martingale } \forall q \in \mathcal{M}_f\}.\end{aligned} \tag{4.20}$$

Thus we can state the following corollary which improves Theorem 2.2. of Delbaen at al. (2002) in a continuous setting.

Corollary 4.9. *Let X be a continuous semimartingale and assume there exists $q \in \mathcal{M}_f^e$ which satisfies $R_{L\log L}(p)$. Then*

$$\max_{\theta\in\widehat{\Theta}_2} \mathbb{E}_p\left[U(W_T(\theta))\right] = \max_{\theta\in\Theta_2} \mathbb{E}_p\left[U(W_T(\theta))\right] = U(\min_{q\in\mathcal{M}_f} D(q\|p)) = U(\min_{q\in\mathcal{M}\cap\mathcal{E}(p)} D(q\|p)). \tag{4.21}$$

The following proposition proves that the martingality of $W(\theta)$, required in the definition of $\widehat{\Theta}_2$, is automatically satisfied when q is in $\mathcal{E}(p)$. It exploits the equality of the Orlicz spaces centered at two connected densities and is interesting on its own right.

Proposition 4.10. *If $W_T(\theta)$ belongs to $L^{\Phi_1}(p)$ then the wealth process $W(\theta)$ is a q-martingale for any $q \in \mathcal{M}\cap\mathcal{E}(p)$.*

Proof. Let $q \in \mathcal{M}\cap\mathcal{E}(p)$, then $W(\theta)$ is a local q-martingale. In order to prove that it is a q-martingale, we show that $\mathbb{E}_q(\sup_{0\le t\le T}|W_t(\theta)|) < \infty$. By Doob's maximal quadratic inequality we have

$$\mathbb{E}_q(\sup_{0\le t\le T}|W_t(\theta)|)^2 \le 4\mathbb{E}_q(|W_T(\theta)|)^2.$$

As q is connected by an exponential arc to p, by condition iv) of Portmanteau Theorem $L^{\Phi_1}(p) = L^{\Phi_1}(q)$ and therefore $W_T(\theta) \in L^{\Phi_1}(q) \subseteq L^2(q)$, from which the thesis follows. □

When q does not belong to $\mathcal{E}(p)$ the martingality of $W(\theta)$ is not a byproduct. For the optimal wealth, the proof strongly exploits the dinamicity of $R_{L\log L}(p)$ condition, as in Delbaen et al (2002).

4.1 Examples

In this section we review some well known examples of financial markets in which we can show that the minimal entropy martingale measure belongs to the maximal exponential model $\mathcal{E}(p)$ with $p = 1$. We can then apply Theorem 4.6 and conclude that the optimal wealth process belongs to the Orlicz space $L^{\Phi_1}(\mu)$.

Example 4.11. *Merton's model*

Let $\mathbb{F}$ be the augmented filtration generated by a Brownian motion $(B_t)_{0\leq t\leq T<\infty}$. Assume that the price process X follows the Black and Scholes dynamics

$$dX_t = X_t\left(\sigma dB_t + mdt\right), \quad \forall\, 0 \leq t \leq T,$$

with $\sigma > 0$ and $m \in \mathbb{R}$. We recall that the market is complete and $\mathcal{M}_f^e = \{q^*\}$. It is known that the minimal entropy martingale measure is

$$q^* = e^{-\frac{m}{\sigma}B_T - \frac{1}{2}\frac{m^2}{\sigma^2}T}.$$

It is straightforward to prove that $R_{L\log L}(p)$ is satisfied since

$$\mathbb{E}_{q^*}\left(-\frac{m}{\sigma}(B_T - B_\tau) - \frac{1}{2}\frac{m^2}{\sigma^2}(T-\tau)\Big|\mathcal{F}_\tau\right) = \frac{1}{2}\frac{m^2}{\sigma^2}(T-\tau) \leq \frac{1}{2}\frac{m^2}{\sigma^2}T,$$

which, by Proposition 4.4, implies $q^* \in \mathcal{E}(1)$, with

$$u^* = -\frac{m}{\sigma}B_T \text{ and } K_1(u^*) = D(1||q^*) = \frac{1}{2}\frac{m^2}{\sigma^2}T.$$

