An application of maximal exponential models to duality theory

(Article begins on next page)

An Application of Maximal Exponential Models to Duality Theory

Marina Santacroce (ID), **Paola Siri** (ID) **and Barbara Trivellato *** (ID)

Dipartimento di Scienze Matematiche "G.L. Lagrange", Dipartimento di Eccellenza 2018-2022, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy; marina.santacroce@polito.it (M.S.); paola.siri@polito.it (P.S.)
*   Correspondence: barbara.trivellato@polito.it; Tel.: +39-011-090-7550

Abstract: We use maximal exponential models to characterize a suitable polar cone in a mathematical convex optimization framework. A financial application of this result is provided, leading to a duality minimax theorem related to portfolio exponential utility maximization.

1. Introduction

In this manuscript, we use the notion of maximal exponential model in a convex duality framework and consider an application to portfolio optimization problems.

The theory of non-parametric maximal exponential models centered at a given positive density p starts with the work by Pistone and Sempi [1]. In that paper, by using the Orlicz space associated with an exponentially growing Young function, the set of positive densities is endowed with a structure of exponential Banach manifold.

More recently, different authors have generalized this structure replacing the exponential function with deformed exponentials (see, e.g., Vigelis and Cavalcante [2], De Andrade et al. [3]) and in other directions (see e.g., Imparato and Trivellato [4]).

The geometry of nonparametric exponential models and its analytical properties in the topology of Orlicz spaces have been also studied in subsequent works, such as Cena and Pistone [5] and Santacroce, Siri and Trivellato [6,7], among others.

Statistical exponential models built on Orlicz spaces have been exploited in several fields, such as differential geometry, algebraic statistics, information theory and physics (see, for example, Brigo and Pistone [8], Lods and Pistone [9]).

In mathematical finance, convex duality is strongly used to tackle portfolio optimization problems. Very recently, exponential models have been exploited for the first time to address the classical exponential utility maximization problem (see Santacroce, Siri and Trivellato [7]).

In this paper, we consider a particular set of densities contained in the maximal exponential model centered at a fixed density p, and we give a characterization of its polar cone with respect to a suitable dual system, studied in Santacroce, Siri and Trivellato [10].

This characterization is then exploited to slightly improve a well known financial result on exponential utility maximization. In particular, we prove that the classical optimal strategy, which solves the primal problem, maximizes the expected utility over a larger set.

The paper is organized as follows. In Section 2, we recall Orlicz spaces and related properties from the literature and some results on maximal exponential models, contained in our previous works and in papers by Pistone and different coauthors. Section 3 contains the main results of the paper, namely the characterization of the polar cone associated with the set of densities belonging to the maximal

exponential model and under which a given convex set of random variables has negative expectation. In Section 4, a financial application of this characterization is given and leads to a minimax theorem for exponential utility maximization.

2. Orlicz Spaces and Maximal Exponential Models

In this section, we recall some results from the theory of Orlicz spaces and exponential models that we will use hereinafter.

We denote with Φ a *Young function*, which is an even, convex function from $\mathbb{R}$ 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.

Two Young functions Φ and Φ' are said to be *equivalent* if there exists $x_0 > 0$, and $0 < c_1 < c_2$ such that $\Phi(c_1 x) \leq \Phi'(x) \leq \Phi(c_2 x)$, for any $x \geq x_0$.

In the following, we will focalize on $\Phi_1(x) = \cosh(x) - 1$, which is equivalent to the more commonly used $\Phi_2(x) = e^{|x|} - |x| - 1$.

Let us recall that the *conjugate convex function F^** of a real function F is defined as

$$F^*(y) = \sup_{x \in \mathbb{R}} \{xy - F(x)\}, \ \forall y \in \mathbb{R}, \tag{1}$$

and that the Fenchel inequality

$$xy \leq F(x) + F^*(y), \ \forall x, y \in \mathbb{R} \tag{2}$$

holds. In the case of a Young function Φ the conjugate, denoted by Ψ, is itself a Young function and the stronger Fenchel–Young inequality

$$|xy| \leq \Phi(x) + \Psi(y), \ \forall x, y \in \mathbb{R} \tag{3}$$

holds.

The conjugate of the Young function Φ_1 is $\Psi_1(y) = \int_0^y \sinh^{-1}(t)dt$ and is equivalent to the conjugate of Φ_2, namely $\Psi_2(y) = (1 + |y|) \log(1 + |y|) - |y|$.

