CARF Working Paper

CARF-F-319

A ROBUST VERSION OF CONVEX INTEGRAL FUNCTIONALS

Keita Owari
The University of Tokyo

May 2013

CARF is presently supported by Bank of Tokyo-Mitsubishi UFJ, Ltd., Dai-ichi Mutual Life Insurance Company, Meiji Yasuda Life Insurance Company, Nomura Holdings, Inc. and Sumitomo Mitsui Banking Corporation (in alphabetical order). This financial support enables us to issue CARF Working Papers.

A ROBUST VERSION OF CONVEX INTEGRAL FUNCTIONALS

KEITA OWARI

Graduate School of Economics, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

We consider the pointwise supremum of a family of convex integral functionals on L^∞, each associated to a common convex integrand and a respective probability measure belonging to a dominated weakly compact convex set. Its conjugate functional is analyzed, providing a pair of upper and lower bounds as direct sums of common regular part and respective singular parts, which coincide when the defining set of probabilities is a singleton, as the classical Rockafellar theorem, and these bounds are generally the best in a certain sense. We then investigate when the conjugate eliminates the singular measures, which a fortiori implies the equality of the upper and lower bounds, and its relation to other finer regularity properties of the original functional and of the conjugate. As an application, a general duality result in the robust utility maximization problem is obtained.

1. INTRODUCTION

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $f : \Omega \times \mathbb{R} \to (-\infty, \infty]$ a measurable mapping with $f(\omega, \cdot)$ convex for a.e. ω. Then $X \mapsto I_f(X) := \mathbb{E}[f(\cdot, X)] = \int_\Omega f(\omega, x)\mathbb{P}(d\omega)$ defines a convex functional on $L^\infty := L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, called a *convex integral functional*. Among many others, Rockafellar obtained in [32] that under mild integrability assumptions on f, the conjugate $I_f^*(\nu) = \sup_{X \in L^\infty}(\nu(X) - I_f(X))$ of I_f is expressed as the direct sum of regular and singular parts (which we call the *Rockafellar theorem*):

$$I_f^*(\nu) = I_{f^*}(d\nu_r/d\mathbb{P}) + \sup_{X \in \mathrm{dom}(I_f)} \nu_s(X), \quad \forall \nu \in (L^\infty)^*, \tag{1.1}$$

where $f^*(\omega, y) := \sup_x(xy - f(\omega, x))$ and ν_r (resp. ν_s) denotes the regular (resp. singular) part of $\nu \in (L^\infty)^*$ (see the end of this section for unexplained notation). In particular, if I_f is finite everywhere on L^∞, the conjugate I_f^* "eliminates" the *singular elements* of $(L^\infty)^*$ in that $I_f^*(\nu) = \infty$ unless ν is σ-additive. The finiteness of I_f implies also the $\sigma(L^1, L^\infty)$-compactness of all the lower level sets of I_f^* (inf-compactness) and the weak* subdifferentiability of I_f at everywhere on L^∞ and so on (see e.g. [34]).

This paper is concerned with a *robust version* of integral functionals. Let $\mathcal{P}$ be a set of probability measures on $(\Omega, \mathcal{F})$ dominated by $\mathbb{P}$ and f as above. Then we let

$$\mathcal{I}_{f,\mathcal{P}}(X) := \sup_{P \in \mathcal{P}} E_P[f(\cdot, X)] = \sup_{P \in \mathcal{P}} \int_\Omega f(\omega, X(\omega))P(d\omega), \ X \in L^\infty.$$

This is the pointwise supremum of the family of classical integral functionals $I_{f,P}(\cdot) = \int_\Omega f(\omega, \cdot)P(d\omega)$, $P \in \mathcal{P}$, so in particular, $\mathcal{I}_{f,\mathcal{P}}$ is convex, lower semicontinuous if so is each $I_{f,P}$, and is even norm-continuous as soon as it is finite everywhere. On the other

---

E-mail address: keita.owari@gmail.com.
Date: This version: 26th May, 2013.
Key words and phrases. Convex Integral Functionals, Convex Duality Method, Robust Utility Maximization.

hand, the description of its convex conjugate $\mathcal{I}^*_{f,\mathcal{P}}$ is not easy. The aim of this paper is to obtain an analogue of the Rockafellar theorem for the robust versions convex integral functionals of this type, as well as to explore the possibility of having finer regularity properties of $\mathcal{I}_{f,\mathcal{P}}$ and $\mathcal{I}^*_{f,\mathcal{P}}$ which are necessary to build a convex duality method for robust convex optimization problems *without involving singular measures*.

In application, a number of stochastic optimization problems are formulated as or reduced to the minimization of an integral functional I_f with suitably chosen f. Then (1.1) combined with the Fenchel duality theorem provides us the precise form of the corresponding *dual problem* as well as conditions for the existence of a "nice solution" to the dual. See e.g. [33] and [28] for general treatments, and e.g. [2], [3] for applications to optimal investment problems in financial mathematics where this type of convex duality technique collaborates with the martingale duality in probability theory.

In recent studies in mathematical finance, the growing awareness of the *model uncertainty* has led to *robust* formulations of those optimization problems. In this context, a probability P is considered as a model under which the quality of a control X is evaluated via the integral functional $I_{f,P}(X) = E_P[f(\cdot, X)]$, while the specification of the model is often unreliable in practice. Then a common *robust* formulation is to take a whole family of possible models as $\mathcal{P}$, then to optimize the worst case which corresponds to the minimization of $\mathcal{I}_{f,\mathcal{P}}(X) = \sup_{P\in\mathcal{P}} I_{f,P}(X)$. The initial motivation of the study of our robust version of integral functionals is to provide an *efficient* way to convex duality methods for this sort of robust optimization problems.

The main results of the paper are summarized as follows. After giving the precise formulation and basic properties of $\mathcal{I}_{f,\mathcal{P}}$ in Section 2, we obtain in Section 3.1 a pair of upper and lower bounds for $\mathcal{I}^*_{f,\mathcal{P}}(\nu)$ ($\nu \in (L^\infty)^*$), both of which are of forms analogous to (1.1) with the difference in the singular parts (Theorem 3.1). These bounds coincide when $\mathcal{P}$ is a singleton as exactly the Rockafellar theorem. In the general case, however, a counter example (Example 3.2) shows that $\mathcal{I}^*_{f,\mathcal{P}}(\nu)$ can be located strictly between the upper and lower bounds, and that those estimates are generally the best in a certain sense. Next, we investigate, in Section 3.2, when the conjugate $\mathcal{I}^*_{f,\mathcal{P}}$ eliminates the singular measures which *a fortiori* implies that the lower and upper bounds agree. In contrast to the classical case, the everywhere finiteness of $\mathcal{I}_{f,\mathcal{P}}$ is not enough. The estimates in Theorem 3.1 provide us a simple sufficient condition (even necessary given the finiteness) in a form of "uniform integrability" condition. The latter condition will turn out to be equivalent to some other finer regularity properties of $\mathcal{I}_{f,\mathcal{P}}$ and $\mathcal{I}^*_{f,\mathcal{P}}$, especially to the weak inf-compactness of $\mathcal{I}^*_{f,\mathcal{P}}$ and the everywhere weak*-subdifferentiability of $\mathcal{I}_{f,\mathcal{P}}$ (Theorem 3.4 and its Corollaries). Some examples are provided in Section 3.3. In Section 4, we apply our main results to an abstract version of *robust utility maximization*, providing the key duality result as well as a representation of a robust version of *utility indifference price* of a claim.

1.1. Basic Notation

We use the probabilistic notation. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $L^0 := L^0(\Omega, \mathcal{F}, \mathbb{P})$ denote the space of (equivalence classes modulo $\mathbb{P}$-a.s. equality of) $\mathbb{R}$-valued (finite) random variables defined on it. As usual, we do not distinguish a random variable and the equivalence class it generates. The $\mathbb{P}$-expectation of $X \in L^0$ is denoted by $\mathbb{E}[X] (:= \int_\Omega X(\omega)\mathbb{P}(d\omega))$ and we write $L^p := L^p(\Omega, \mathcal{F}, \mathbb{P})$ and $\|\cdot\|_p := \|\cdot\|_{L^p}$ for each $1 \leq p \leq \infty$. For other probability measures P absolutely continuous with respect to $\mathbb{P}$ ($P \ll \mathbb{P}$), we write $E_P[\cdot]$ for P-expectation, $L^p(P) := L^p(\Omega, \mathcal{F}, P)$ etc, explicitly indicating the probability that involves. Any probabilistic notation without reference to a probability is to be understood with respect to $\mathbb{P}$. Especially, "a.s." means "$\mathbb{P}$-a.s.", and the identification of random variables is always made by $\mathbb{P}$. We write $Q \sim P$ to mean Q and P are equivalent ($Q \ll P$ and $P \ll Q$). Also, for any $A \in \mathcal{F}$, $\mathbb{1}_A$ denotes its indicator function *in the sense of measure theory*, i.e., $\mathbb{1}_A(\omega) = 1$ if $\omega \in A$, and otherwise 0.

