J. Appl. Prob. **55,** 682–700 (2018)
*doi:*10.1017/jpr.2018.45

EXPONENTIAL MODELS BY ORLICZ SPACES AND APPLICATIONS

MARINA SANTACROCE,* **
PAOLA SIRI * *** AND
BARBARA TRIVELLATO,* **** *Politecnico di Torino*

Abstract

The geometric structure of the nonparametric 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.

Keywords: Orlicz space; maximal exponential model; exponential utility maximization; minimal entropy martingale measure; reverse Hölder condition

2010 Mathematics Subject Classification: Primary 46E30; 46N30; 62B10; 91G80

1. Introduction

Statistical exponential models built on Orlicz spaces arise in several fields, such as differential geometry, algebraic statistics, and information theory. To the best of the authors' knowledge, their application to finance has not been investigated, although the use of Orlicz spaces in utility maximization and in risk measure theory is known; see, e.g. Cheridito and Li (2009), and 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 of Santacroce *et al.* (2016).

The theory of nonparametric maximal exponential models centered at a given positive density p was initiated by the work of 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 was endowed with a structure of an exponential Banach manifold. Such a manifold setting turns out to be well suited for applications in physics as some authors have recently demonstrated; 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 vice versa (by *open*, we essentially mean that the two densities are not the extremal points of the arc). Santacroce *et al.* (2016) proved the equivalence between the equality of the maximal exponential models centered at two (connected) densities p and q and the equality of the Orlicz spaces referred to the same densities.

This work is a natural continuation of Santacroce *et al.* (2016) and, moreover, it includes applications to finance; see Santacroce *et al.* (2017), (2018) for related topics.

Received 13 February 2017; revision received 5 February 2018.
* Postal address: Dipartimento di Scienze Matematiche G. L. Lagrange, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy.
** Email address: marina.santacroce@polito.it
*** Email address: paola.siri@polito.it
**** Email address: barbara.trivellato@polito.it

The paper is essentially composed of two parts. In the first part we provide new theoretical results concerning exponential models which can be useful in understanding 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 density 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.

In the second part of the paper we address 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 so by exploiting the equivalent conditions proved in Santacroce *et al.* (2016). We conclude with Subsection 4.1, where our results are illustrated in some classical examples of financial markets taken from the literature, and Subsection 4.2 where an application under model uncertainty is shown.

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, we refer the reader 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 $f(x) = |x|^a/a$ with $a > 1$ and, consequently, Orlicz spaces are generalizations of the Lebesgue spaces $L^a(\mu)$. The definition of a 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 $\Phi\colon \mathbb{R} \to [0, +\infty]$ 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 $\Psi(y) = \sup_{x\in\mathbb{R}}\{xy - \Phi(x)\}$ for all $y \in \mathbb{R}$ and it is itself a Young function.

Let L^0 denote the set of all measurable functions $u\colon X \to \mathbb{R}$ 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

$$L^\Phi(\mu) = \left\{u \in L^0\colon \text{there exists } \alpha > 0 \text{ such that } \int_{\mathcal{X}} \Phi(\alpha u)\,\mathrm{d}\mu < +\infty\right\}.$$

https://doi.org/10.1017/jpr.2018.45 Published online by Cambridge University Press

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*

$$\|u\|_{\Phi,\mu} = \inf\left\{k > 0\colon \int_{\mathcal{X}} \Phi\left(\frac{u}{k}\right) \mathrm{d}\mu \le 1\right\}.$$

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 ball. Observe that

$$u \in B(0,1) \iff \text{there exists } \alpha > 1 \text{ such that } \int_{\mathcal{X}} \Phi(\alpha u)\,\mathrm{d}\mu \le 1,$$

$$u \in \overline{B(0,1)} \iff \int_{\mathcal{X}} \Phi(u)\,\mathrm{d}\mu \le 1.$$

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

$$N_{\Phi,\mu}(u) = \sup_{v \in L^{\Psi}(\mu)\colon \int_{\mathcal{X}} \Psi(v)\,\mathrm{d}\mu \le 1} \left\{\int_{\mathcal{X}} |uv|\,\mathrm{d}\mu\right\},$$

where Ψ is the conjugate function of Φ.

It is worth recalling 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 exist $x_0 > 0$, and two positive constants $c_1 < c_2$ such that, for all $x \ge x_0$,

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

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 by $\mathcal{P}$ the set of all densities which are positive μ-almost surely. Moreover, we use the notation $\mathbb{E}_p$ to indicate the expectation with respect to $p\,\mathrm{d}\mu$ 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 $\Phi_2(x) = \mathrm{e}^{|x|} - |x| - 1$.

We recall that the conjugate function of $\Phi_1(x)$ is $\Psi_1(y) = \int_0^y \sinh^{-1}(t)\,\mathrm{d}t$, which, in 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 associated to Φ_1, defined with respect to the measure induced by p, i.e.

$$L^{\Phi_1}(p) = \{u \in L^0\colon \text{ there exists } \alpha > 0 \text{ such that } \mathbb{E}_p(\Phi_1(\alpha u)) < +\infty\}.$$

It is worth noting that in order to prove that a random variable u belongs to $L^{\Phi_1}(p)$, it is sufficient to check that $\mathbb{E}_p(\mathrm{e}^{\alpha u}) < +\infty$ with α belonging to an open interval containing 0.

Finally, we note the following chain of inclusions:

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

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 $\xi \in I$.

In the following proposition we provide an equivalent definition of exponential connection by arc. Its proof can be found in Santacroce *et al.* (2016).

Proposition 3.1. *It holds that $p, q \in \mathcal{P}$ 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 $p(\xi) \propto \mathrm{e}^{\xi u} p$ belongs to $\mathcal{P}$ for every $\xi \in I$, and $p(0) = p$ and $p(1) = q$.*

The connection 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 the maximal exponential model. In the next section, we prove that the maximal exponential model at p coincides with the set of all densities $q \in \mathcal{P}$ which are connected to p by an open exponential arc.