From now on, we will work on a fixed probability space $(\mathcal{X}, \mathcal{F}, \mu)$. We denote with $\mathcal{P}$ the set of all densities that are positive μ-a.s. and with $\mathbb{E}_p$ the expectation with respect to $pd\mu$, for each fixed $p \in \mathcal{P}$.

Given $p \in \mathcal{P}$, we consider the Orlicz space associated with a Young function Φ, defined by

$$L^\Phi(p) = \{u : \mathcal{X} \to \mathbb{R} \text{ measurable} : \exists\, \alpha > 0 \text{ s.t. } \mathbb{E}_p(\Phi(\alpha u)) < +\infty\}. \tag{4}$$

$L^\Phi(p)$ is a Banach space when endowed with the *Luxembourg norm*

$$\|u\|_{\Phi,p} = \inf \left\{ k > 0 \ : \ \mathbb{E}_p \left(\Phi \left(\frac{u}{k} \right) \right) \leq 1 \right\}. \tag{5}$$

The Orlicz spaces $L^\Phi(p)$ and $L^{\Phi'}(p)$ associated with two equivalent Young functions are equal as sets and have equivalent norms as Banach spaces.

Finally, it is worth noting the following chain of inclusions:

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

Definition 1. *$p, q \in \mathcal{P}$ are connected by an open exponential arc if there exists an open interval $I \supset [0, 1]$ such that the following equivalent relations are satisfied:*

1.　*$p(\theta) \propto p^{(1-\theta)} q^\theta \in \mathcal{P}, \ \forall \theta \in I$;*
2.　*$p(\theta) \propto e^{\theta u} p \in \mathcal{P}, \ \forall \theta \in I$, where $u \in L^{\Phi_1}(p)$ and $p(0) = p, \ p(1) = q$.*

Observe that connection by open exponential arcs is an equivalence relation.

Definition 2. *Let us denote $L_0^{\Phi_1}(p) = \{u \in L^{\Phi_1}(p) : \mathbb{E}_p(u) = 0\}$. The cumulant generating functional is the map $K_p : L_0^{\Phi_1}(p) \longrightarrow [0, +\infty]$ defined by the relation $K_p(u) = \log \mathbb{E}_p(e^u)$.*

We recall from Pistone and Sempi [1] that K_p is a positive convex and lower semicontinuous function, vanishing at zero. Moreover, the interior of its proper domain

$$\text{dom}\, K_p = \{u \in L_0^{\Phi_1}(p) : K_p(u) < +\infty\}, \tag{6}$$

denoted here by $\overset{\circ}{\text{dom}}\, K_p$, is a non-empty convex set.

Definition 3. *For every density $p \in \mathcal{P}$, the maximal exponential model at p is defined as*

$$\mathcal{E}(p) = \left\{ q = e^{u - K_p(u)} p : u \in \overset{\circ}{dom}\, K_p \right\} \subseteq \mathcal{P}. \tag{7}$$

In the sequel, we use the notation $D(q\|p)$ to indicate the Kullback–Leibler divergence of $\mathbb{Q} = q \cdot \mu$ with respect to $\mathbb{P} = p \cdot \mu$ and we simply refer to it as the divergence of q from p.

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

Proposition 2. *Let $p, q \in \mathcal{P}$. If $D(q\|p) < +\infty$, then $L^{\Phi_1}(p) \subseteq L^1(q)$.*

The proof of Proposition 1 can be found in Cena and Pistone [5], while Proposition 2 is proved in Santacroce, Siri and Trivellato [6].

We now state one of the central results of [5–7], which gives equivalent conditions to open exponential connection by arcs, in a complete version, containing all the recent improvements.

Theorem 1. *(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) $\log \frac{q}{p} \in L^{\Phi_1}(p) \cap L^{\Phi_1}(q)$;
(v) $L^{\Phi_1}(p) = L^{\Phi_1}(q)$;
(vi) $\frac{q}{p} \in L^{1+\varepsilon}(p)$ *and* $\frac{p}{q} \in L^{1+\varepsilon}(q)$, *for some $\varepsilon > 0$*;
(vii)   *the mixture transport mapping*

$$^m\mathbb{U}_p^q : L^{\Psi_1}(p) \longrightarrow L^{\Psi_1}(q) \tag{8}$$
$$v \mapsto \frac{p}{q}v,$$

is an isomorphism of Banach spaces.