The norm dual of L^∞ is $ba := ba(\Omega, \mathcal{F}, \mathbb{P})$, the space of all *bounded finitely additive signed measures* ν respecting $\mathbb{P}$-null sets, i.e., $\sup_{A\in\mathcal{F}} |\nu(A)| < \infty$, $\nu(A \cup B) = \nu(A) + \nu(B)$ if $A \cap B = \emptyset$, and $\nu(A) = 0$ if $\mathbb{P}(A) = 0$ ([42, Ch. IV, Sec. 9] or [15, pp. 354-357]). The bilinear form of (L^∞, ba) is given by the (Radon) integral $\nu(X) = \int_\Omega X d\nu$ which agree with the usual integral when ν is σ-additive. A $\nu \in ba$ is said to be *purely finitely additive* if there exists a sequence (A_n) in $\mathcal{F}$ such that $\mathbb{P}(A_n) \nearrow 1$ but $|\nu|(A_n) = 0$ for all n. Then any $\nu \in ba$ admits a unique *Yosida-Hewitt decomposition* $\nu = \nu_r + \nu_s$ where ν_r is the *regular* (σ-additive) part, and ν_s is the *purely finitely additive part* (e.g. [9, Th. III.7.8]). We denote by ba_+ (resp. ba^σ, ba^s) the set of elements of ba which are positive (resp. σ-additive, singular), and $ba^\sigma_+ := ba_+ \cap ba^\sigma$ etc.

Finally, we use the convention to identify a σ-additive finite signed measure $\nu \ll \mathbb{P}$ with its Radon-Nikodým density $d\nu/d\mathbb{P}$, thus ba^σ is regarded as L^1. In particular, the set of all probability measures $P \ll \mathbb{P}$ is identified with $\{Z \in L^1 : Z \geq 0, \mathbb{E}[Z] = 1\}$.

2. Robust Version of Integral Functionals: Basic Properties

Throughout, we work with a set $\mathcal{P}$ of probability measures $P \ll \mathbb{P}$ which we assume

$$\mathcal{P} \text{ is convex and } \sigma(L^1, L^\infty)\text{-compact in } L^1. \tag{2.1}$$

This means, of course, that $\{dP/d\mathbb{P} : P \in \mathcal{P}\}$ is convex and weakly compact in L^1, and the weak compactness of a convex set in L^1 is equivalent to the norm-closedness plus the *uniform integrability* (the Dunford-Pettis theorem). We assume also for non-redundancy

$$\mathcal{P} \sim \mathbb{P} \text{ in the sense that for all } A \in \mathcal{F}, \mathbb{P}(A) = 0 \Leftrightarrow P(A) = 0, \ \forall P \in \mathcal{P}. \tag{2.2}$$

This implies in particular that for any random variable X (not *a priori* assumed finite-valued), $\sup_{P\in\mathcal{P}} E_P[|X|] < \infty$ implies $X \in L^0$ (i.e., a.s. finite). (2.2) is just a simplifying assumption. In fact, all of the results remain correct if the qualification "a.s." is replaced by "$\mathcal{P}$-q.s." (quasi surely $\Leftrightarrow$ P-a.s. for all $P \in \mathcal{P}$). The latter is further equivalent to the usual "a.s." with respect to another probability $\mathbb{P}' \ll \mathbb{P}$ satisfying (2.2), that we can construct under (2.1) by means of an exhaustion argument à la Halmos-Savage. But we do not prefer that the presentation gets notationally messy for straightforward generalization. So, *in the sequel, (2.1) and (2.2) will be always in force without further notice*.

Let us introduce an "L^1-type" space with respect to $\mathcal{P}$:

$$L^1(\mathcal{P}) := \left\{ X \in L^0 : \|X\|_{1,\mathcal{P}} := \sup_{P\in\mathcal{P}} E_P[|X|] < \infty \right\}. \tag{2.3}$$

It is easy to see that $L^1(\mathcal{P})$ is a *solid* vector space ($|X| \leq |Y|$ a.s. and $Y \in L^1(\mathcal{P}) \Rightarrow X \in L^1(\mathcal{P})$), and is equal to L^1 if $\mathcal{P} = \{\mathbb{P}\}$. In fact, $L^1(\mathcal{P})$ with the norm $\|\cdot\|_{1,\mathcal{P}}$ is a Banach space, but we do not use this here. We introduce another "L^1-type" space:

$$L^1_u(\mathcal{P}) := \left\{ X \in L^0 : \lim_{N\to\infty} \|X \mathbb{1}_{\{|X|\geq N\}}\|_{1,\mathcal{P}} = 0 \right\}. \tag{2.4}$$

Again $L^1_u(\mathcal{P})$ agree with L^1 (hence with $L^1(\mathcal{P})$) if $\mathcal{P} = \{\mathbb{P}\}$. In general, $L^1_u(\mathcal{P})$ is a closed subspace of $L^1(\mathcal{P})$ (hence itself a Banach space), and the inclusion can be strict.

Example 2.1 ($L^1_u(\mathcal{P}) \subsetneq L^1(\mathcal{P})$, cf. [8]). Let us consider the probability space $(\mathbb{N}, 2^{\mathbb{N}}, \mathbb{P})$ where $\mathbb{P}(\{n\}) = 2^{-n}$. For the set $\mathcal{P}$, we take the closed convex hull (in L^1) of a sequence of probabilities given by $P_1(\{1\}) = 1$, $P_n(\{1\}) = 1 - 1/n$ and $P_n(\{n\}) = 1/n$. Then it is easy to see that $\sup_{P\in\mathcal{P}} E_P[X] = \sup_n E_{P_n}[X]$ if $X \geq 0$, and $\mathcal{P}$ is weakly compact. Now consider a random variable ξ given by $\xi(n) = n$. Then $\sup_n E_{P_n}[\xi] = \sup_n \{1 \cdot (1 - 1/n) + n/n\} = \sup_n \{2 - 1/n\} = 2 < \infty$, hence $\xi \in L^1(\mathcal{P})$, while $\sup_n E_{P_n}[\xi \mathbb{1}_{\{|\xi|\geq N\}}] = \sup_n \mathbb{1}_{\{n\geq N\}} = 1$ for every $N \geq 2$, hence $\xi \notin L^1_u(\mathcal{P})$.

The subscript "u" stands for "uniformly integrable" in view of the next lemma.

Lemma 2.2. *$X \in L^1_u(\mathcal{P})$ if and only if $\{XdP/d\mathbb{P}\}_{P\in\mathcal{P}}$ is uniformly integrable. In particular, $L^1_u(\mathcal{P})$ is a solid vector subspace of $L^1(\mathcal{P})$.*

Proof. If $\{XdP/d\mathbb{P}\}_{P\in\mathcal{P}}$ is uniformly integrable, then $X \in L^0$ by (2.2), thus $\mathbb{P}(|X| > N) \to 0$, and hence $\sup_{P\in\mathcal{P}} E_P[|X|\mathbb{1}_{\{|X|>N\}}] = \sup_{P\in\mathcal{P}} \mathbb{E}[|XdP/d\mathbb{P}|\mathbb{1}_{\{|X|>N\}}] \to 0$.