Denote

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

Definition 3.2. The cumulant generating functional is the map

$$K_p : L_0^{\Phi_1}(p) \to [0, +\infty], \qquad u \to \log \mathbb{E}_p(\mathrm{e}^u).$$

Theorem 3.1. *The cumulant generating functional K_p satisfies the following properties:*

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

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

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

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

Proof. See Pistone and Sempi (1995). □

Definition 3.3. For every density $p \in \mathcal{P}$, the maximal exponential model at p is

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

Remark 3.1. 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 $D(q \parallel p)$ 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).

https://doi.org/10.1017/jpr.2018.45 Published online by Cambridge University Press

Proposition 3.2. *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 3.1. *Let $p, q \in \mathcal{P}$. Then*

$$D(q \parallel p) < +\infty \iff \frac{q}{p} \in L^{\Psi_1}(p) \iff \log \frac{q}{p} \in L^1(q).$$

The following theorem is an important improvement of Cena and Pistone (2007, Theorem 21). Its proof can be found in Santacroce *et al.* (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 $L^{\Phi_1}(q)$ (Theorem 3.2(iv)), and the integrability conditions on the ratios q/p and p/q (Theorem 3.2(vi)).

Theorem 3.2. *(Portmanteau theorem.) Let $p, q \in \mathcal{P}$. The following 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) $\log(q/p) \in L^{\Phi_1}(p) \cap L^{\Phi_1}(q)$;

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

Corollary 3.1. *It holds that $u \in \mathring{\mathrm{dom}}\, K_p$ if and only if $u \in L_0^{\Phi_1}(p)$ and $\mathrm{e}^u \in L^{1+\varepsilon}(p)$ for some $\varepsilon > 0$.*

Proof. It immediately follows from the equivalence of (i) and (vi) in Theorem 3.2. □

Corollary 3.2. *If $q \in \mathcal{E}(p)$ then the divergences $D(q \parallel p) < +\infty$ and $D(p \parallel q) < +\infty$.*

The converse of this corollary does not hold. In Santacroce *et al.* (2016) a counterexample was shown; here we provide a simpler one.

Example 3.1. Let $\mathcal{X} = (2, \infty)$ be endowed with the probability measure μ defined by $\mu(\mathrm{d}x) \propto 1/(x^2(\log x)^3)\,\mathrm{d}x$. 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.

Observe that $q \notin L^{1+\varepsilon}(p) = L^{1+\varepsilon}(\mu)$ for any $\varepsilon > 0$. In fact, if $0 < \varepsilon < 1$, we have

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

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

$$D(q \parallel p) \le C\bigg(\int_2^{\infty} \frac{1}{x(\log x)^3}\,\mathrm{d}x + \int_2^{\infty} \frac{1}{x(\log x)^2}\,\mathrm{d}x\bigg) < \infty$$

and

$$D(p \parallel q) \le C\bigg(\int_2^{\infty} \frac{1}{x^2(\log x)^3}\,\mathrm{d}x - \int_2^{\infty} \frac{1}{x^2(\log x)^2}\,\mathrm{d}x\bigg) < \infty.$$

It is worth noting that, among all the conditions of Theorem 3.2, (iv) and (vi) are the most useful from a practical point of view. As we will see later, the first one allows us to switch from one Orlicz space to another at our convenience, while the second one permits us 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, ${}^{\mathrm{e}}\mathbb{U}_p^q\colon u \to u - \mathbb{E}_q(u)$ from $L_0^{\Phi_1}(p)$ to $L_0^{\Phi_1}(q)$ is well defined. On the other hand, it also implies that $L^{\Psi_1}(q) = (p/q)L^{\Psi_1}(p)$. As a consequence, the *mixture transport mapping,* or m-transport, ${}^{\mathrm{m}}\mathbb{U}_p^q\colon v \to (p/q)v$ from $L^{\Psi_1}(p)$ to $L^{\Psi_1}(q)$ is well defined and is a Banach isomorphism; see Cena and Pistone (2007, Proposition 22). In the following, we prove the converse statement, thereby obtaining an additional equivalent condition in Theorem 3.2.

Theorem 3.3. *It holds that $L^{\Phi_1}(p) = L^{\Phi_1}(q)$ if and only if the mapping*

$$ {}^{\mathrm{m}}\mathbb{U}_p^q\colon L^{\Psi_1}(p) \to L^{\Psi_1}(q), \qquad v \mapsto \frac{p}{q}v $$

is an isomorphism of Banach spaces.

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

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

We show that

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

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

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

From the continuity of the mapping $({}^{\mathrm{m}}\mathbb{U}_p^q)^{-1} = {}^{\mathrm{m}}\mathbb{U}_q^p$, we obtain $\overline{B_q^{\psi_1}(0,1)} \subseteq (p/q)\overline{B_p^{\psi_1}(0,\alpha)}$ for some $\alpha > 0$. Therefore, $\mathbb{E}_q(\psi_1((p/q)v)) \le 1$ implies $\mathbb{E}_p(\psi_1(v)) \le \alpha$ for some $\alpha > 0$. We deduce that $N_{\Phi_1,q}(u) \le C N_{\Phi_1,p}(u) < +\infty$ 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 literature on the subject; see, e.g. Pistone (2013), Lods and Pistone (2015), and Brigo and Pistone (2016).

3.2. Density projections and exponential sub-models

We now state some results concerning the projection on sub-σ-algebras induced by conditional expectation.

We consider the probability space $(\mathcal{X}, \mathcal{F}, \mu)$ and a sub-σ-algebra $\mathcal{G} \subseteq \mathcal{F}$. Let $p \in \mathcal{P}$ and denote $p_{\mathcal{G}} = \mathbb{E}_\mu(p \mid \mathcal{G})$.