The equivalence of conditions $(i) \div (iv)$ is proved in Cena and Pistone [5]. Statements (v) and (vi) have been added by Santacroce, Siri and Trivellato [6], while statement (vii) by Santacroce, Siri and Trivellato [7]. It is worth noting that, among all conditions of the Portmanteau Theorem, (v) and (vi) are the most useful from a practical point of view: the first one allows for switching from one Orlicz space to the other at one's convenience, while the second one permits working with Lebesgue spaces. On the other hand condition (vii), involving the mixture transport mapping could be a useful

tool in physics applications of exponential models, as the recent research on the subject demonstrates (see, e.g., Pistone [11], Lods and Pistone [9], Brigo and Pistone [8]). In these applications, finiteness of Kullback–Leibler divergence, implied from Portmanteau Theorem, is a desirable property.

Corollary 1. *If $q \in \mathcal{E}(p)$, then the Kullback–Leibler divergences $D(q\|p)$ and $D(p\|q)$ are both finite.*

The converse of this corollary does not hold, as the counterexamples in Santacroce, Siri and Trivellato [6,7] show.

Proposition 3. *Let $p \in \mathcal{P}$. Then, $\mathcal{E}(p)$ is a convex set.*

For the proof, see Santacroce, Siri and Trivellato [6].

3. Duality Results

This section contains the main results of the paper, namely it shows an application of the maximal model to duality in a general framework.

We start with introducing as basic tools $\mathcal{K} \subseteq L^0(p)$ a convex cone containing 0 and

$$\mathcal{M}(\mathcal{K}) = \{q \in \mathcal{P}_{\geq} : -\infty < \mathbb{E}_q(k) \leq 0, \ \forall k \in \mathcal{K}\}, \tag{9}$$

where $\mathcal{P}_{\geq}$ denotes the set of non negative densities. Let us define

$$\mathcal{M}_{\mathcal{E}}(\mathcal{K}) = \mathcal{M}(\mathcal{K}) \cap \mathcal{E}(p) \tag{10}$$

and suppose $\mathcal{M}_{\mathcal{E}}(\mathcal{K}) \neq \emptyset$.

Furthermore, let us introduce the linear spaces

$$L = \bigcap_{q \in \mathcal{M}_{\mathcal{E}}(\mathcal{K})} L^1(q) \supseteq \mathcal{K} \ \text{and} \ L' = Lin\{\mathcal{M}_{\mathcal{E}}(\mathcal{K})\} \subseteq L^1(p). \tag{11}$$

L and L' are dual spaces with the duality given by the bilinear map $(l, l') \to \langle l, l' \rangle = \mathbb{E}_\mu(ll')$.

The dual system is separated in both L and L' and, endowed with the weak topologies, they become locally convex Hausdorff topological vector spaces (see Santacroce, Siri and Trivellato [10] for details on this duality).

In the following, we denote with $\mathcal{M}_{\mathcal{E}}^0(\mathcal{K})$ the polar cone of $\mathcal{M}_{\mathcal{E}}(\mathcal{K})$, i.e.,

$$\mathcal{M}_{\mathcal{E}}^0(\mathcal{K}) = \{h \in L \,|\, \mathbb{E}_q(h) \leq 0 \ \forall q \in \mathcal{M}_{\mathcal{E}}(\mathcal{K})\}. \tag{12}$$

The following result gives a characterization of the density that minimizes the divergence over $\mathcal{M}_{\mathcal{E}}(\mathcal{K})$ in a formulation that is very close to the one given by Biagini and Frittelli [12]. Therefore, we omit the proof, even if the fact that we work on the exponential model makes some checking more straighforward.

Proposition 4. *$\bar{q} \in \mathcal{M}_{\mathcal{E}}(\mathcal{K})$ minimizes $D(q \| p)$ over $\mathcal{M}_{\mathcal{E}}(\mathcal{K})$ if and only if $D(\bar{q} \| p) - \log \frac{\bar{q}}{p} \in \mathcal{M}_{\mathcal{E}}^0(\mathcal{K})$.*

In the next theorem, we give the main result of the paper, namely a useful description of the polar cone in terms of $\mathcal{K}$.

Theorem 2.

$$\mathcal{M}_{\mathcal{E}}^0(\mathcal{K}) = \bigcap_{q \in \mathcal{M}_{\mathcal{E}}(\mathcal{K})} \overline{\mathcal{K} - L_+^1(q)}^q, \tag{13}$$

where $\overline{C}^q$ represents the $L^1(q)$-closure of a set C.