Conversely, suppose $X \in L^1_u(\mathcal{P})$, and observe that for any $A \in \mathcal{F}$,

$$\begin{aligned}\sup_{P\in\mathcal{P}} E[|XdP/d\mathbb{P}|\mathbb{1}_A] &\leq \sup_{P\in\mathcal{P}} E_P[|X|\mathbb{1}_{A\cap\{|X|\geq N\}}] + \sup_{P\in\mathcal{P}} E_P[|X|\mathbb{1}_{A\cap\{|X|<N\}}] \\ &\leq \sup_{P\in\mathcal{P}} E_P[|X|\mathbb{1}_{\{|X|\geq N\}}] + N \sup_{P\in\mathcal{P}} P(A).\end{aligned}$$

For any $\varepsilon > 0$, the first term can be made less than $\varepsilon/2$ for a large N since $X \in L^1_u(\mathcal{P})$, while the uniform integrability of $\mathcal{P}$ shows the existence of $\delta > 0$ such that $\mathbb{P}(A) \leq \delta$ implies $\sup_{P\in\mathcal{P}} P(A) \leq 1/2N$. Summing up, $\sup_{P\in\mathcal{P}} \mathbb{E}[|XdP/d\mathbb{P}|\mathbb{1}_A] \leq \varepsilon$ whenever $\mathbb{P}(A) \leq \delta$, which establishes the uniform integrability of $\{XdP/d\mathbb{P}\}_{P\in\mathcal{P}}$. □

A sufficient condition for $X \in L^1_u(\mathcal{P})$ is that $|X|^p \in L^1(\mathcal{P})$ for a $p > 1$. Indeed, $E_P[|X|\mathbb{1}_{\{|X|>N\}}] \leq E_P[|X|^p]^{1/p} P(|X| > N)^{1/q}$ (with $\frac{1}{p} + \frac{1}{q} = 1$) by Hölder's inequality, hence $\sup_{P\in\mathcal{P}} E_P[|X|\mathbb{1}_{\{|X|\geq N\}}] \leq \sup_{P\in\mathcal{P}} E_P[|X|^p]^{1/p} \sup_{P'\in\mathcal{P}} P'(|X| \geq N)^{1/q}$, while $\sup_{P'\in\mathcal{P}} P'(|X| \geq N) \to 0$ by (2.1).

2.1. Robust Version of Integral Functionals and Fatou Property

Definition 2.3 (Normal convex integrands). A mapping $f : \Omega \times \mathbb{R} \mapsto \mathbb{R} \cup \{+\infty\}$ is called a *normal convex integrand* if f is $\mathcal{F} \otimes \mathcal{B}(\mathbb{R})$-measurable (where $\mathcal{B}(E)$ denotes the Borel σ-field on the topological space E), and the section $x \mapsto f(\omega, x)$ is a lower semicontinuous proper convex function for a.e. ω.

There are several formulations of normality equivalent to the above one when $\mathcal{F}$ is complete with respect to *some* measure μ, as we are assuming (w.r.t. $\mathbb{P}$). See [35, Ch. 14] for a general reference. Here are a couple of immediate but crucial consequences of normality:

1. If f is a normal, then for any $X \in L^0$, $\omega \mapsto f(\omega, X(\omega))$ is $\mathcal{F}$-measurable.
2. If f is normal, so is the "ω-wise" conjugate:

$$f^*(\omega, y) := \sup_{x\in\mathbb{R}} (xy - f(\omega, x)), \quad \forall y \in \mathbb{R}, \tag{2.5}$$

Given a normal convex integrand f and a set $\mathcal{P}$ verifying (2.1) and (2.2), we shall consider a robust analogue of convex integral functional

$$\mathcal{I}_{f,\mathcal{P}}(X) := \sup_{P\in\mathcal{P}} E_P[f(\cdot, X)] = \sup_{P\in\mathcal{P}} \int_\Omega f(\omega, X(\omega)) P(d\omega), \quad \forall X \in L^\infty. \tag{2.6}$$

If no confusion may arise, we suppress the subscript $\mathcal{P}$ and simply write $\mathcal{I}_f$. This functional is indeed well-defined as a proper convex functional on L^∞ if we assume:

$$\exists X_0 \in L^\infty \text{ s.t. } f(\cdot, X_0)^+ \in L^1(\mathcal{P}) \tag{2.7}$$

$$\exists Y_0 \in L^1(\mathcal{P}) \text{ s.t. } f^*(\cdot, Y_0)^+ \in L^1(\mathcal{P}). \tag{2.8}$$

Indeed, since $f(\cdot, X) \geq XY_0 - f^*(\cdot, Y_0)^+$ by Young's inequality, (2.8) implies $f(\cdot, X)^- \in L^1(\mathcal{P}) \subset \bigcap_{P\in\mathcal{P}} L^1(P)$ for any $X \in L^\infty$. Thus $\mathcal{I}_f(X) = \sup_{P\in\mathcal{P}} E_P[f(\cdot, X)]$ is well-defined and clearly convex on L^∞, and it is finite if and only if $f(\cdot, X)^+ \in L^1(\mathcal{P})$. In particular, $\mathcal{I}_f$ is proper by (2.7), and

$$\mathrm{dom}\mathcal{I}_f := \{X \in L^\infty : \mathcal{I}_f(X) < \infty\} = \{X \in L^\infty : f(\cdot, X)^+ \in L^1(\mathcal{P})\}. \tag{2.9}$$

We next check that $\mathcal{I}_f$ has a nice regularity on L^∞.

Lemma 2.4. *Assume (2.1), (2.7) and (2.8). Then $\mathcal{I}_f$ has the following property:*

$$\sup_n \|X_n\|_\infty < \infty,\ X_n \to X \text{ a.s.} \ \Rightarrow \mathcal{I}_f(X) \leq \liminf_n \mathcal{I}_f(X_n). \tag{2.10}$$

Proof. Suppose $a := \sup_n \|X_n\|_\infty < \infty$ and $X_n \to X$ a.s., then automatically $X \in L^\infty$. With Y_0 as in (2.8), $f(\cdot, X_n) \geq -a|Y_0| - f^*(\cdot, Y_0)^+ \in L^1(\mathcal{P})$. Hence for each $P \in \mathcal{P}$, Fatou's lemma shows that $E_P[f(\cdot, X)] \leq \liminf_n E_P[f(\cdot, X_n)]$, so

$$\sup_{P\in\mathcal{P}} E_P[f(\cdot, X)] \leq \sup_{P\in\mathcal{P}} \liminf_n E_P[f(\cdot, X_n)] \leq \liminf_n \sup_{P\in\mathcal{P}} E_P[f(\cdot, X_n)].$$

This establishes (2.10). □

Remark 2.5. The property (2.10) is called the *Fatou property* on L^∞, or the *order lower semicontinuity*, which is equivalent to the $\sigma(L^\infty, L^1)$-lower semicontinuity of $\mathcal{I}_f$. Indeed, the latter property is equivalent to the $\sigma(L^\infty, L^1)$-closedness of the level sets $\{X :\in L^\infty : \mathcal{I}_f(X) \leq c\}$ for all $c \geq \mathbb{R}$, while a *convex* subset A of L^∞ is $\sigma(L^\infty, L^1)$-closed if and only if $A \cap \{X : \|X\|_\infty \leq c\}$ is L^0-closed for every $c \geq 0$, as a consequence of the Krein-Šmulian and the Mackey-Arens theorems (cf. [13]). In particular, $\mathcal{I}_f$ is lower semicontinuous with respect to the norm topology as well.

2.2. Robust f^*-divergence

We proceed to the functional that plays the role of I_{f^*} in the classical case. Let

$$\tilde{f}^*(\omega, y, z) = \sup_{x\in\mathrm{dom} f(\omega,\cdot)} (xy - zf(\omega, x)). \tag{2.11}$$

Noting that $(a^-_{f^*}, a^+_{f^*}) \subset \mathrm{dom} f \subset [a^-_{f^*}, a^+_{f^*}]$ where $a^\pm_{f^*} := \lim_{k\to\pm} f^*(\cdot, k)/k$, a simple computation shows more explicitly that

$$\tilde{f}^*(\omega, y, z) = \begin{cases} 0 & \text{if } y = z = 0 \\ y \cdot a^-_{f^*}(\omega) & \text{if } y < 0, z = 0 \\ y \cdot a^+_{f^*}(\omega) & \text{if } y > 0, z = 0 \\ zf^*(\omega, y/z) & \text{if } z > 0. \end{cases} \tag{2.12}$$

Lemma 2.6. *$\tilde{f}^* : \Omega \times \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R} \cup \{+\infty\}$ is a normal convex integrand on $\mathbb{R} \times \mathbb{R}_+$, i.e., it is $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}_+)$-measurable and $(y, z) \mapsto \tilde{f}^*(\omega, y, z)$ is a lower semicontinuous proper convex function for a.e. $\omega \in \Omega$. Also, for a.e. $\omega \in \Omega$,*

$$xy \leq zf(\omega, x) + \tilde{f}^*(\omega, y, z), \quad \forall x \in \mathrm{dom} f(\omega, \cdot),\ y \in \mathbb{R},\ \forall z \geq 0. \tag{2.13}$$