In the following proposition we state 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.

https://doi.org/10.1017/jpr.2018.45 Published online by Cambridge University Press

Proposition 3.3. *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 Theorem 3.2(vi), we can find $\varepsilon > 0$ such that $q/p \in L^{1+\varepsilon}(p)$. Moreover, $q_{\mathcal{G}} = \mathbb{E}_\mu(q \mid \mathcal{G}) = p_{\mathcal{G}} \mathbb{E}_p(q/p \mid \mathcal{G})$. Then, using Jensen's inequality, we obtain

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

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

Remark 3.2. 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.

Example 3.2. (*Counterexample.*) We now provide an example where $p(\xi)$ belongs to the arc connecting p and q, but its projection $\mathbb{E}_\mu(p(\xi) \mid \mathcal{G})$ does not belong to the arc connecting $p_{\mathcal{G}}$ and $q_{\mathcal{G}}$.

Denote $\mathcal{X} = [-1, 1]$, $\mathcal{F} = \mathcal{B}([-1, 1])$, and μ the corresponding normalized Lebesgue measure. Consider the densities $p(x) = \frac{1}{2}(1 + x)$ and $q(x) = \frac{1}{2}(1 - x)$, and the σ-algebra $\mathcal{G}$ generated by the symmetric intervals.

It is easy to prove that $q \in \mathcal{E}(p)$ by exploiting Theorem 3.2(vi). In particular, if we fix $\xi \in I \supset [0, 1]$ then $p(\xi) \propto (1 + x)^{1-\xi}(1 - x)^{\xi}$ belongs to $\mathcal{P}$.

Since p is the symmetric function of q (and vice versa), we have $p_{\mathcal{G}} = q_{\mathcal{G}} = \frac{1}{2}(p+q) = \frac{1}{2}$, i.e. 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

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

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-models. 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 state an equivalent definition suitable for our aims.

Definition 3.4. 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 we show in the next proposition.

Proposition 3.4. (Abraham *et al.* (1988, Corollary 2.2.18).) *It holds that V splits in E if and only if there exists a continuous linear projection Π from E to itself such that $V = \operatorname{Im}\Pi$. In such a case, $E = V \oplus_a \ker \Pi$.*

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

https://doi.org/10.1017/jpr.2018.45 Published online by Cambridge University Press

Definition 3.5. 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

$$\mathcal{E}_V(p) = \{q = \mathrm{e}^{u-K_p(u)}p \colon u \in \overset{\circ}{\mathrm{dom}} K_p \cap V\}. \tag{3.1}$$

In the literature, V is usually chosen to split in $L_0^{\Phi_1}(p)$ so that $\mathcal{E}_V(p)$ preserves the manifold structure inherited by $\mathcal{E}(p)$.

Many statistical models can be seen as exponential sub-models; see, e.g. Imparato and Trivellato (2009). Here we focus only 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

$$\mathbb{E}_p(\cdot \mid \mathcal{G}) \colon L_0^{\Phi_1}(p) \to V_{\mathcal{G}}$$

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

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

Moreover, it is surjective since $V_{\mathcal{G}}$ is mapped in itself and continuous. As a consequence, $W_{\mathcal{G}} = \{u \in L_0^{\Phi_1}(p) \colon \mathbb{E}_p(u \mid \mathcal{G}) = 0\}$ is closed since it is 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 $W_{\mathcal{G}}$, i.e.

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

With this choice of $V_{\mathcal{G}}$ and $W_{\mathcal{G}}$, $\mathcal{E}_{V_{\mathcal{G}}}(p)$ as defined by (3.1) 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 3.2. *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 $q = \mathrm{e}^{u-K_p(u)}p \in \mathcal{E}(p)$ 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 $q = \mathrm{e}^{u-K_p(u)}p \in \mathcal{E}(p)$ is $\mathcal{G}$-measurable then trivially $u \in V_{\mathcal{G}}$ and, therefore, $q \in \mathcal{E}_{V_{\mathcal{G}}}(p)$. □

Lemma 3.3. *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. (i) This trivially follows by expectation properties.

(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 $\overset{\circ}{\mathrm{dom}} K_p$. From (i), we immediately see 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 result follows, completing the proof. □

https://doi.org/10.1017/jpr.2018.45 Published online by Cambridge University Press

Proposition 3.5. *The map*

$$\mathbb{E}_\mu(\cdot \mid \mathcal{G})\colon \mathcal{E}(p) \to \mathcal{E}_{V_\mathcal{G}}(p_\mathcal{G}), \qquad q \mapsto q_\mathcal{G}$$

is surjective.

Proof. We first prove that the map is well defined. If $q \in \mathcal{E}(p)$ then, by Proposition 3.3, $q_\mathcal{G} \in \mathcal{E}(p_\mathcal{G})$. Since $q_\mathcal{G}$ is obviously $\mathcal{G}$-measurable, by Lemma 3.2, $q_\mathcal{G} \in \mathcal{E}_{V_\mathcal{G}}(p_\mathcal{G})$. In order to show the surjectivity, we fix $r \in \mathcal{E}_{V_\mathcal{G}}(p_\mathcal{G})$. Then $r = \mathrm{e}^{u-K_{p_\mathcal{G}}(u)} p_\mathcal{G}$ with $u \in \overset{\circ}{\operatorname{dom}} K_{p_\mathcal{G}}$. Moreover, $u \in L^0(\mathcal{X}, \mathcal{G}, \mu)$ so that, by Lemma 3.3, $u \in \overset{\circ}{\operatorname{dom}} K_p$ and $K_p(u) = K_{p_\mathcal{G}}(u)$. As a consequence, $q := \mathrm{e}^{u-K_p(u)} p \in \mathcal{E}(p)$. By measurability, we deduce that $q_\mathcal{G} = \mathbb{E}_\mu(\mathrm{e}^{u-K_p(u)} p \mid \mathcal{G}) = \mathrm{e}^{u-K_p(u)} p_\mathcal{G} = r$ and the result follows. □

Remark 3.3. The surjectivity in Lemma 3.5 can be alternatively proved using Theorem 3.2(vi).