Nevertheless this can be directly proved by condition $vi)$ of Portmanteau Theorem because

$$\mathbb{E}((q^*)^{1+\varepsilon}) = \mathbb{E}\left(e^{-\frac{m}{\sigma}(1+\varepsilon)B_T - \frac{1}{2}\frac{m^2}{\sigma^2}(1+\varepsilon)T}\right) < +\infty \text{ and } \mathbb{E}((q^*)^{-\varepsilon}) = \mathbb{E}\left(e^{\frac{m}{\sigma}\varepsilon B_T + \frac{1}{2}\frac{m^2}{\sigma^2}\varepsilon T}\right) < +\infty.$$

By (4.17) we can compute the final wealth

$$W_T(\theta^*) = -u^* + D(q^*\|1) + D(1\|q^*) = \frac{m}{\sigma^2}\left(\sigma B_T + mT\right) = \int_0^T \theta_s^* dX_s$$

where $\theta_s^* = \frac{m}{\sigma^2}$, $\forall\, 0 \leq s \leq T$.
Finally the optimal expected utility in (4.21) turns out to be

$$\mathbb{E}\left[U(W_T(\theta^*))\right] = U(D(q^*\|1)) = -e^{-\frac{1}{2}\frac{m^2}{\sigma^2}T}.$$

Example 4.12. *BMO Martingales*

In the literature, the optimal martingale measure q^* of a duality problem often turns out to be the stochastic exponential of a BMO continuous martingale (see, for example Mania et al. (2003)). This fact is achieved assuming a Reverse Hölder condition, which in our framework, by Proposition 4.4, guarantees that q^* belongs to an exponential model.
The link between BMO continuous martingales and Reverse Hölder inequalities has been established by Kazamaki (1994), Grandits and Rheinländer (2002), and is expressed by the following

equivalent statements:
i) M is a BMO martingale, i.e. it is square integrable and $\mathbb{E}(\langle M\rangle_T - \langle M\rangle_\tau | F_\tau) \leq C$ for a fixed constant $C > 0$ and for any stopping time τ;
ii) the stochastic exponential $\mathscr{E}(M)$ is uniformly integrable and satisfies $R_{1+\varepsilon}$ for some $\varepsilon > 0$;
iii) the stochastic exponential $\mathscr{E}(M)$ is uniformly integrable and satisfies $R_{L\log L}$.

In the following, we explore the connection between BMO exponential martingales and densities in $\mathcal{E}(1)$, exploiting a setting from Mania et al. (2003). We suppose the price process X is a continuous semimartingale satisfying the structure condition, i.e. X admits the decomposition

$$X_t = X_0 + M_t + \int_0^t \lambda_s d\langle M\rangle_s, \; t \leq T < \infty,$$

where M is a continuous local martingale and λ is a predictable process such that the *mean variance tradeoff* $\int_0^T \lambda_s^2 d\langle M\rangle_s$ is finite.
Under the standing assumption $\mathcal{M}_f^e \neq \emptyset$, we introduce the value process associated to the problem of finding the minimal entropy martingale measure

$$V_t = \underset{q\in\mathcal{M}_f^e}{essinf}\, \mathbb{E}_q\left(\ln \frac{q}{q_t}\Big| F_t\right).$$

For any $q \in \mathcal{M}_f^e$, it is known that q_t can be written as the stochastic exponential of the local martingale $M^q = -\lambda\cdot M + N^q$, where $\lambda\cdot M$ stands for the stochastic integral of λ with respect to M and N^q is a local martingale strongly orthogonal to M. In addition, if the local martingale $\hat{q} = \mathscr{E}(-\lambda\cdot M)$ is a true martingale, $\hat{q}$ defines an equivalent probability measure called the minimal martingale measure for X.
The following result follows from Lemma 3.1 and Theorem 3.1 of Mania et al. (2003).