Proof. Since f is a normal, there exists a sequence of measurable functions $(X_n)_{n\in\mathbb{N}} \subset L^0$ such that $(X_n(\omega))_n \cap \mathrm{dom} f(\omega,\cdot)$ is dense in $\mathrm{dom} f(\omega,\cdot)$ ([34, Proposition 2D]). Modifying the sequence as $\bar{X}_n := X_n \mathbb{1}_{\{f(\cdot,X_n)<\infty\}} + X_0 \mathbb{1}_{\{f(\cdot,X_n)=\infty\}}$ where $X_0 \in \mathrm{dom}\mathcal{I}_f$, we have for a.e. ω, $\bar{X}_n(\omega,\cdot) \in \mathrm{dom} f(\omega,\cdot)$ and $(\bar{X}_n(\omega))_n$ is dense in $\mathrm{dom} f(\omega,\cdot)$. Thus

$$\tilde{f}^*(\omega,y,z) = \sup_n (\bar{X}_n(\omega)y - zf(\cdot,\bar{X}_n(\omega))).$$

Consequently, $\tilde{f}^*$ is a normal convex integrand as the countable supremum of affine integrands with $\tilde{f}^*(\cdot,0,0) = 0$. (2.13) is obvious from the definition. $\square$

Note that the condition (2.8) implies that

$$\forall P \in \mathcal{P},\ \exists Y_P \in L^1 \text{ s.t. } \tilde{f}^*(\cdot,Y_P,dP/d\mathbb{P})^+ \in L^1. \tag{2.14}$$

Indeed, if Y_0 is as in (2.8), then $Y_P := Y_0 dP/d\mathbb{P} \in L^1$ and $\tilde{f}^*(\cdot,Y_P,dP/d\mathbb{P})^+ = f^*(\cdot,Y_0)^+ dP/d\mathbb{P} \in L^1$ (with the convention $0\cdot\infty = 0$). Now we define

$$\mathcal{H}_{f^*}(Y|P) := \mathbb{E}[\tilde{f}^*(\cdot,Y,dP/d\mathbb{P})], \quad \forall Y \in L^1,\ \forall P \in \mathcal{P}, \tag{2.15}$$

$$\mathcal{H}_{f^*}(Y|\mathcal{P}) := \inf_{P\in\mathcal{P}} \mathcal{H}_{f^*}(Y|P) = \inf_{P\in\mathcal{P}} \mathbb{E}[\tilde{f}^*(\cdot,Y,dP/d\mathbb{P})], \quad \forall Y \in L^1. \tag{2.16}$$

In view of identification of ba^σ and L^1, we also write $\mathcal{H}_{f^*}(\nu|\mathcal{P}) = \mathcal{H}_{f^*}\left(\frac{d\nu}{d\mathbb{P}}|\mathcal{P}\right)$ etc. Since $\tilde{f}^*(\cdot,y,1) = f^*(\cdot,y)$, we recover $\mathcal{H}_{f^*}(Y|\mathbb{P}) = \mathcal{H}_{f^*}(Y|\{\mathbb{P}\}) = I_{f^*}(Y)$ in case $\mathcal{P} = \{\mathbb{P}\}$. Under the above assumptions, $\mathcal{H}_{f^*}(\cdot|\cdot)$ and $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$ are well-defined.

Lemma 2.7. *Assume (2.7) and (2.8). Then $\mathcal{H}_{f^*}(\cdot|\cdot)$ and $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$ are well-defined as a proper convex functionals respectively on $L^1\times\mathcal{P}$ and L^1, and it holds*

$$\mathbb{E}[XY] \le \mathcal{I}_{f,\mathcal{P}}(X) + \mathcal{H}_{f^*}(Y|\mathcal{P}), \quad \forall X \in L^\infty,\ \forall Y \in L^1. \tag{2.17}$$

If, in addition, there exists an $X_0' \in L^\infty$ with $f(\cdot,X_0')^+ \in L^1_u(\mathcal{P})$, then both $\mathcal{H}_{f^*}(\cdot|\cdot)$ and $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$ are weakly lower semicontinuous on $L^1\times\mathcal{P}$ and L^1, respectively.

Proof. As a consequence of (2.2), $X \in \mathrm{dom}\mathcal{I}_f (\neq \emptyset)$ implies $X(\omega) \in \mathrm{dom} f(\omega,\cdot)$ for a.e. ω. Then by (2.13), we have for any $X \in \mathrm{dom}\mathcal{I}_f$, $Y \in L^1$ and $P \in \mathcal{P}$,

$$\tilde{f}^*(\cdot,Y,dP/d\mathbb{P}) \ge XY - \frac{dP}{d\mathbb{P}} f(\cdot,X) \in L^1. \tag{2.18}$$

Hence $\tilde{f}^*(\cdot,Y,dP/d\mathbb{P})^- \in L^1$ for any $(Y,P) \in L^1\times\mathcal{P}$, so $\mathcal{H}_{f^*}(\cdot|\cdot)$ is well-defined, convex (since $\tilde{f}^*$ is convex), and $> -\infty$ on $L^1\times\mathcal{P}$, while $\mathcal{H}_{f^*}(Y_P|P) < \infty$ for $Y_P \in L^1$ as in (2.14) ($\Leftarrow$ (2.8)). Also, by (2.18), for any $X \in \mathrm{dom}\mathcal{I}_f$ and $Y \in L^1$,

$$\inf_{P\in\mathcal{P}} \mathbb{E}[\tilde{f}^*(\cdot,Y,dP/d\mathbb{P})] \ge \inf_{P\in\mathcal{P}} (\mathbb{E}[XY] - E_P[f(\cdot,X)]) = \mathbb{E}[XY] - \mathcal{I}_f(X) > -\infty,$$

hence $\mathcal{H}_{f^*}(Y|\mathcal{P}) > -\infty$ for all $Y \in L^1$ and we have (2.17) (which is trivial if $X \notin \mathrm{dom}\mathcal{I}_f$). The convexity of $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$ follows from that of $\mathcal{H}_{f^*}(\cdot|\cdot)$ and $\mathcal{P}$.

We now suppose that $f(\cdot,X_0)^+ \in L^1_u(\mathcal{P})$ for some $X_0 \in L^\infty$ and let $\{(Y_n,P_n)\}_n \subset L^1\times\mathcal{P}$ converge in norm to (Y,P). Passing to a subsequence, we may assume that the convergence takes place also in the a.s. sense. Then (2.18) applied to X_0 as well as the L^1-convergence (hence uniform integrability) of $(Y_n)_n$ show that $\{\tilde{f}^*(\cdot,Y_n,dP_n/d\mathbb{P})^-\}_n$ is uniformly integrable, hence Fatou's lemma yields that

$$\mathcal{H}_{f^*}(Y|P) = \mathbb{E}[\tilde{f}^*(\cdot,Y,dP/d\mathbb{P})] \le \liminf_n \mathbb{E}[\tilde{f}^*(\cdot,Y_n,dP_n/d\mathbb{P})].$$

This implies that for any $c \in \mathbb{R}$, the lower level set $\{(Y, P) \in L^1 \times \mathcal{P} : \mathcal{H}_{f^*}(Y|P) \leq c\}$ is norm-closed, and hence weakly closed by the convexity. We thus deduce that $\mathcal{H}_{f^*}(\cdot|\cdot)$ is weakly lower semicontinuous.

Finally, suppose that $(Y_n)_n \subset L^1$ converges weakly to some $Y \in L^1$ and $\mathcal{H}_{f^*}(Y_n|\mathcal{P}) \leq c$ for all n. Since $P \mapsto \mathcal{H}(Y_n|P)$ is weakly lower semicontinuous (from the previous paragraph) and $\mathcal{P}$ is weakly compact, we can pick a $P_n \in \mathcal{P}$ so that $\mathcal{H}_{f^*}(Y_n|\mathcal{P}) = \mathcal{H}_{f^*}(Y_n|P_n)$ for each n. Then another application of the weak compactness enables us to pass to a subsequence (still denoted by $\{(Y_n, P_n)\}$) with P_n converging weakly to some $P \in \mathcal{P}$ as well, and consequently, $\mathcal{H}_{f^*}(Y|\mathcal{P}) \leq \mathcal{H}_{f^*}(Y|P) \leq \liminf_n \mathcal{H}_{f^*}(Y_n|P_n) \leq c$, by the weak lower semicontinuity of $\mathcal{H}_{f^*}(\cdot|\cdot)$. This establishes the weak lower semicontinuity of $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$. □

Remark 2.8 (Robust f^*-divergence etc)**.** A couple of remarks are in order.