In fact, if $r \in \mathcal{E}_{V_\mathcal{G}}(p_\mathcal{G})$ then $r/p_\mathcal{G} \in L^{1+\varepsilon}(p_\mathcal{G})$ and $p_\mathcal{G}/r \in L^{1+\varepsilon}(r)$, i.e. $p_\mathcal{G}/r \in L^{\varepsilon}(p_\mathcal{G})$. Since r is $\mathcal{G}$-measurable (see Lemma 3.2), $r/p_\mathcal{G} \in L^{1+\varepsilon}(p)$ and $p_\mathcal{G}/r \in L^{\varepsilon}(p)$. Now, choosing $q = (r/p_\mathcal{G})p$, we immediately obtain $q/p \in L^{1+\varepsilon}(p)$ and $p/q \in L^{\varepsilon}(p)$, 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, and information theory are well known in the literature, the same cannot be said for applications to finance. To the best of the authors' knowledge, this paper is the first in which the connection between martingale measure theory in finance and maximal exponential models is investigated. 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 Theorem 3.2(vi). Furthermore, thanks to Theorem 3.2(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 at the end of the section.

We endow the probability space $(\mathcal{X}, \mathcal{F}, \mu)$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ 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 $\mathbb{P} = \int p \, \mathrm{d}\mu$. Let $X = (X)_{0\le t\le T}$ 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 $q = \mathrm{d}\mathbb{Q}/\mathrm{d}\mu$, where $\mathbb{Q}$ is a $\mathbb{P}$-absolutely continuous *local martingale measure* for X, i.e. 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 any misunderstanding, when we say that X is a q-local martingale with $q \in \mathcal{M}$, we mean that X is a local martingale with respect to $\mathbb{Q}$. Moreover, let $\mathcal{M}^\mathrm{e}$ be the subset of $\mathcal{M}$ consisting of those densities q which are strictly positive μ-almost surely and define

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

Note that, by Lemma 3.1, $\mathcal{M}_f = \mathcal{M} \cap pL^{\psi_1}(p)$ and, by Corollary 3.2,

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

A *self-financing trading strategy* is denoted by $\theta = (\theta_t)_{0\le t\le T}$, where θ_t represents the number of shares invested in the asset. We assume that θ is in $L(X)$, i.e. an $\mathbb{F}$-predictable and X-integrable process. The stochastic integral process $W(\theta) = \theta \cdot X = \int \theta \, \mathrm{d}X$ is then well defined and, assuming an initial capital equal to 0, $W_t(\theta)$ represents the portfolio wealth at time t.

https://doi.org/10.1017/jpr.2018.45 Published online by Cambridge University Press

Let $U(x) = -\mathrm{e}^{-\gamma x}$ be the exponential utility function with risk aversion parameter $\gamma \in (0, +\infty)$ (without loss of generality, we set $\gamma = 1$). Consider the related problem of maximizing the expected utility of the final wealth

$$\sup_{\theta \in \Theta} \mathbb{E}_p(U(W_T(\theta)))$$

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

$$\sup_{\theta \in \Theta} \mathbb{E}_p(U(W_T(\theta))) = U\Big(\inf_{q \in \mathcal{M}_f} D(q \,\|\, p)\Big). \tag{4.1}$$

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

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, given a fixed probability measure μ, we study the 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 model uncertainty, the investigation when p ranges on a certain set of densities without any *a priori* choice of reference measure is our ongoing research interest. Nevertheless, in the last subsection we include an example of a complete market model in which we summarize the basic ideas.

In the case of an exponential utility, such as in (4.1), the dual problem of finding the minimal entropy martingale measure was treated by several authors assuming that a *reverse Hölder condition* is satisfied; see, e.g. Grandits and Rheinländer (2002), Delbaen *et al.* (2002), and Mania *et al.* (2003).

In the following, we introduce all the inequalities we need using notation consistent with exponential models.

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

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

$$\mathbb{E}_p\bigg(\frac{q/p}{q_\tau/p_\tau}\log\bigg(\frac{q/p}{q_\tau/p_\tau}\bigg) \,\bigg|\, \mathcal{F}_\tau\bigg) \leq C \quad \text{for all stopping times } \tau \leq T.$$

Definition 4.2. ($R_{1+\varepsilon}(p)$.) 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

$$\mathbb{E}_p\bigg(\bigg(\frac{q/p}{q_\tau/p_\tau}\bigg)^{1+\varepsilon} \,\bigg|\, \mathcal{F}_\tau\bigg) \leq C \quad \text{for all stopping times } \tau \leq T.$$

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

$$\mathbb{E}_p\bigg(\bigg(\frac{q/p}{q_\tau/p_\tau}\bigg)^{-\varepsilon} \,\bigg|\, \mathcal{F}_\tau\bigg) \leq C \quad \text{for all stopping times } \tau \leq T.$$

https://doi.org/10.1017/jpr.2018.45 Published online by Cambridge University Press

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

Proposition 4.1. *Let X be a continuous semimartingale and assume there exists $q \in \mathcal{M}_f^{\mathrm{e}}$ which satisfies $R_{L\log L}(p)$. Then $q^* \in \mathcal{E}(p)$.*

Proof. As observed above, q^* satisfies $R_{L\log L}(p)$. From Grandits and Rheinländer (2002, Lemmas 2.2 and 4.6), this implies that q^* also satisfies $R_{1+\varepsilon}(p)$ for some $\varepsilon > 0$. As a consequence, we find that $q^*/p \in L^{1+\varepsilon}(p)$ and the first integrability condition of Theorem 3.2(vi) is satisfied. The validity of the second integrability condition follows from Doléans-Dade and Meyer (1979, Proposition 5). 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 $p/q^* \in L^\varepsilon(p)$ for some $\varepsilon > 0$. □

Remark 4.1. Due to Proposition 3.5, if $q^* \in \mathcal{E}(p)$ then $q_t^* \in \mathcal{E}_{V_{\mathcal{F}_t}}(p_t)$ for all $t \le T$.