Proof. The technique used in the proof is in the spirit of that in Biagini and Frittelli [12]. On the other hand, here we need to work with measures belonging to $\mathcal{E}(p)$. This adds a difficulty to the proof which can be solved by exploiting the properties of maximal exponential models given by the Portmanteau Theorem (i.e., Theorem 1).

We prove only that

$$\mathcal{M}_{\mathcal{E}}^0(\mathcal{K}) \subseteq \underset{q \in \mathcal{M}_{\mathcal{E}}(\mathcal{K})}{\cap} \overline{\mathcal{K} - L_+^1(q)}^q,$$

since the opposite inclusion essentially follows from the definition of $\mathcal{M}_{\mathcal{E}}(\mathcal{K})$.

Let $h \in \mathcal{M}_{\mathcal{E}}^0(\mathcal{K})$ and suppose by contradiction that there exists $q_0 \in \mathcal{M}_{\mathcal{E}}(\mathcal{K})$ such that $h \notin \overline{\mathcal{K} - L_+^1(q_0)}^{q_0}$.

By the Hahn–Banach Theorem, there exists $\eta \in L^\infty(q_0)$ such that for every $k \in \mathcal{K} - L_+^1(q_0)$

$$\mathbb{E}_{q_0}[\eta k] \leq 0 < \mathbb{E}_{q_0}[\eta h]. \tag{14}$$

If we consider $k = -\mathbb{1}_{(\eta<0)}$, by (14), we get $\mathbb{E}_{q_0}[\eta \mathbb{1}_{(\eta<0)}] = 0$ and thus $\eta \geq 0$ μ-a.s..

Then, we can define a new density $q_1 = \frac{\eta}{\mathbb{E}_{q_0}[\eta]} q_0$ (μ-a.s.). Let us note that $q_1 \geq 0$ does not belong to $\mathcal{P}$ but, by (14), $q_1 \in \mathcal{M}(\mathcal{K})$ and, $\mathbb{E}_{q_1}[h] > 0$.

Now consider the convex combination $q_\lambda = \lambda q_1 + (1 - \lambda)q_0$, with $0 \leq \lambda < 1$.

We can prove that $q_\lambda \in \mathcal{M}_{\mathcal{E}}(\mathcal{K})$, for any $0 \leq \lambda < 1$. It is clear that $q_\lambda \in \mathcal{M}(\mathcal{K}) \cap \mathcal{P}$. In order to check that it belongs to the maximal exponential model, we use statement $vi)$ of Theorem 1. In fact,

$$\mathbb{E}_p\left[\left(\frac{q_\lambda}{p}\right)^{1+\epsilon}\right] = \mathbb{E}_p\left[\left(\lambda \frac{\eta}{\mathbb{E}_{q_0}[\eta]}\frac{q_0}{p} + (1 - \lambda)\frac{q_0}{p}\right)^{1+\epsilon}\right] \leq C\,\mathbb{E}_p\left[\left(\frac{q_0}{p}\right)^{1+\epsilon}\right], \tag{15}$$

$$\mathbb{E}_p\left[\left(\frac{p}{q_\lambda}\right)^{\epsilon}\right] = \mathbb{E}_p\left[\frac{1}{\left(\lambda\frac{\eta}{\mathbb{E}_{q_0}[\eta]} + 1 - \lambda\right)^{\epsilon}}\left(\frac{p}{q_0}\right)^{\epsilon}\right] \leq C\,\mathbb{E}_p\left[\left(\frac{p}{q_0}\right)^{\epsilon}\right], \tag{16}$$

where the inequalities are due to the fact that η is non negative and bounded.

Since $q_0 \in \mathcal{E}(p)$, by $vi)$ of Theorem 1, we get $\mathbb{E}_p\left[\left(\frac{q_0}{p}\right)^{1+\epsilon}\right] < +\infty$ and $\mathbb{E}_p\left[\left(\frac{p}{q_0}\right)^{\epsilon}\right] < +\infty$. Moreover, from (15) and (16), we deduce $\mathbb{E}_p\left[\left(\frac{q_\lambda}{p}\right)^{1+\epsilon}\right] < +\infty$ and $\mathbb{E}_p\left[\left(\frac{p}{q_\lambda}\right)^{\epsilon}\right] < +\infty$. Therefore, $q_\lambda \in \mathcal{E}(p)$ and consequently $q_\lambda \in \mathcal{M}_{\mathcal{E}}(\mathcal{K})$.