1. When f^* (hence f as well) is non-random, $\mathcal{H}_{f^*}(\cdot|\cdot)$ is nothing other than the *f^*-divergence*, and $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$ is called the *robust f^*-divergence*. In this case, the integrability conditions (2.7) and (2.16) (even with $L^1_u(\mathcal{P})$ instead of $L^1(\mathcal{P})$) automatically hold, and Lemma 2.7 is standard (e.g. [11, Lemma 2.7]). We shall still use the terminologies f^*-divergence and robust f^*-divergence for *random* f^*.
2. The integrand f is finite-valued (in the sense that $\mathbb{P}(f(\cdot, x) < \infty, \forall x \in \mathbb{R}) = 1$) if and only if $\lim_{|y|\to\infty} f^*(\cdot, y)/y = +\infty$ a.s. In this case, $\mathcal{H}_{f^*}(\nu|P) < \infty$ implies $\nu \ll P$, or more precisely, (2.12) with $a^{\pm}_{f^*} = \pm\infty$ shows that

$$\mathcal{H}_{f^*}(\nu|P) = \begin{cases} E_P[f^*(\cdot, d\nu/dP)] & \text{if } \nu \ll P, \\ +\infty & \text{otherwise.} \end{cases}$$

3. Main Results

This section presents the main results of the paper. Recall that $\mathcal{I}_f := \mathcal{I}_{f,\mathcal{P}}$ and $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$ are defined respectively by (2.6) and (2.16), and that *(2.1) and (2.2) are always in force without particular mention.*

3.1. A Rockafellar-Type Theorem for the Convex Conjugate

We are interested in the convex conjugate of $\mathcal{I}_f$:

$$\mathcal{I}_f^*(\nu) := \sup_{X \in L^\infty} (\nu(X) - \mathcal{I}_f(X)), \quad \forall \nu \in ba = ba(\Omega, \mathcal{F}, \mathbb{P}).$$

We have the following *robust analogue* of the Rockafellar theorem for the conjugate $\mathcal{I}_f^*$. The proof will be given in Section 3.4.

Theorem 3.1. *Suppose (2.1), (2.2) and*

$$\exists X_0 \in L^\infty \text{ such that } f(\cdot, X_0)^+ \in L^1_u(\mathcal{P}) \tag{3.1}$$

$$\exists Y_0 \in L^1(\mathcal{P}) \text{ such that } f^*(\cdot, Y_0)^+ \in L^1(\mathcal{P}). \tag{2.8}$$

Then for any $\nu \in ba$ with the Yosida-Hewitt decomposition $\nu = \nu_r + \nu_s$,

$$\mathcal{H}_{f^*}(\nu_r|\mathcal{P}) + \sup_{X \in \mathcal{D}_f} \nu_s(X) \leq (\mathcal{I}_f)^*(\nu) \leq \mathcal{H}_{f^*}(\nu_r|\mathcal{P}) + \sup_{X \in \mathrm{dom}\mathcal{I}_f} \nu_s(X), \tag{3.2}$$

where $\mathrm{dom}\mathcal{I}_f = \{X \in L^\infty : \mathcal{I}_f(X) < \infty\}$ *and*

$$\mathcal{D}_f := \{X \in L^\infty : f(\cdot, X)^+ \in L^1_u(\mathcal{P})\}. \tag{3.3}$$

In the classical case $\mathcal{P} = \{\mathbb{P}\}$, both $L^1_u(\mathcal{P})$ and $L^1(\mathcal{P})$ agree with the standard L^1, hence (3.1) (resp. (2.8)) reduces to the existence of $X \in L^\infty$ with $f(\cdot, X)^+ \in L^1$ (resp. $Y \in L^1$ with $f^*(\cdot, Y)^+ \in L^1$), and the two inequalities in (3.2) reduce to a single equality. This is exactly the original Rockafellar theorem [32, Theorem 1] for the integral functional $I_f(X) = \mathbb{E}[f(\cdot, X)]$ on $L^\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$. The original version of [32] is a little bit more general, where the integral functional is defined with respect to a σ*-finite* (rather than probability) measure μ for $\mathbb{R}^d$-valued random variables $X \in L^\infty(\Omega, \mathcal{F}, \mu; \mathbb{R}^d)$. There are also some extensions with $L^\infty(\Omega, \mathcal{F}, \mu; \mathbb{R}^d)$ replaced by a certain class of decomposable spaces of measurable functions taking values in a Banach space, say E. See in this line [30], [18, 19], [4] and [34] for a general reference when $E = \mathbb{R}^d$.

With a non-trivial compact convex set $\mathcal{P}$, this type of result is new to the best of our knowledge. In this general case, a possible complaint would be the difference between the singular parts of the left and the right hand sides of (3.2). Since $\mathrm{dom}\mathcal{I}_f = \{X \in L^\infty : f(\cdot, X)^+ \in L^1(\mathcal{P})\}$ (see (2.9) and the paragraph that precedes), we have always $\mathcal{D}_f \subset \mathrm{dom}\mathcal{I}_f$, and the two sets coincide when $L^1_u(\mathcal{P}) = L^1(\mathcal{P})$ (especially when $\mathcal{P}$ is generated by a finite number of extreme points). But as seen in Example 2.1, this is not generally the case. Even then, one would ask some possibility (density argument etc) to obtain a single equality. At the level of generality of Theorem 3.1, however, both inequalities in (3.2) can really be strict:

Example 3.2 (Badly Behaving Integrand)**.** Consider the probability space $(\mathbb{N}, 2^{\mathbb{N}}, \mathbb{P})$ and the set $\mathcal{P} = \overline{\mathrm{conv}}(P_n; n \in \mathbb{N})$ of Example 2.1. Then, L^∞ is identified as the space l^∞ of bounded sequences with the norm $\|X\|_\infty = \sup_n |X(n)|$, and $\nu \in ba^s(\mathbb{N}, 2^{\mathbb{N}}, \mathbb{P})$ if and only if ν vanishes on any finite set, thus in particular

$$\nu \in ba^s_+ \Rightarrow \|\nu\| \cdot \liminf_n X(n) \le \nu(X) \le \|\nu\| \cdot \limsup_n X(n), \ \forall X \in L^\infty \simeq l^\infty.$$

Note that such $\nu \neq 0$ exists, thus $ba^+ \neq \emptyset$ (see [1, Ch. 16] for detail). Now we set

$$f(n, x) = nx^+ e^x, \quad \forall n \in \mathbb{N} = \Omega, \ \forall x \in \mathbb{R}. \tag{3.4}$$

Then $\mathcal{I}_f(X) = \sup_n \left(\left(1 - \frac{1}{n}\right) X(1)^+ e^{X(1)} + X(n)^+ e^{X(n)^+}\right) \le 2\|X\|_\infty e^{\|X\|_\infty}$, hence $\mathrm{dom}\mathcal{I}_f = L^\infty$, while $\lim_{N\to\infty} \sup_n E_{P_n}[f(\cdot, X)\mathbb{1}_{\{f(\cdot,X)\ge N\}}] = \limsup_n X(n)^+ e^{X(n)^+}$, hence $0 \in \mathcal{D}_f = \{X : \limsup_n X(n) \le 0\} \subsetneq \mathrm{dom}\mathcal{I}_f$ (see Lemma A.1). Thus, for $\nu \in ba^s_+$, $\sup_{X \in \mathcal{D}_f} \nu(X) = 0$ while $\sup_{X \in \mathrm{dom}\mathcal{I}_f} \nu(X) = +\infty$ (both are $+\infty$ when $\nu \in ba^s \setminus ba^s_+$).

As for $\mathcal{I}^*_f$, observe first that since f is increasing in x, $\mathcal{I}^*_f(\nu)$ is finite only if $\nu \ge 0$, and since $f^*(n, y) = \sup_{x>0} x(y - ne^x)$, $\mathcal{H}_{f^*}(0|\mathcal{P}) = 0$. Thus for $\nu \in ba^s_+$, (3.2) reads as $\sup_{X \in \mathcal{D}_f} \nu_s(X) \le (\mathcal{I}_f)^*(\nu_s) \le \sup_{X \in \mathrm{dom}(\mathcal{I}_f)} \nu_s(X)$. More explicitly (see Lemma A.2):

$$(\mathcal{I}_f)^*(\nu) = \sup_{x\ge 0} x(\|\nu_s\| - e^x), \ \forall \nu \in ba^s_+.$$

In particular, $\mathcal{I}^*_f(\nu) = 0$ if $\nu \in A' := \{\nu \in ba^s_+ : \|\nu\| = \nu(\mathbb{N}) \le 1\}$, $0 < \mathcal{I}^*_f(\nu) < \infty$ if $\nu \in A'' := \{\nu \in ba^s_+ : \|\nu\| > 1\}$, and $\lim_{\|\nu\|\to\infty, \nu \in A''} \mathcal{I}^*_f(\nu) = \infty$. Hence in this case,

1. $\mathcal{I}^*_f$ agrees with the lower bound $\sup_{X \in \mathcal{D}_f} \nu(X) = 0$ on a non-empty region $A' \subset ba^s$;
2. On another region A'', $\mathcal{I}^*_f$ is strictly between the upper and lower bounds and it runs through the whole interval of these bounds (in this specific case, $[0, \infty]$).