Theorem 4.13. *Assume the filtration $\mathbb{F}$ is continuous and the minimal martingale measure $\hat{q}$ exists and satisfies $R_{L\log L}$. Then, the value process V is the unique bounded solution of the semimartingale backward equation*

$$\begin{aligned} Y_t &= Y_0 - \frac{1}{2}\int_0^t \lambda_s\left(\lambda_s - 2\varphi_s\right) d\langle M\rangle_s + \frac{1}{2}\langle\tilde{L}\rangle_t + \int_0^t \varphi_s dM_s + \tilde{L}_t, \quad t < T, \\ Y_T &= 0, \end{aligned} \tag{4.22}$$

where $\varphi\cdot M + \tilde{L}$ is a BMO martingale and $\langle\tilde{L}, M\rangle = 0$. Moreover, the minimal entropy martingale measure is given by

$$q^* = \mathscr{E}_T(-\lambda\cdot M - \tilde{L}). \tag{4.23}$$

Note that q^* is written as the martingale exponential of $-(\lambda\cdot M + \tilde{L})$ which is in BMO and thus q^* belongs to $\mathcal{E}(1)$. Therefore, comparing (4.23) with (4.16), we can express also u^* in terms the BMO martingale $\tilde{L}$ appearing in the solution of BSDE (4.22). In fact,

$$u^* - K_1(u^*) = -(\lambda\cdot M)_T - \tilde{L}_T - \frac{1}{2}\left(\langle\lambda\cdot M + \tilde{L}\rangle_T\right).$$

Exploiting $\mathbb{E}(u^*) = 0$, we get

$$K_1(u^*) = D(1||q^*) = \frac{1}{2}\mathbb{E}\left(\langle\lambda\cdot M + \tilde{L}\rangle_T\right) \tag{4.24}$$

and

$$u^* = (\lambda \cdot M)_T - \tilde{L}_T - \frac{1}{2}\left(\langle \lambda \cdot M + \tilde{L} \rangle_T\right) + \frac{1}{2}\mathbb{E}\left(\langle \lambda \cdot M + \tilde{L} \rangle_T\right). \tag{4.25}$$

Since V is the solution of BSDE (4.22), we can write its initial value as

$$V_0 = \frac{1}{2}\langle \lambda \cdot M \rangle_T - \langle \varphi \cdot M, \lambda \cdot M \rangle_T - \frac{1}{2}\langle \tilde{L} \rangle_T - (\varphi \cdot M)_T - \tilde{L}_T. \tag{4.26}$$

Furthermore recall that, by definition, $V_0 = \mathbb{E}\left(q^* \ln q^*\right) = D(q^*||1)$.
By (4.17), and exploiting (4.24), (4.25) and (4.26), we can compute the final wealth

$$W_T(\theta^*) = -u^* + D(q^*\|1) + D(1\|q^*) = ((\lambda - \varphi) \cdot M)_T + \langle \lambda \cdot M \rangle_T - \langle \varphi \cdot M, \lambda \cdot M \rangle_T = \int_0^T \theta_s^* dX_s$$

where $\theta_s^* = \lambda - \varphi,\ \forall\ 0 \leq s \leq T$.
Finally the optimal expected utility in (4.21) turns out to be

$$\mathbb{E}\left[U(W_T(\theta^*))\right] = U(D(q^*\|1)) = -e^{-V_0}.$$

Example 4.14.

Here we explore two generalizations of Merton's model which represent particular cases of the previous example in a diffusion setting. The corresponding solutions of the BSDE (4.22) are characterized by the two extremal situations $\varphi = 0$ and $\tilde{L} = 0$, respectively (see, Mania et al. (2004)).