Finally, let us denote by $\beta = \mathbb{E}_{q_1}[h] > 0$ and $\alpha = \mathbb{E}_{q_0}[h] \leq 0$, since $q_0 \in \mathcal{M}_{\mathcal{E}}(\mathcal{K})$ and $h \in \mathcal{M}_{\mathcal{E}}^0(\mathcal{K})$. It is easy to see that $\mathbb{E}_{q_\lambda}[h] = \lambda\beta + (1 - \lambda)\alpha > 0$ if and only if $\lambda > \frac{-\alpha}{\beta - \alpha}$.

By definition of the polar cone, we get to a contradiction and this concludes the proof. □

4. Financial Application

In this 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 continuous 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, which 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 that are strictly positive μ-a.s. and define

$$\mathcal{M}_f = \{q \in \mathcal{M} : D(q\|p) < \infty\}, \qquad \mathcal{M}_f^e = \mathcal{M}_f \cap \mathcal{M}^e. \tag{17}$$

We assume throughout the paper that $\mathcal{M}_f^e \neq \emptyset$.

Note that, by Corollary 1,

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

A *self-financing trading strategy* is denoted by $\theta = (\theta_t)_{0 \le t \le T}$, 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

$$\sup_{\theta \in \Theta} \mathbb{E}_p \left[U(W_T(\theta)) \right] \tag{18}$$

over a set Θ of *admissible strategies*.

In this framework, Θ is the classical set of strategies where wealth is uniformly bounded from below (see for example Schachermayer [13]):

$$\Theta = \{ \theta \in L(X) : W_t(\theta) \ge c \ \mu\text{-a.s.}, \ \forall \ 0 \le t \le T, \text{ for some } c \in \mathbb{R} \}. \tag{19}$$

It is well known that the maximization problem in (18) can be written in the duality formulation

$$\sup_{\theta \in \Theta} \mathbb{E}_p \left[U(W_T(\theta)) \right] = U(\inf_{q \in \mathcal{M}_f} D(q \| p)). \tag{20}$$

Recall 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 Theorem 2.1 in Frittelli [19]). This q^* is called the *minimal entropy martingale (density) measure*. If, in addition, $\mathcal{M}_f^e \neq \varnothing$, then $q^* > 0 \ \mu$-a.s..

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

$$\mathbb{E}_p \left[\frac{q/p}{q_\tau / p_\tau} \log \left(\frac{q/p}{q_\tau / p_\tau} \right) \Big| \mathcal{F}_\tau \right] \le C \text{ for all stopping times } \tau \le T. \tag{21}$$

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. [14]). When the process X is continuous, this fact then implies that $q^* \in \mathcal{E}(p)$, as the following proposition shows.

Proposition 5. *Let X be a continuous semimartingale and assume there exists $q \in \mathcal{M}_f^e$ that satisfies $R_{L \log L}(p)$. Then, $q^* \in \mathcal{E}(p)$.*

(For the proof, see Santacroce, Siri and Trivellato [7]).

Let us consider here the convex cone whose elements are the final values of the portfolio wealth

$$\mathcal{K} = \{ W_T(\theta) : \theta \in \Theta \}. \tag{22}$$

The maximization problem can then be rewritten in terms of $\mathcal{K}$:

$$\sup_{k \in \mathcal{K}} \mathbb{E}_p \left[U(k) \right]. \tag{23}$$

Moreover, $\mathcal{M}(\mathcal{K})$ in (9) coincides with the set of martingale measures $\mathcal{M}$. For the sake of coherence, we drop the dependence on $\mathcal{K}$ also in $\mathcal{M}_{\mathcal{E}}(\mathcal{K})$ in (10) and $\mathcal{M}_{\mathcal{E}}^0(\mathcal{K})$ in (12), denoting them respectively $\mathcal{M}_{\mathcal{E}}$ and $\mathcal{M}_{\mathcal{E}}^0$.

It is common knowledge that the optimal solution to the primal utility maximization problem (23) on the class $\mathcal{K}$ does not exist, but it exists on the larger set $\mathcal{K}_f = \underset{q \in \mathcal{M}_f}{\cap} \overline{\mathcal{K} - L_+^1(q)}^q$ (see, e.g., Biagini and Frittelli [12]).