Therefore, one can not hope sharper bound than (3.2) in the full generality of Theorem 3.1. A little more detail will be given in Appendix.

Nevertheless, Theorem 3.1 is not so bad. First, the next is an immediate consequence of (3.2) restricted to ba^σ (i.e., to the case $\nu_s = 0$) and the Fatou property of $\mathcal{I}_f$ (Lemma 2.4).

Corollary 3.3 (Restriction to $ba^\sigma \simeq L^1$). *For any $\nu \in ba^\sigma$, we have*

$$(\mathcal{I}_f)^*(\nu) = \mathcal{H}_{f^*}(\nu|\mathcal{P}) = \inf_{P\in\mathcal{P}} E[\tilde{f}^*(\cdot, d\nu/d\mathbb{P}, dP/d\mathbb{P})].$$

In particular,

$$\mathcal{I}_f(X) = \sup_{Y\in L^1} \big(\mathbb{E}[XY] - \mathcal{H}_{f^*}(Y|\mathcal{P})\big), \quad X \in L^\infty. \tag{3.5}$$

Proof. The first assertion is clear. The second one is a consequence of $\sigma(L^\infty, L^1)$-lower semicontinuity (Lemma 2.4 and Remark 2.5) via the Fenchel-Moreau theorem. □

3.2. Fine Properties in the Finite-Valued Case

We now focus on the case where $\mathcal{I}_f$ is finite everywhere in L^∞. Then $\mathcal{I}_f$ is norm-continuous on the whole L^∞ as a finite-valued lower semicontinuous convex function on a Banach space (see [10, Ch.1, Corollary 2.5]), and the integrand f must be finite-valued.

In the classical case $\mathcal{P} = \{\mathbb{P}\}$, the finiteness of $I_f := \mathcal{I}_{\{\mathbb{P}\}}$ implies immediately that the singular part of I_f^* is trivial (i.e., it is 0 if $\nu \in L^1$, and otherwise $+\infty$), hence I_f^* reduces entirely to I_{f^*}. This property implies that all the lower level sets of I_{f^*} are $\sigma(L^1, L^\infty)$-compact, and the latter is in fact equivalent to the continuity of I_f for the Mackey topology $\tau(L^\infty, L^1)$. Also, I_f admits σ-additive subgradient at everywhere (weak* subdifferentiable). Consequently, regarding applications, we can work with the dual pair $\langle L^\infty, L^1\rangle$ rather than $\langle L^\infty, ba\rangle$. See e.g. [31].

In our robust case, the "triviality of the singular part" of $\mathcal{I}_f^*$ should be understood as the property that $\mathcal{I}_f^*$ *eliminates the singular measures*:

$$\forall \nu \in ba,\ (\mathcal{I}_f)^*(\nu) < \infty \ \Rightarrow\ \nu \text{ is } \sigma\text{-additive}, \tag{3.6}$$

since $\mathcal{I}_f^*$ *itself* is not guaranteed to be the direct sum of the regular and singular parts. In contrast to the classical case, Example 3.2 tells us that the finiteness of $\mathcal{I}_f$ is *not* enough for this property. A trivial sufficient condition is that $\mathcal{D}_f = L^\infty$, i.e., $f(\cdot, X)^+ \in L^1_u(\mathcal{P})$ for all $X \in L^\infty$, which is (possibly strictly) stronger than the finiteness of $\mathcal{I}_f$. If this is the case, we have *a fortiori* the "equality" in (3.2). In fact, given the finiteness of $\mathcal{I}_f$, the condition $\mathcal{D}_f = L^\infty$ turn out to be equivalent to the elimination of singular measures by $\mathcal{I}_f^*$, and these properties are further equivalent to several fine properties of $\mathcal{I}_f$ and of $\mathcal{I}_f^*$ which are implied solely by the finiteness in the classical case. We collect the key equivalences in the next theorem, and some remarks and immediate but useful consequences will follow it. The proof of the theorem will be given in Section 3.5.

Theorem 3.4. *Suppose (2.1), (2.2), (2.8) and (3.1) and that $\mathcal{I}_f$ is finite everywhere on L^∞. Then the following are equivalent:*

(1) $\mathcal{I}_f$ satisfies the Lebesgue property, that is,

$$\sup_n \|X_n\|_\infty < \infty \text{ and } X_n \to X \text{ a.s.} \ \Rightarrow \mathcal{I}_f(X) = \lim_n \mathcal{I}_f(X_n). \tag{3.7}$$

(2) $\mathcal{D}_f = L^\infty$, i.e., $f(\cdot, X)^+ \in L^1_u(\mathcal{P})$ for all $X \in L^\infty$;
(3) $\mathbb{R} \subset \mathcal{D}_f$, i.e., $f(\cdot, x)^+ \in L^1_u(\mathcal{P})$ for all $x \in \mathbb{R}$;
(4) $\mathcal{I}_f^*$ eliminates singular measures in the sense of (3.6);
(5) $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$ is $\sigma(L^1, L^\infty)$-inf-compact, i.e.,

$$\{Y \in L^1 : \mathcal{H}_{f^*}(Y|\mathcal{P}) \le c\} \text{ is } \sigma(L^1, L^\infty)\text{-compact for all } c \in \mathbb{R}. \tag{3.8}$$

(6) for each $X \in L^\infty$, $\sup_{Y\in L^1}(\mathbb{E}[YX] - \mathcal{H}_{f^*}(Y|\mathcal{P}))$ is finite and attained, i.e.,

$$\mathcal{I}_f(X) = \max_{Y\in L^1}\left(\mathbb{E}[XY] - \mathcal{H}_{f^*}(Y|\mathcal{P})\right) < \infty, \quad \forall X \in L^\infty. \tag{3.9}$$

Remark 3.5 (Lebesgue Property)**.** The Lebesgue property (3.7) often plays a key role in financial mathematics, and the equivalence between (1), (5) and (6) for *convex risk measures* (monotone decreasing (w.r.t. the a.s. pointwise order) convex function $\rho : L^\infty \to \mathbb{R}$ with the property $\rho(X + c) = \rho(X) - c$ for $c \in \mathbb{R}$) is known ([17] and [6]). For an arbitrary function on L^∞, (3.7) is equivalent to the *order-continuity* with respect to the partial order of a.s. pointwise inequality (see [1, Ch. 8] for the definition) while the Fatou property (2.10) is the order *lower semicontinuity*. Of course, (3.7) implies (2.10).

Several consequences of Theorem 3.4 deserve attention. In the sequel, we suppose all the assumptions of Theorem 3.4 without further notice. We noted in Remark 2.5 that the Fatou property (order lower semicontinuity) on L^∞ is equivalent to the weak*-lower semicontinuity, and the Lebesgue property is the order continuity with respect to the a.s. order. As a consequence of (1) ⇔ (5), we have a topological characterization of the Lebesgue property of $\mathcal{I}_f$.

Corollary 3.6. *$\mathcal{I}_f$ has the Lebesgue property on L^∞ if and only if it is (finite-valued and) continuous for the Mackey topology $\tau(L^\infty, L^1)$.*

Proof. Given the $\sigma(L^\infty, L^1)$-lower semicontinuity and finiteness of $\mathcal{I}_f$ (Lemma 2.4 and Remark 2.5), the inf-compactness of the conjugate $\mathcal{I}_f^*|_{L^1} = \mathcal{H}_{f^*}(\cdot|\mathcal{P})$ is equivalent to the continuity of $\mathcal{I}_f$ for the *Mackey topology* $\tau(L^\infty, L^1)$ (see e.g. [23, Propositions 1,2]). □

Item (6) of Theorem 3.4 (attainment of supremum) is alternatively stated in terms of *subdifferentiability*. Recall that the subdifferential of the functional $\mathcal{I}_f$ at the point $X \in L^\infty$ is the set $\partial_X\mathcal{I}_f$ of $\nu \in ba$, called subgradients of $\mathcal{I}_f$ at X, such that

$$\nu(X) - \mathcal{I}_f(X) \geq \nu(X') - \mathcal{I}_f(X'), \quad \forall X' \in L^\infty. \tag{3.10}$$

If this set is non-empty, we say that $\mathcal{I}_f$ is subdifferentiable at X. If further there exists $Y \in \partial_X\mathcal{I}_f \cap L^1$, we say that Y is a *σ-additive subgradient* of $\mathcal{I}_f$ at X.