In the next subsection we provide some examples in which we 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 $\mathcal{M}_f^{\mathrm{e}} \neq \varnothing$ then q^* has the form

$$q^* = c^* \mathrm{e}^{-W_T(\theta^*)} p, \tag{4.2}$$

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

Theorem 4.1. *If $q^* \in \mathcal{E}(p)$ then*

(i) $\mathrm{e}^{-W_T(\theta^*)} \in L^{1+\varepsilon}(p)$ *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

$$q^* = \mathrm{e}^{u^* - K_p(u^*)} p, \tag{4.3}$$

where $u^* \in \mathring{\mathrm{dom}} K_p$ and $K_p(u^*) = D(p \,\|\, q^*)$. Comparing (4.3) with (4.2), we deduce that

$$W_T(\theta^*) = -u^* + D(q^* \,\|\, p) + D(p \,\|\, q^*). \tag{4.4}$$

From Corollary 3.1, 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 Theorem 3.2(iv), i.e. the equality $L^{\Phi_1}(p) = L^{\Phi_1}(q^*)$. In fact, since the process $W(\theta^*)$ is a q^*-martingale and taking into account (4.4), we obtain

$$W_t(\theta^*) = -\mathbb{E}_{q^*}(u^* \mid \mathcal{F}_t) + D(q^* \,\|\, p) + D(p \,\|\, q^*),$$

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

$$\mathbb{E}_{q^*}(\mathrm{e}^{\alpha \mathbb{E}_{q^*}(u^* \mid \mathcal{F}_t)}) \le \mathbb{E}_{q^*}(\mathbb{E}_{q^*}(\mathrm{e}^{\alpha u^*} \mid \mathcal{F}_t)) = \mathbb{E}_{q^*}(\mathrm{e}^{\alpha u^*}) < +\infty. \qquad \square$$

Remark 4.2. Theorem 4.1(i) 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 4.3. From (4.4), we deduce that

$$\mathbb{E}_p(W_T(\theta^*)) = \mathbb{E}_{q^*}(u^*) = D(q^* \parallel p) + D(p \parallel q^*)$$

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

We now return to deal with the duality problem (4.1), assuming the $R_{L\log L}(p)$ condition and $\mathcal{M}_f^{\mathrm{e}} \neq \varnothing$. 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

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

Specifically, they showed that the supremum in (4.1) is attained as a maximum when computed over the class Θ_2 (the same cannot be stated for Θ_3).

Furthermore, we recall that the infimum is attained and the minimal entropy martingale measure q^* is in $\mathcal{M}_f^{\mathrm{e}}$.

Assuming the price process X is continuous, from Proposition 4.1 we see that, in fact, q^* belongs to the maximal exponential model $\mathcal{E}(p)$. As a consequence, Theorem 4.1 proves that the maximum of the utility problem is reached on the smaller class $\widehat{\Theta}_2 \subseteq \Theta_2$, defined as

$$\widehat{\Theta}_2 = \{\theta \in L(X) \mid \mathrm{e}^{-W_T(\theta)} \in L^{1+\varepsilon}(p),\, W_t(\theta) \in L^{\Phi_1}(p) \text{ for all } t \in [0, T]$$
$$\text{and } W(\theta) \text{ is a } q\text{-martingale for all } q \in \mathcal{M}_f\}.$$

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

Corollary 4.1. *Let X be a continuous semimartingale and assume there exists $q \in \mathcal{M}_f^{\mathrm{e}}$ which satisfies $R_{L\log L}(p)$. Then*

$$\begin{aligned}\max_{\theta\in\widehat{\Theta}_2} \mathbb{E}_p(U(W_T(\theta))) &= \max_{\theta\in\Theta_2} \mathbb{E}_p(U(W_T(\theta)))\\ &= U\Big(\min_{q\in\mathcal{M}_f} D(q \parallel p)\Big)\\ &= U\Big(\min_{q\in\mathcal{M}\cap\mathcal{E}(p)} D(q \parallel p)\Big).\end{aligned} \tag{4.5}$$

In the following proposition we prove 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 in its own right.

Proposition 4.2. *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

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

https://doi.org/10.1017/jpr.2018.45 Published online by Cambridge University Press

As q is connected by an exponential arc to p, by Theorem 3.2(iv), $L^{\Phi_1}(p) = L^{\Phi_1}(q)$ and, therefore, $W_T(\theta) \in L^{\Phi_1}(q) \subseteq L^2(q)$, from which the result 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 dynamic properties of the $R_{L\log L}(p)$ condition, as in Delbaen *et al.* (2002).

4.1. Examples

In this subsection 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 4.1 and conclude that the optimal wealth process belongs to the Orlicz space $L^{\Phi_1}(\mu)$.

Example 4.1. (*Merton's model.*) Let $\mathcal{F}$ be the augmented filtration generated by a Brownian motion $(B_t)_{0\le t\le T<\infty}$. Assume that the price process X follows the Black–Scholes dynamics

$$\mathrm{d}X_t = X_t(\sigma\,\mathrm{d}B_t + m\,\mathrm{d}t) \quad \text{for all } 0 \le t \le T$$

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

$$q^* = \exp\biggl(-\frac{m}{\sigma}B_T - \frac{1}{2}\frac{m^2}{\sigma^2}T\biggr).$$

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

$$\mathbb{E}_{q^*}\biggl(-\frac{m}{\sigma}(B_T - B_\tau) - \frac{1}{2}\frac{m^2}{\sigma^2}(T-\tau) \biggm| \mathcal{F}_\tau\biggr) = \frac{1}{2}\frac{m^2}{\sigma^2}(T-\tau) \le \frac{1}{2}\frac{m^2}{\sigma^2}T,$$

which, by Proposition 4.1, implies that $q^* \in \mathcal{E}(1)$ with

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

Nevertheless, this can be directly proved using Theorem 3.2(vi) since

$$\mathbb{E}((q^*)^{1+\varepsilon}) = \mathbb{E}\biggl(\exp\biggl(-\frac{m}{\sigma}(1+\varepsilon)B_T - \frac{1}{2}\frac{m^2}{\sigma^2}(1+\varepsilon)T\biggr)\biggr) < +\infty,$$

$$\mathbb{E}((q^*)^{-\varepsilon}) = \mathbb{E}\biggl(\exp\biggl(\frac{m}{\sigma}\varepsilon B_T + \frac{1}{2}\frac{m^2}{\sigma^2}\varepsilon T\biggr)\biggr) < +\infty.$$

By (4.4), we can compute the final wealth

$$W_T(\theta^*) = -u^* + D(q^* \,\|\, 1) + D(1 \,\|\, q^*) = \frac{m}{\sigma^2}\Bigl(\sigma B_T + mT\Bigr) = \int_0^T \theta_s^*\,\mathrm{d}X_s,$$

where $\theta_s^* = m/\sigma^2$ for all $0 \le s \le T$.