The following theorem, exploiting the results in the previous section, proves that this set can be further enlarged if the minimal martingale measure q^* belongs to $\mathcal{E}(p)$.

Theorem 3. *Let X be a continuous semimartingale and assume there exists $q \in \mathcal{M}_f^e$ which satisfies $R_{L \log L}(p)$. Then,*

$$\sup_{k \in \mathcal{K}} \mathbb{E}_p\left(U(k)\right) = \max_{k \in \mathcal{K}_{\mathcal{E}}} \mathbb{E}_p\left(U(k)\right) = U\left(\min_{q \in \mathcal{M}_{\mathcal{E}}} D(q \parallel p)\right) = U\left(\min_{q \in \mathcal{M}_f} D(q \parallel p)\right), \tag{24}$$

where

$$\mathcal{K}_{\mathcal{E}} = \underset{q \in \mathcal{M}_{\mathcal{E}}}{\cap} \overline{\mathcal{K} - L_+^1(q)}^q. \tag{25}$$

Proof. It is trivial to note that $\mathcal{K}_f \subseteq \mathcal{K}_{\mathcal{E}}$ so that $\underset{k \in \mathcal{K}_f}{\sup} \mathbb{E}_p\left(U(k)\right) \le \underset{k \in \mathcal{K}_{\mathcal{E}}}{\sup} \mathbb{E}_p\left(U(k)\right)$.

Using the duality results from Biagini and Frittelli [12], we get

$$\sup_{k \in \mathcal{K}} \mathbb{E}_p\left(U(k)\right) = \sup_{k \in \mathcal{K}_f} \mathbb{E}_p\left(U(k)\right) = \max_{k \in \mathcal{K}_f} \mathbb{E}_p\left(U(k)\right) = U\left(\min_{q \in \mathcal{M}_f} D(q \parallel p)\right).$$

On the other side, using Fenchel inequality (2) and exploiting Theorem 2, we get

$$\sup_{k \in \mathcal{K}_{\mathcal{E}}} \mathbb{E}_p\left(U(k)\right) \le U\left(\inf_{q \in \mathcal{M}_{\mathcal{E}}} D(q \parallel p)\right).$$

In fact, for any $k \in \mathcal{K}_{\mathcal{E}} = \mathcal{M}_{\mathcal{E}}^0$ and $q \in \mathcal{M}_{\mathcal{E}}$, by definition of polar cone

$$\mathbb{E}_p\left(U(k)\right) \le \mathbb{E}_p\left(k e^{-D(q\|p)}\frac{q}{p}\right) + \mathbb{E}_p\left(V\left(e^{-D(q\|p)}\frac{q}{p}\right)\right) \le \mathbb{E}_p\left(V\left(e^{-D(q\|p)}\frac{q}{p}\right)\right),$$

where $V(x) = x \log(x) - x$ is the convex conjugate of the exponential function.

Since $\mathbb{E}_p\left(V\left(e^{-D(q\|p)}\frac{q}{p}\right)\right) = -e^{-D(q\|p)} = U(D(q \parallel p))$, we get

$$\sup_{k \in \mathcal{K}_{\mathcal{E}}} \mathbb{E}_p\left(U(k)\right) \le \inf_{q \in \mathcal{M}_{\mathcal{E}}} U(D(q \parallel p)) = U\left(\inf_{q \in \mathcal{M}_{\mathcal{E}}} D(q \parallel p)\right).$$

Finally, by Proposition 5, $q^* \in \mathcal{E}(p)$, from which we deduce that

$$U\left(\inf_{q \in \mathcal{M}_{\mathcal{E}}} D(q \parallel p)\right) = U\left(\min_{q \in \mathcal{M}_{\mathcal{E}}} D(q \parallel p)\right) = U\left(\min_{q \in \mathcal{M}_f} D(q \parallel p)\right)$$

and this concludes the proof. □

5. Conclusions

We studied maximal exponential models and their use in a mathematical convex optimization framework.

The main result characterizes the polar cone associated with the set of densities belonging to the maximal exponential model and under which a given convex set of random variables has non-positive expectation.

In a financial context, this set describes a "good" set of claims, which is both economically sound and well suited for exploiting convex duality techniques.

As an application we considered an exponential utility maximization problem in a continuous semimartingale incomplete market model and, under some technical assumptions, we proved a minimax theorem.