Corollary 3.7. *The equivalent conditions (1) – (6) are further equivalent to*

(7) $\emptyset \neq \partial_X\mathcal{I}_f \subset L^1$ for each $X \in L^\infty$.

Proof. If we suppose (6) and if $Y \in L^1$ is a corresponding maximizer for $X \in L^\infty$, then $\mathbb{E}[XY] - \mathcal{I}_f(X) = \mathcal{I}_{f^*}(Y|\mathcal{P}) = \sup_{X'\in L^\infty}(\mathbb{E}[X'Y] - \mathcal{I}_f(X'))$ by Corollary 3.3, hence $Y \in \partial_X f$. Also, if $\nu \in \partial_X\mathcal{I}_f$, then $\mathcal{I}_f^*(\nu) = \sup_{X'\in L^\infty}(\nu(X') - \mathcal{I}_f(X')) \leq \nu(X) - \mathcal{I}_f(X) < \infty$, hence (4) (⇔ (6)) implies $\nu \in L^1$, thus $\emptyset \neq \partial_X\mathcal{I}_f \subset L^1$.

Conversely, if $Y \in \partial_X\mathcal{I}_f \cap L^1$, then $\mathbb{E}[XY] - \mathcal{I}_f(X) \geq \sup_{X'\in L^\infty}(\mathbb{E}[X'Y] - \mathcal{I}_f(X')) = \mathcal{H}_{f^*}(Y|\mathcal{P})$, hence $\mathbb{E}[XY] - \mathcal{H}_{f^*}(Y|\mathcal{P}) \geq \mathcal{I}_f(X) = \sup_{Y'\in L^1}(\mathbb{E}[XY'] - \mathcal{H}_{f^*}(Y'|\mathcal{P}))$ by (3.5), which implies that Y attains the supremum. □

The weak inf-compactness of the functional $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$ is of independent interest as a generalization of the *de la Vallée-Poussin criterion* for uniform integrability. Recall that a family $\mathcal{Y}$ in L^1 is uniformly integrable if and only if there exists a function $g : [0, \infty) \to (-\infty, \infty]$ such that $\lim_{y\to\infty} g(y)/y = \infty$ and $\sup_{Y\in\mathcal{Y}} \mathbb{E}[g(|Y|)] = \sup_{Y\in\mathcal{Y}} \mathcal{H}_g(|Y||\mathbb{P}) < \infty$ (e.g. [7, Theorem II.22]). Such a function (when exists) can be taken convex and can be supposed to be defined on the entire $\mathbb{R}$ replacing it by $g(|y|)$. The *coercivity condition* $\lim_{|y|\to\infty} g(y)/y = \infty$ is equivalent to saying that g^* is finite on the entire real line. Now we have as a consequence of (3) ⇔ (5):

Corollary 3.8. *A family $\mathcal{Y}$ in L^1 is uniformly integrable if and only if there exists a weakly compact convex set of probabilities $\mathcal{Q} \sim \mathbb{P}$ as well as a normal convex integrand g with $g^*(\cdot, x) \in L^1_u(\mathcal{Q})$, $\forall x \in \mathbb{R}$, such that $\sup_{Y \in \mathcal{Y}} \mathcal{H}_g(Y|\mathcal{Q}) < \infty$.*

Proof. Changing the roles of f and f^* and noting that $f^{**} = f$ since normal, the sufficiency is nothing but (3) $\Rightarrow$ (5), while the necessity follows from the de la Vallée-Poussin criterion since the singleton $\{\mathbb{P}\}$ is obviously convex and weakly compact. $\square$

Remember that when g is deterministic (non-random), the condition $g^*(\cdot, x)^+ \in L^1_u(\mathcal{P})$ is automatically true as soon as it is finite ($\Leftrightarrow$ g is coercive). In this case, a part of Corollary 3.8 is already obtained by [11], but with a different proof, in their study of *robust f-projection*. In financial mathematics and statistics, the divergence $\mathcal{H}_g(Q|P)$ is often considered as a "something like a *distance*" between two probability measures Q and P. Fixing a probability P and a set $\mathcal{Q}$ of probabilities, an element Q of $\mathcal{Q}$ that minimizes $\mathcal{H}_g(Q|P)$ is called a g-projection of P on the set $\mathcal{Q}$, and similarly, a $Q \in \mathcal{Q}$ that minimizes $\mathcal{H}_g(Q|\mathcal{P})$ is called a *robust g-projection of $\mathcal{P}$ onto $\mathcal{Q}$*. Now the following is an immediate consequence of Corollary 3.8 and the lower semicontinuity of $\mathcal{H}_g(\cdot|\mathcal{P})$, which generalizes the result of [11] to the case of random g.

Corollary 3.9. *Let g be a normal convex integrand with $g^*(\cdot, x) \in L^1_u(\mathcal{P})$ for all constants $x \in \mathbb{R}$. Then for any convex and norm closed set $\mathcal{Q}$ of probability measures absolutely continuous w.r.t. $\mathbb{P}$, there exists a robust g-projection of $\mathcal{Q}$ on $\mathcal{P}$.*

3.3. Examples of "Nice" Integrands

Practically, we need to check whether a given normal convex integrand f satisfies (2.8) and one of equivalent conditions in Theorem 3.4. When f is *non-random* finite convex function, $\mathbb{R} \subset \mathcal{D}_f$ is automatic, while taking a constant $y \in \mathrm{dom} f^* \neq \emptyset$, we see that $f^*(y) \in L^1(\mathcal{P})$. Here are a couple of ways to generate "nice" random integrands.

Example 3.10 (Random scaling)**.** Let $g : \mathbb{R} \to \mathbb{R}$ be a (non-random) finite convex function and W be a strictly positive random variable (i.e., $\mathbb{P}(W > 0) = 1$). Then put

$$f(\omega, x) := g(W(\omega)x), \quad \forall(\omega, x) \in \Omega \times \mathbb{R}. \tag{3.11}$$

In this case, $f^*(\omega, y) = g^*(y/W(\omega))$ and $\mathbb{R} \subset \mathcal{D}_f$ is true if (see Appendix A.1)

$$\exists\delta > 0,\ p > 1 \text{ such that } g(-\delta W^p)^+ \vee g(\delta W^p)^+ \in L^1_u(\mathcal{P}). \tag{3.12}$$

Also, (2.8) is satisfied as soon as $W \in L^1(\mathcal{P})$ since then $Y = yW \in L^1(\mathcal{P})$ with $y \in \mathrm{dom} g^*$ satisfies $f^*(\cdot, Y) = g^*(y) \in L^1(\mathcal{P})$. A couple of remarks are in order.

1. If $\lim_{x\to\infty} g(x) = +\infty$ (i.e., g is eventually increasing), (3.12) already implies $W \in L^1(\mathcal{P})$. Indeed, the assumption implies that $\mathrm{dom} g^*$ contains some $y > 0$,
$$\delta y E_P[W]^p \le E_P[y(\delta W^p)] \le E_P[g(\delta W^p)] + g^*(y),$$
thus $\|W\|_{1,\mathcal{P}} \le \frac{1}{\delta y}(\|g(\delta W^p)\|_{1,\mathcal{P}} + g^*(y)) < \infty$.
2. If g is monotone increasing, $g(-W^p)^+ \le g(0)^+$, thus the half of (3.12) is automatic.

Another type of transformation of normal convex integrands is the parallel shift.