In the first model the price is described by the SDE

$$dX_t = X_t\big(\sigma(t, X_t)\ dB_t + m(t, X_t)\ dt\big), \quad \forall\ 0 \leq t \leq T, \tag{4.27}$$

where B is a standard Brownian motion, m and σ are bounded measurable functions such that the SDE (4.27) admits a unique strong solution and $\sigma^2 \geq c > 0$. It is easy to see that, by the hypotheses on the model coefficients, the mean variance tradeoff is bounded and, thus, the minimal martingale measure exists and satisfies $R_{L \log L}(1)$.
Theorem 4.13 and the Markovianity of the coefficients imply that $V_t = V(t, X_t)$; therefore the optimal solution is characterized by a PDE and the minimal entropy martingale measure q^* coincides with the minimal martingale measure $\widehat{q}$. We immediately get

$$D(q^*||1) = V_0 = \frac{1}{2}\mathbb{E}_{\widehat{q}}\left(\int_0^T \frac{m^2(t, X_t)}{\sigma^2(t, X_t)} dt\right)$$

and

$$D(1||q^*) = K_1(u^*) = \frac{1}{2}\mathbb{E}\left(\int_0^T \frac{m^2(t, X_t)}{\sigma^2(t, X_t)} dt\right),$$

where we observe that the two divergences are expectations of the mean variance tradeoff with respect to the minimal martingale measure and the reference measure, respectively. Furthermore

$$u^* = -\int_0^T \frac{m(t, X_t)}{\sigma(t, X_t)} dB_t - \frac{1}{2}\int_0^T \frac{m^2(t, X_t)}{\sigma^2(t, X_t)} dt + \frac{1}{2}\mathbb{E}\left(\int_0^T \frac{m^2(t, X_t)}{\sigma^2(t, X_t)} dt\right)$$

and the optimal strategy is

$$\theta^* = \frac{m(t, X_t)}{\sigma^2(t, X_t)} - \sigma(t, X_t) X_t \frac{\partial V}{\partial x}(t, X_t).$$

The second example is a stochastic volatility model described by the SDEs

$$\begin{aligned} dX_t =& X_t\big(\sigma(t, Y_t) dB_t + m(t, Y_t) dt\big), \\ dY_t =& \sigma^{\perp}(t, Y_t) dB_t^{\perp} + b(t, Y_t) dt, \quad \forall\, 0 \le t \le T, \end{aligned} \tag{4.28}$$

where B and $B^{\perp}$ are independent standard Brownian motions and the coefficients satisfy some regularity conditions (similarly to the previous example) such that the SDEs (4.28) admits a unique strong solution. In this case the market price of risk $\frac{m(t,Y_t)}{\sigma(t,Y_t)}$ is $\mathbb{F}^{B^{\perp}}$ adapted and, as a consequence, $V_t = V(t, Y_t)$. The optimal strategy boils down to the ratio $\theta^* = \frac{m(t,Y_t)}{\sigma^2(t,Y_t)}$ which is function of the exogenous Y_t, thus, generalizing the classical Merton strategy.

Example 4.15.

Many examples where the optimal solution q^* belongs to the maximal exponential model can be found in the literature, even for non locally bounded X. We analyze one of them taken from Biagini and Frittelli (2008). Let N be a Poisson Process of parameter $\lambda > 0$ with jump times $\mathbb{T} = (T_j)_{j\ge1}$, $T_0 = 0$. Denote by $Y_0 = 0$ and $(Y_j)_{j\ge1}$ a sequence of i.i.d. random variables independent from $\mathbb{T}$, with density $f(y) = \frac{\nu}{2} e^{-\nu|y-1|}$, $\nu > 0$. We define the price process

$$X_t = \sum_{j: T_j \le t\wedge T} Y_j$$

where $0 \le T < +\infty$.
In this case q^* turns out to be the optimal solution of the dual problem on a generalization of the set $\mathcal{M}_f$ and has the form