Finally, the optimal expected utility in (4.5) turns out to be

$$\mathbb{E}(U(W_T(\theta^*))) = U(D(q^* \,\|\, 1)) = -\exp\biggl(-\frac{1}{2}\frac{m^2}{\sigma^2}T\biggr).$$

Example 4.2. (*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 (bounded mean oscillation) continuous martingale; see, e.g. Mania *et al.* (2003). This fact is achieved assuming a reverse Hölder condition which, in our framework using Proposition 4.1, guarantees that q^* belongs to an exponential model.

The link between BMO continuous martingales and reverse Hölder inequalities was established by Kazamaki (1994) and 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 $\mathbb{E}(\langle M\rangle_T - \langle M\rangle_\tau \mid F_\tau) \leq C$ for a fixed constant $C > 0$ and for any stopping time τ;

(ii) the stochastic exponential $\mathfrak{E}(M)$ is uniformly integrable and satisfies $R_{1+\varepsilon}$ for some $\varepsilon > 0$;

(iii) the stochastic exponential $\mathfrak{E}(M)$ 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

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

where M is a continuous local martingale and λ is a predictable process such that the *mean variance tradeoff* $\int_0^T \lambda_s^2 \,\mathrm{d}\langle M\rangle_s$ is finite.

Under the standing assumption $\mathcal{M}_f^{\mathrm{e}} \neq \varnothing$, we introduce the value process associated to the problem of finding the minimal entropy martingale measure

$$V_t = \operatorname*{ess\,inf}_{q\in\mathcal{M}_f^{\mathrm{e}}} \mathbb{E}_q\left(\ln \frac{q}{q_t} \;\middle|\; F_t\right).$$

For any $q \in \mathcal{M}_f^{\mathrm{e}}$, 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 $\lambda \cdot M$ denotes 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 $\hat{q} = \mathfrak{E}(-\lambda \cdot M)$ is a true martingale, $\hat{q}$ defines an equivalent probability measure called the minimal martingale measure for X.

Our next result follows from Mania *et al.* (2003, Lemma 3.1 and Theorem 3.1).

Theorem 4.2. *Assume the filtration $\mathcal{F}$ 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*

$$Y_t = Y_0 - \frac{1}{2}\int_0^t \lambda_s\Big(\lambda_s - 2\varphi_s\Big)\,\mathrm{d}\langle M\rangle_s + \frac{1}{2}\langle\tilde{L}\rangle_t + \int_0^t \varphi_s \,\mathrm{d}M_s + \tilde{L}_t, \qquad t < T, \tag{4.6}$$
$$Y_T = 0,$$

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

$$q^* = \mathfrak{E}_T(-\lambda \cdot M - \tilde{L}). \tag{4.7}$$

https://doi.org/10.1017/jpr.2018.45 Published online by Cambridge University Press

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 (4.7) with (4.3), we can also express u^* in terms of the BMO martingale $\tilde{L}$ appearing in the solution of the backward stochastic differential equation (BSDE) (4.6). In fact,

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

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

$$K_1(u^*) = D(1 \,\|\, q^*) = \tfrac{1}{2}\mathbb{E}(\langle \lambda \cdot M + \tilde{L} \rangle_T) \tag{4.8}$$

and

$$u^* = (\lambda \cdot M)_T - \tilde{L}_T - \tfrac{1}{2}(\langle \lambda \cdot M + \tilde{L} \rangle_T) + \tfrac{1}{2}\mathbb{E}(\langle \lambda \cdot M + \tilde{L} \rangle_T). \tag{4.9}$$

Since V is the solution of BSDE (4.6), we can write its initial value as

$$V_0 = \tfrac{1}{2}\langle \lambda \cdot M \rangle_T - \langle \varphi \cdot M, \lambda \cdot M \rangle_T - \tfrac{1}{2}\langle \tilde{L} \rangle_T - (\varphi \cdot M)_T - \tilde{L}_T. \tag{4.10}$$

Furthermore, recall that, by definition, $V_0 = \mathbb{E}(q^* \ln q^*) = D(q^* \,\|\, 1)$.

By (4.4), and exploiting (4.8)–(4.10), we can compute the final wealth

$$\begin{aligned} W_T(\theta^*) &= -u^* + D(q^* \,\|\, 1) + D(1 \,\|\, q^*) \\ &= ((\lambda - \varphi) \cdot M)_T + \langle \lambda \cdot M \rangle_T - \langle \varphi \cdot M, \lambda \cdot M \rangle_T \\ &= \int_0^T \theta_s^* \, \mathrm{d}X_s, \end{aligned}$$

where $\theta_s^* = \lambda - \varphi$ for all $0 \le s \le T$.