It turns out that the solution of the primal problem, which cannot be attained in the set of replicable claims by admissible strategies at zero initial cost, is reached on this new larger set of claims described by means of the maximal exponential model.

Further investigation in which the Kullback–Leibler divergence is replaced by other statistical divergences, such as Bregman's (see e.g. [15,16]), could be done. However, this interesting research would require a suitable theory with powerful tools like the Portmanteau Theorem.

Moreover, the dual results obtained in the present paper could probably be extended in the framework described by [17,18], where a portfolio optimization problem, which involves deformed exponentials, is investigated.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pistone, G.; Sempi, C. An infinite-dimensional geometric structure on the space of all the probability measures equivalent to a given one. *Ann. Stat.* **1995**, *23*, 1543–1561.
2. Vigelis, R.F.; Cavalcante, C.C. On φ-families of probability distributions. *J. Theor. Prob.* **2013**, *26*, 870–884.
3. De Andrade, L.H.F.; Vieira, F.L.J.; Vigelis, R.F.; Cavalcante, C.C. Mixture and Exponential Arcs on Generalized Statistical Manifold. *Entropy* **2018**, *20*, 147, doi:10.3390/e20030147.
4. Imparato, D.; Trivellato, B. Geometry of Extendend Exponential Models. In *Algebraic and Geometric Methods in Statistics*; Cambridge University Press: Cambridge, UK, 2009; pp. 307–326.
5. Cena, A.; Pistone, G. Exponential Statistical Manifold. *Ann. Inst. Stat. Math.* **2007**, *59*, 27–56.
6. Santacroce, M.; Siri, P.; Trivellato, B. New results on mixture and exponential models by Orlicz spaces. *Bernoulli* **2016**, *22*, 1431–1447, doi:10.3150/15-BEJ698.
7. Santacroce, M.; Siri, P.; Trivellato, B. Exponential models by Orlicz spaces and Applications. *Bernoulli* **2017**, in press.
8. Brigo, D.; Pistone, G. Projection based dimensionality reduction for measure valued evolution equations in statistical manifolds. *arXiv* **2016**, arXiv:1601.04189v.
9. Lods, B.; Pistone, G. Information geometry formalism for the spatially homogeneous Boltzmann equation. *Entropy* **2015**, *17*, 4323–4363.
10. Santacroce, M.; Siri, P.; Trivellato, B. On Mixture and Exponential Connection by Open Arcs. In *Geometric Science of Information*; Nielsen, F., Barbaresco, F., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 577–584, ISBN 978-3-319-68444-4.
11. Pistone, G. Examples of the application of nonparametric information geometry to statistical physics. *Entropy* **2013**, *15*, 4042–4065.
12. Biagini, S.; Frittelli, M. Utility maximization in incomplete markets for unboundee processes. *Financ. Stoch.* **2005**, *9*, 493–517.
13. Schachermayer, W. Optimal investment in incomplete markets when wealth may become negative. *Ann. Appl. Prob.* **2001**, *11*, 694–734.
14. Delbaen, F.; Grandits, P.; Rheinländer, T.; Samperi, D.; Schweizer, M.; Stricker, C. Exponential hedging and entropic penalties. *Math. Financ.* **2002**, *12*, 99–123.

15. Nock, R.; Magdalou, B.; Briys, E.; Nielsen, F. Mining Matrix Data with Bregman Matrix Divergences for Portfolio Selection In *MAtrix Information Geometry*; Nielsen, F., Bathia, R., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 373–402.

16. Nock, R.; Magdalou, B.; Briys, E.; Nielsen, F. On tracking portfolios with certainty equivalents on a generalization of Markowitz model: The fool, the wise and the adaptive. In Proceedingsof the 28th International Conference on Machine Learning, Bellevue, WA, USA, 28 June–2 July 2011; Omnipress: Madison, WI, USA, 2011; pp. 73–80.

17. Rodrigues, A.F.P.; Cavalcante, C.C. Principal Curves for Statistical Divergences and an Application to Finance. *Entropy* **2018**, *20*, 333, doi:10.3390/e20050333.

18. Rodrigues, A.F.P.; Guerreiro, M.; Cavalcante, C.C. Deformed Exponentials and Portfolio Selection. *Int. J. Mod. Phys. C* **2018**, *29*, doi:10.1142/S0129183118500298.

19. Frittelli, M. The minimal entropy martingale measure and the valuation problem in incomplete markets. *Math. Financ.* **2000**, *10*, 39–52.