$$q^* = exp\Bigg(-a^* X_T - \lambda T \left(\frac{\nu^2}{\nu^2 - (a^*)^2} e^{a^*} - 1\right)\Bigg),$$

where $a^* = \sqrt{1+\nu^2} - 1$.
It can be checked, exploiting condition $vi)$ of Portmanteau Theorem, that $q^* \in \mathcal{E}(1)$. In fact, $q^* \in L^{1+\epsilon}(\mu)$ and $\frac{1}{q^*} \in L^{\epsilon}(\mu)$ for some $\epsilon > 0$ if and only if $M_{X_T}(-a^*(1+\epsilon)) < \infty$ and $M_{X_T}(a^*\epsilon) < \infty$, where M_{X_T} is the moment generating function of X_T given by

$$M_{X_T}(s) = e^s \frac{\nu^2}{\nu^2 - s^2}, \quad -\nu < s < \nu.$$

Both conditions are verified for any $0 < \epsilon < \frac{\nu}{a^*} - 1$.

Example 4.16.

Here we show an example where the optimal martingale measure does not belong to the maximal exponential model $\mathcal{E}(1)$. It is borrowed from Grandits and Rheinländer (2002) and adapted to the notation of exponential models.

Let $\mathbb{F}$ be the augmented filtration generated by a Brownian motion $(B_t)_{0\leq t<\infty}$. Consider the price process X given by

$$X_t = B_t^\tau - t \wedge \tau, \quad \forall 0 \leq t < \infty,$$

where $\tau = \inf\{t \geq 0 : B_t = 1\}$ is a stopping time such that $\mathbb{P}(\tau < \infty) = 1$ and $\mathbb{E}(\tau) = \infty$.
It can be checked that $\mathcal{M}_f^e = \{q^*\}$, where

$$q^* = e^{B_\tau - \frac{1}{2}\tau} = e^{1-\frac{1}{2}\tau}.$$

Since q^* is bounded, obviously $q^* \in L^{1+\varepsilon}(\mu)$ for some $\varepsilon > 0$. However, $\frac{1}{q^*} \notin L^\varepsilon(\mu)$ for any $\varepsilon > 0$ because

$$\mathbb{E}((q^*)^{-\varepsilon}) = \mathbb{E}(e^{-\varepsilon+\frac{1}{2}\varepsilon\tau}) \geq e^{-\varepsilon+\frac{1}{2}\varepsilon\mathbb{E}(\tau)} = \infty.$$

By condition $vi)$ of Portmanteau Theorem we conclude that $q^* \notin \mathcal{E}(1)$.
Note that

$$D(q^*\|1) = \mathbb{E}_{q^*}\left(1 - \frac{1}{2}\tau\right).$$

Since Kullback-Leibler divergence is always non-negative, we may conclude that $\mathbb{E}_{q^*}(\tau) < \infty$ and therefore $D(q^*\|1) < \infty$. On the other hand,

$$D(1\|q^*) = \mathbb{E}\left(\frac{1}{2}\tau - 1\right) = \infty,$$

which, by Corollary 3.11, leads to the same conclusion $q^* \notin \mathcal{E}(1)$.
As observed by Delbaen et al. (2002), in this example the duality result (4.21) holds for Θ_2 even though $R_{L\log L}(p)$ is not satisfied.

4.2 An example under model uncertainty

In this section we present a simple example which shows the possible impact of maximal exponential models on financial applications in an uncertainty framework.
We refer to the Merton's model described in Example 4.11.
Let us recall that the unique martingale measure q^* is the minimal entropy martingale measure with respect to $p = 1$ and belongs to $\mathcal{E}(1)$.
Due to the Portmanteau Theorem, item $iii)$, chosen a density r connected to $p = 1$ by an open arc, then $q^* \in \mathcal{E}(r) = \mathcal{E}(1)$ and trivially, since the set of the equivalent martingale measures is a singleton, it turns out to be the minimal entropy martingale measure with respect to r.
Let us now consider the generical element $p(\xi)$ of the open arc connecting r and p. From Definition 3.1, we recall that $p(\xi) \propto r^\xi$, where ξ ranges in an open interval I strictly containing $[0, 1]$.
Our aim is to solve the min-max problem