Finally, the optimal expected utility in (4.5) turns out to be

$$\mathbb{E}(U(W_T(\theta^*))) = U(D(q^* \,\|\, 1)) = -\mathrm{e}^{-V_0}.$$

Example 4.3. 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 BSDE (4.6) 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

$$\mathrm{d}X_t = X_t(\sigma(t, X_t)\,\mathrm{d}B_t + m(t, X_t)\,\mathrm{d}t) \quad \text{for all } 0 \le t \le T, \tag{4.11}$$

where B is a standard Brownian motion, m and σ are bounded measurable functions such that SDE (4.11) admits a unique strong solution, and $\sigma^2 \ge c > 0$. 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 4.2 and the Markovianity of the coefficients imply that $V_t = V(t, X_t)$; therefore, the optimal solution is characterized by a partial differential equation and the minimal entropy martingale measure q^* coincides with the minimal martingale measure $\widehat{q}$. We immediately obtain

$$\begin{aligned} D(q^* \,\|\, 1) = V_0 &= \frac{1}{2}\mathbb{E}_{\widehat{q}}\bigg(\int_0^T \frac{m^2(t, X_t)}{\sigma^2(t, X_t)}\,\mathrm{d}t\bigg), \\ D(1 \,\|\, q^*) = K_1(u^*) &= \frac{1}{2}\mathbb{E}\bigg(\int_0^T \frac{m^2(t, X_t)}{\sigma^2(t, X_t)}\,\mathrm{d}t\bigg), \end{aligned}$$

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.

https://doi.org/10.1017/jpr.2018.45 Published online by Cambridge University Press

Furthermore,

$$u^* = -\int_0^T \frac{m(t, X_t)}{\sigma(t, X_t)}\,\mathrm{d}B_t - \frac{1}{2}\int_0^T \frac{m^2(t, X_t)}{\sigma^2(t, X_t)}\,\mathrm{d}t + \frac{1}{2}\mathbb{E}\bigg(\int_0^T \frac{m^2(t, X_t)}{\sigma^2(t, X_t)}\,\mathrm{d}t\bigg)$$

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

$$\begin{aligned} \mathrm{d}X_t &= X_t(\sigma(t, Y_t)\,\mathrm{d}B_t + m(t, Y_t)\,\mathrm{d}t), \\ \mathrm{d}Y_t &= \sigma^{\perp}(t, Y_t)\,\mathrm{d}B_t^{\perp} + b(t, Y_t)\,\mathrm{d}t \quad \text{for all } 0 \le t \le T, \end{aligned} \tag{4.12}$$

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 (4.12) admit a unique strong solution. In this case, the market price of the risk $m(t, Y_t)/\sigma(t, Y_t)$ is $\mathcal{F}^{B^{\perp}}$-adapted and, as a consequence, $V_t = V(t, Y_t)$. The optimal strategy is essentially the ratio $\theta^* = m(t, Y_t)/\sigma^2(t, Y_t)$ which is a function of the exogenous Y_t, thus, generalizing the classical Merton strategy.

Example 4.4. Many examples where the optimal solution q^* belongs to the maximal exponential model can be found in the literature, even for nonlocally bounded X. We analyze the one from Biagini and Frittelli (2008). Let N be a Poisson process of parameter $\lambda > 0$ with jump times $\mathsf{T} = (T_j)_{j\ge1}$, $T_0 = 0$. Denote $Y_0 = 0$, and $(Y_j)_{j\ge1}$ a sequence of independent and identically distributed random variables independent from T with density $f(y) = \frac{1}{2}\nu\mathrm{e}^{-\nu|y-1|}$, $\nu > 0$. We define the price process

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

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

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

where $a^* = \sqrt{1+\nu^2} - 1$.

Using Theorem 3.2(vi), it can be checked that $q^* \in \mathcal{E}(1)$. In fact, $q^* \in L^{1+\varepsilon}(\mu)$ and $1/q^* \in L^{\varepsilon}(\mu)$ for some $\varepsilon > 0$ if and only if $M_{X_T}(-a^*(1+\varepsilon)) < \infty$ and $M_{X_T}(a^*\varepsilon) < \infty$, where M_{X_T} is the moment generating function of X_T, i.e.

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

Both conditions are verified for any $0 < \varepsilon < \nu/a^* - 1$.

Example 4.5. We provide 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 is adapted to the notation of exponential models.

Let $\mathcal{F}$ be the augmented filtration generated by a Brownian motion $(B_t)_{0\le t<\infty}$. Consider the price process X, i.e.

$$X_t = B_t^\tau - t \wedge \tau \quad \text{for all } 0 \le t < \infty,$$

where $\tau = \inf\{t \ge 0\colon B_t = 1\}$ is a stopping time such that $\mathbb{P}(\tau < \infty) = 1$ and $\mathbb{E}(\tau) = \infty$.

We can check that $\mathcal{M}_f^{\mathrm{e}} = \{q^*\}$, where

$$q^* = \exp\big(B_\tau - \tfrac{1}{2}\tau\big) = \exp\big(1 - \tfrac{1}{2}\tau\big).$$

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

$$\mathbb{E}((q^*)^{-\varepsilon}) = \mathbb{E}\big(\exp\big(-\varepsilon + \tfrac{1}{2}\varepsilon\tau\big)\big) \ge \exp\big(-\varepsilon + \tfrac{1}{2}\varepsilon\mathbb{E}(\tau)\big) = \infty.$$

From Theorem 3.2(vi), we conclude that $q^* \notin \mathcal{E}(1)$. Note that

$$D(q^* \,\|\, 1) = \mathbb{E}_{q^*}\big(1 - \tfrac{1}{2}\tau\big).$$

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

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

which, by Corollary 3.2, leads to the same conclusion $q^* \notin \mathcal{E}(1)$.

As observed by Delbaen *et al.* (2002), in this example the duality result (4.5) holds for Θ_2 even though $R_{L\log L}(p)$ is not satisfied.

4.2. An example under model uncertainty

In this subsection we present a simple example in which we show the possible impact of maximal exponential models on financial applications in an uncertainty framework. We refer to Merton's model described in Example 4.1.

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 Theorem 3.2(iii), choose 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.