$$\min_{\xi\in I} \max_{\theta\in\Theta_2(\xi)} \mathbb{E}_{p(\xi)}\left[U(W_T(\theta))\right],$$

which, from (4.21), is expressed through the dual problem by

$$\min_{\xi\in I} \max_{\theta\in\Theta_2(\xi)} \mathbb{E}_{p(\xi)}\left[U(W_T(\theta))\right] = U(\min_{\xi\in I} \min_{q\in\mathcal{M}_f(\xi)} D(q\|p(\xi))) = U(\min_{\xi\in I} D(q^*\|p(\xi))). \tag{4.29}$$

After some computations, we obtain

$$D(q^*\|p(\xi)) = D(q^*\|1) - \xi\mathbb{E}_{q^*}(\log r) + \log\mathbb{E}(r^\xi) = D(q^*\|1) - \xi\mathbb{E}_{q^*}(u) + \log\mathbb{E}(e^{\xi u}), \tag{4.30}$$

where u is the random variable characterizing the representation of r in the exponential model $\mathcal{E}(1)$, that is $r = e^{u-K_1(u)}$. Recall also the representation of q^* in $\mathcal{E}(1)$, which is $q^* = e^{u^*-K_1(u^*)}$, where $u^* = -\frac{m}{\sigma}B_T$.
A particular choice of r is made by selecting u such that (u, u^*) is a non-degenerate Gaussian vector with covariance matrix $\begin{pmatrix} \gamma^2 & c \\ c & \frac{m^2}{\sigma^2}T \end{pmatrix}$, picking arbitrarily the parameters γ^2 and c.
In this case we can choose $I = \mathbb{R}$. For any $\xi \in I$, the divergence takes the form

$$D(q^*\|p(\xi)) = D(q^*\|1) - \xi c + \frac{1}{2}\xi^2\gamma^2 \tag{4.31}$$

and it is minimized by $\bar{\xi} = \frac{c}{\gamma^2}$.
Therefore the solution of the dual problem in (4.29) is given by

$$U\left(D(q^*\|p(\bar{\xi}))\right) = U\left(D(q^*||1) - \frac{1}{2}\frac{c^2}{\gamma^2}\right) = U\left(\frac{1}{2}\frac{m^2}{\sigma^2}\left(T - \frac{cov(u, B_T)^2}{var(u)}\right)\right). \tag{4.32}$$

The representation in $\mathcal{E}(1)$ of the corresponding optimal density is then $p(\bar{\xi}) = e^{\bar{u}-K_1(\bar{u})}$, where

$$\bar{u} = \frac{c}{\gamma^2}u = -\frac{m}{\sigma}\,\frac{cov(u, B_T)}{var(u)}u, \qquad K_1(\bar{u}) = \frac{1}{2}\frac{c^2}{\gamma^2} = \frac{1}{2}\frac{m^2}{\sigma^2}\,\frac{cov(u, B_T)^2}{var(u)}. \tag{4.33}$$

In the primal problem, the optimal wealth can be explicitly identified adapting (4.17) for $p = p(\bar{\xi})$:

$$\begin{aligned} W_T(\theta^*(\bar{\xi})) &= -u^*(\bar{\xi}) + D(q^*\|p(\bar{\xi})) + D(p(\bar{\xi})\|q^*) = -u^* + \bar{u} - \bar{\xi}c + \frac{m^2}{\sigma^2}T \\ &= W_T(\theta^*) - \frac{m}{\sigma}\,\frac{cov(u, B_T)}{var(u)}\left(u + \frac{m}{\sigma}cov(u, B_T)\right). \end{aligned} \tag{4.34}$$

In some cases, the optimal strategy can be explicitly computed from (4.34). If, for instance, $u = \int_0^T \varphi(s)\, dB_s$, with $\varphi \in L^2([0, T])$, then

$$(\theta^*(\bar{\xi}))_t = \theta_t^* + \bar{\xi}\,\frac{\varphi(t)}{\sigma X_t} = \frac{m}{\sigma}\left(1 - \left(\frac{\int_0^T \varphi(s)\, ds}{\int_0^T \varphi^2(s)\, ds}\right)\,\frac{\varphi(t)}{\sigma X_t}\right). \tag{4.35}$$

Acknowledgements.
The authors are indebted to the Associate Editor and to the referee for their comments and suggestions, which led to a significant improvement of the manuscript.

References

[1] Abraham, R., Marsden, J.E. and Ratiu, T. Manifold, Tensor Analysis and Applications. Second edition. Springer-Verlag, New York, 1988.

[2] Biagini, S. and Frittelli, M. A unified framework for utility maximization problems: An Orlicz space approach. *Ann. Appl. Probab.* 18 (3), pp. 929-966, 2008.

[3] Brigo, D. and Pistone G. Projection based dimensionality reduction for measure valued evolution equations in statistical manifolds. arXiv:1601.04189v3, 2016.

[4] Cena, A. and Pistone, G. Exponential Statistical Manifold. *AISM* 59, pp. 27-56, 2007.

[5] Cheridito, P. and Li, T. Risk measures on Orlicz hearts. *Math. Finance* 19 (2), pp. 189-214, 2009.

[6] Delbaen, F., Grandits, P., Rheinländer, T., Samperi, D., Schweizer, M. and Stricker, C. Exponential hedging and entropic penalties. *Math. Finance* 12 (2), pp. 99-123, 2002.

[7] Doléans-Dade, C. and Meyer, P.A. Inégalités de normes avec poids. *In Sém. Prob. XIII, Lecture Notes in Mathematics* 721, Springer-Verlag New York, pp. 260-280, 1979.

[8] Frittelli, M. The minimal entropy martingale measure and the valuation problem in incomplete markets. *Math. Finance* 10 (1), pp. 39-52, 2000.

[9] Grandits, P. and Rheinländer, T. On the minimal entropy martingale measure. *Ann. Probab.* 30 (3), pp. 1003-1038, 2002.

[10] Imparato, D. and Trivellato, B. Geometry of Extendend Exponential Models. *In Algebraic and Geometric Methods in Statistics*, pp. 307-326, 2009.

[11] Kazamaki, N. Continuous Exponential Martingales and BMO. *Lecture Notes in Mathematics* 1579, Springer-Verlag New York, 1994.

[12] Lods, B. and Pistone G. Information geometry formalism for the spatially homogeneous Boltzmann equation. *Entropy* 17, pp. 4323-4363, 2015.

[13] Mania M., Santacroce M. and Tevzadze R. A semimartingale BSDE related to the minimal entropy martingale measure. *Finance Stoch.* 7 (3), pp. 385-402, 2003.

[14] Mania M., Santacroce M. and Tevzadze R. The Bellman equation related to the minimal entropy martingale measure. *Georgian Math. J.* 11 (1), pp. 125-135, 2004.

[15] Pistone G. Examples of the application of nonparametric information geometry to statistical physics. *Entropy* 15, pp. 4042-4065, 2013.

[16] Pistone G. and Rogantin M.P. The exponential statistical manifold: mean parameters, orthogonality and space transformations. *Bernoulli* 5 (4), pp. 571-760, 1999.

[17] Pistone G. and Sempi C. An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one. *Ann. Stat.* 23 (5), pp.1543-1561, 1995.

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

[19] Rao M.M. and Ren Z.D. Applications of Orlicz Spaces. Marcel Dekker Inc., New York, 2002.

[20] Santacroce M., Siri P. and Trivellato B. New results on mixture and exponential models by Orlicz spaces. *Bernoulli* 22 (3), pp. 1431-1447, 2016.