We 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

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

which, from (4.5), is expressed through the dual problem by

$$\begin{aligned}\min_{\xi\in I} \max_{\theta\in\Theta_2(\xi)} \mathbb{E}_{p(\xi)}(U(W_T(\theta))) &= U\Big(\min_{\xi\in I} \min_{q\in\mathcal{M}_f(\xi)} D(q \,\|\, p(\xi))\Big) \\ &= U\Big(\min_{\xi\in I} D(q^* \,\|\, p(\xi))\Big). \end{aligned} \tag{4.13}$$

https://doi.org/10.1017/jpr.2018.45 Published online by Cambridge University Press

After some computation, we obtain

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

where u is the random variable characterizing the representation of r in the exponential model $\mathcal{E}(1)$, i.e. $r = \mathrm{e}^{u - K_1(u)}$. Recall also the representation of q^* in $\mathcal{E}(1)$, which is $q^* = \mathrm{e}^{u^* - K_1(u^*)}$, where $u^* = -(m/\sigma)B_T$.

A particular choice of r is made by selecting u such that (u, u^*) is a nondegenerate 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

$$D(q^* \parallel p(\xi)) = D(q^* \parallel 1) - \xi c + \tfrac{1}{2}\xi^2 \gamma^2$$

and is minimized by $\bar{\xi} = c/\gamma^2$.

Therefore, the solution of the dual problem in (4.13) is

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

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

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

In the primal problem, the optimal wealth can be explicitly identified adapting (4.4) for $p = p(\bar{\xi})$, i.e.

$$\begin{aligned} W_T(\theta^*(\bar{\xi})) &= -u^*(\bar{\xi}) + D(q^* \parallel p(\bar{\xi})) + D(p(\bar{\xi}) \parallel q^*) \qquad (4.14) \\ &= -u^* + \bar{u} - \bar{\xi} c + \frac{m^2}{\sigma^2} T \\ &= W_T(\theta^*) - \frac{m}{\sigma}\frac{\operatorname{cov}(u, B_T)}{\operatorname{var}(u)}\left(u + \frac{m}{\sigma}\operatorname{cov}(u, B_T)\right). \qquad (4.15) \end{aligned}$$

In some cases, the optimal strategy can be explicitly computed from (4.15). If, for instance, $u = \int_0^T \varphi(s)\,\mathrm{d}B_s$ with $\varphi \in L^2([0, T])$, then

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

Acknowledgements

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

References

ABRAHAM, R., MARSDEN, J. E. AND RATIU, T. (1988). *Manifold, Tensor Analysis and Applications*, 2nd edn. Springer, New York.

BIAGINI, S. AND FRITTELLI, M. (2008). A unified framework for utility maximization problems: an Orlicz space approach. *Ann. Appl. Prob.* **18,** 929–966.

BRIGO, D. AND PISTONE, G. (2017). Dimensionality Reduction for Measure Valued Evolution Equations in Statistical Manifolds. In *Computational Information Geometry. Signals and Communication Technology,* Springer, Cham, pp. 217–265.

CENA, A. AND PISTONE, G. (2007). Exponential statistical manifold. *Ann. Inst. Statist. Math.* **59,** 27–56.

CHERIDITO, P. AND LI, T. (2009). Risk measures on Orlicz hearts. *Math. Finance* **19,** 189–214.

DELBAEN, F. *et al.* (2002). Exponential hedging and entropic penalties. *Math. Finance* **12,** 99–123.

DOLÉANS-DADE, C. AND MEYER, P.-A. (1979). Inégalités de normes avec poids. In *Séminaire de Probabilités XIII* (Lecture Notes Math. **721**), Springer, Berlin, pp. 313–331.

FRITTELLI, M. (2000). The minimal entropy martingale measure and the valuation problem in incomplete markets. *Math. Finance* **10,** 39–52.

GRANDITS, P. AND RHEINLÄNDER, T. (2002). On the minimal entropy martingale measure. *Ann. Prob.* **30,** 1003–1038.

IMPARATO, D. AND TRIVELLATO, B. (2009). Geometry of extendend exponential models. In *Algebraic and Geometric Methods in Statistics*, Cambridge University Press, pp. 307–326.

KAZAMAKI, N. (1994). *Continuous Exponential Martingales and BMO* (Lecture Notes Math. **1579**). Springer, Berlin.

LODS, B. AND PISTONE, G. (2015). Information geometry formalism for the spatially homogeneous Boltzmann equation. *Entropy* **17,** 4323–4363.

MANIA, M., SANTACROCE, M. AND TEVZADZE, R. (2003). A semimartingale BSDE related to the minimal entropy martingale measure. *Finance Stoch.* **7,** 385–402.

MANIA, M., SANTACROCE, M. AND TEVZADZE, R. (2004). The Bellman equation related to the minimal entropy martingale measure. *Georgian Math. J.* **11,** 125–135.

PISTONE, G. (2013). Examples of the application of nonparametric information geometry to statistical physics. *Entropy* **15,** 4042–4065.

PISTONE, G. AND ROGANTIN, M. P. (1999). The exponential statistical manifold: mean parameters, orthogonality and space transformations. *Bernoulli* **5,** 721–760.

PISTONE, G. AND SEMPI, C. (1995). An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one. *Ann. Statist.* **23,** 1543–1561.

RAO, M. M. AND REN, Z. D. (1991). *Theory of Orlicz Spaces*. Marcel Dekker, New York.

RAO, M. M. AND REN, Z. D. (2002). *Applications of Orlicz Spaces*. Marcel Dekker, New York.

SANTACROCE, M., SIRI, P. AND TRIVELLATO, B. (2016). New results on mixture and exponential models by Orlicz spaces. *Bernoulli* **22,** 1431–1447.

SANTACROCE, M., SIRI, P. AND TRIVELLATO, B. (2017). *On Mixture and Exponential Connection by Open Arcs.* In *Geometric Science of Information, GSI 2017* (Lecture Notes Comp. Sci. **10589**), Springer, Cham. pp. 577–584.

SANTACROCE, M., SIRI, P. AND TRIVELLATO, B. (2018). An Application of Maximal Exponential Models to Duality Theory. *Entropy* **20,** 495.

https://doi.org/10.1017/jpr.2018.45 Published online by Cambridge University Press