Example 3.11 (Random parallel shift)**.** Let f be a finite normal convex integrand satisfying (2.8) and $\mathbb{R} \subset \mathcal{D}_f$, and B be a random variable. Then put

$$f_B(\omega, x) = f(\omega, x + B(\omega)), \quad (\omega, x) \in \Omega \times \mathbb{R}. \tag{3.13}$$

A direct computation shows $f_B^*(\cdot, y) = f^*(\cdot, y) - yB$, and letting $\Gamma_\alpha(x) = f(\cdot, \alpha x)^+/\alpha$, we have the following estimates (see Appendix A.1):

$$\frac{1+\varepsilon}{\varepsilon} f\left(\cdot, \frac{\varepsilon}{1+\varepsilon} x\right) - \Gamma_\varepsilon(-B) \leq f_B(\cdot, x) \leq \frac{\varepsilon}{1+\varepsilon} f\left(\cdot, \frac{1+\varepsilon}{\varepsilon} x\right) + \Gamma_{1+\varepsilon}(B), \tag{3.14}$$

$$f^*(\cdot, y) - \Gamma_{1+\varepsilon}(B) \leq f_B^*(\cdot, y) \leq f^*(\cdot, y) + \Gamma_\varepsilon(-B). \tag{3.15}$$

Thus if we suppose that

$$\exists \varepsilon > 0 \text{ such that } \Gamma_{1+\varepsilon}(B) \in L^1_u(\mathcal{P}) \text{ and } \Gamma_\varepsilon(-B) \in L^1(\mathcal{P}), \tag{3.16}$$

then $\mathbb{R} \subset \mathcal{D}_{f_B}$ and f_B satisfies (2.8). Moreover, (3.15) implies in this case that

$$\mathcal{H}_{f_B^*}(Y|\mathcal{P}) < \infty \Leftrightarrow \mathcal{H}_{f^*}(Y|\mathcal{P}) < \infty \Rightarrow YB \in L^1, \tag{3.17}$$

and $\mathcal{H}_{f_B^*}(\cdot|\mathcal{P})$ is explicitly given as

$$\mathcal{H}_{f_B^*}(Y|\mathcal{P}) = \begin{cases} \mathcal{H}_{f^*}(Y|\mathcal{P}) - \mathbb{E}[YB] & \text{if } \mathcal{H}_{f^*}(Y|\mathcal{P}) < \infty, \\ +\infty & \text{otherwise.} \end{cases} \tag{3.18}$$

We can combine the preceding two examples:

Example 3.12. Let $g : \mathbb{R} \to \mathbb{R}$, W and B be as in Examples 3.10 and 3.11, and put

$$h(\cdot, x) := g(Wx + B) = g(W(x + B/W))$$

This h satisfies (2.8) and $\mathbb{R} \subset \mathcal{D}_h$ if (g, W) satisfies (3.12) and (3.16) holds with $f = g$. Note that if we apply Example 3.11 to $f(\cdot, x) = g(Wx)$ and B/W, then for instance $f(\cdot, (1+\varepsilon)B/W) = g((1+\varepsilon)B)$.

3.4. Proof of Theorem 3.1

The upper bound is easy.

Proof of Theorem 3.1: the upper bound. Since $\nu(X) = \nu_r(X) + \nu_s(X) = \mathbb{E}[X d\nu_r/d\mathbb{P}] + \nu_s(X)$, the inequality (2.17) in Lemma 2.7 shows that

$$\begin{aligned} \mathcal{I}_f^*(\nu) &= \sup_{X \in \mathrm{dom}(\mathcal{I}_f)} (\nu(X) - \mathcal{I}_f(X)) = \sup_{X \in \mathrm{dom}(\mathcal{I}_f)} \big(\mathbb{E}[X d\nu_r/d\mathbb{P}] - \mathcal{I}_f(X) + \nu_s(X)\big) \\ &\overset{(2.17)}{\leq} \mathcal{H}_{f^*}(\nu_r|\mathcal{P}) + \sup_{X \in \mathrm{dom}(\mathcal{I}_f)} \nu_s(X) \end{aligned}$$

as claimed. $\square$

The lower bound is more involved. We split the proof into several lemmas.

Lemma 3.13. *For any $\nu \in ba$,*

$$\mathcal{I}_f^*(\nu) = \sup_{X \in L^\infty} (\nu(X) - \mathcal{I}_f(X)) \geq \inf_{P \in \mathcal{P}} \sup_{X \in \mathcal{D}_f} (\nu(X) - E_P[f(\cdot, X)]). \tag{3.19}$$

Proof. Observe that the mapping $(X, P) \mapsto \nu(X) - E_P[f(\cdot, X)]$ is finite on $\mathcal{D}_f \times \mathcal{P}$, concave in $X \in \mathcal{D}_f$ for each $P \in \mathcal{P}$, and convex and weakly lower semicontinuous ($\Leftrightarrow$ norm lsc since convex) in $P \in \mathcal{P}$. Indeed, let $(P_n)_n \subset \Lambda_c := \{P' \in \mathcal{P} : \nu(X) - E_{P'}[f(\cdot, X)] \leq c\}$ converge in L^1 to some $P \in \mathcal{P}$. Then we can suppose, passing to a subsequence, that the convergence takes place also in the almost sure sense (in terms of densities w.r.t. $\mathbb{P}$), while $\{f(\cdot, X)^+ dP_n/d\mathbb{P}\}_n$ is uniformly integrable by Lemma 2.2 since $X \in \mathcal{D}_f$. Thus (reverse) Fatou's lemma shows $\nu(X) - c \leq \limsup_n E_{P_n}[f(\cdot, X)] \leq$

$E_P[f(\cdot, X)]$, hence $P \in \Lambda_c$. Therefore, since $\mathcal{P}$ is weakly compact, we can apply a minimax theorem to conclude

$$\begin{aligned}\sup_{X\in L^\infty}(\nu(X)-\mathcal{I}_f(X)) &= \sup_{X\in L^\infty}\inf_{P\in\mathcal{P}}(\nu(X)-E_P[f(\cdot,X)])\\ &\geq \sup_{X\in\mathcal{D}_f}\inf_{P\in\mathcal{P}}(\nu(X)-E_P[f(\cdot,X)])\\ &= \inf_{P\in\mathcal{P}}\sup_{X\in\mathcal{D}_f}(\nu(X)-E_P[f(\cdot,X)])\end{aligned}$$

as claimed. □

Now the first inequality in (3.2) amounts to proving the next lemma:

Lemma 3.14. *For any $\alpha < \mathcal{H}_{f^*}(\nu_r|\mathcal{P})$ and $\beta < \sup_{X\in\mathcal{D}_f}\nu_s(X)$, we have*

$$\sup_{X\in\mathcal{D}_f}(\nu(X)-E_P[f(\cdot,X)]) > \alpha+\beta, \quad \forall P\in\mathcal{P}. \tag{3.20}$$

We need a measurable selection result:

Lemma 3.15. *Let $Y, Z, D \in L^1$ such that $D \geq 0$ a.s. and*

$$\tilde{f}^*(\cdot, Y, D) = \sup_{x\in\operatorname{dom} f}(xY - Df(\cdot,x)) > Z \text{ a.s.} \tag{3.21}$$

Then there exists an $\widehat{X} \in L^0$ such that

$$f(\cdot,\widehat{X}) < \infty \text{ a.s. and } \widehat{X}Y - Df(\cdot,\widehat{X}) \geq Z \text{ a.s.} \tag{3.22}$$

Proof of Lemma 3.15. This amounts to proving that the multifunction

$$S(\omega) := \{x \in \operatorname{dom} f(\omega,\cdot) : xY(\omega) - D(\omega)f(\omega,x) \geq Z(\omega)\}$$

admits a measurable selection. This is a measurable multifunction since $g(\omega,x) := D(\omega)f(\omega,x) - xY(\omega)$ (with the convention $0\cdot\infty = 0$) is a normal convex integrand (see [35, Prop. 14.44, Cor. 14.46]), and $S(\omega) = \operatorname{dom} f(\omega,\cdot) \cap \{x : g(\omega,x) \leq -Z(\omega)\}$. Also, on the set $A_D = \{D > 0\}$, we have more simply $S = \{x : f(\cdot,x) - x\frac{Y}{D} \leq -\frac{Z}{D}\}$ which is non-empty on A_D by (3.21), and *closed* since f is normal. Thus

$$S'(\omega) = \begin{cases} S(\omega) & \text{if } \omega \in A_D, \\ \emptyset & \text{if } \omega \in A_D^c, \end{cases}$$

is a closed-valued measurable multifunction with $\operatorname{dom} S' = \{\omega : S'(\omega) \neq \emptyset\} = A_D$, thus the standard measurable selection theorem [e.g. 35, Cor. 14.6] shows the existence of $X' \in L^0$ such that $X'(\omega) \in S'(\omega) = S(\omega)$ if $\omega \in A_D$.

A little subtlety is that $S(\omega)$ is not necessarily closed when $\omega \in A_D^c$. So we explicitly construct a selector. Let $a^{\pm}_{f^*}(\omega) = \lim_{x\to\pm\infty} f^*(\omega,x)/x$ as in (2.12) and recall that if $a^-_{f^*}(\omega) = a^+_{f^*}(\omega)$, then $\operatorname{dom} f(\omega,\cdot) = \{a^+_{f^*}(\omega)\}$, and otherwise $(a^-_{f^*}(\omega), a^+_{f^*}) \subset \operatorname{dom} f(\omega,\cdot) \subset [a^-_{f^*}(\omega), a^+_{f^*}]$. In the reminder of this proof, we fix an $\bar{X} \in \operatorname{dom}\mathcal{I}_f$ which exists by assumption.

We split A_D^c into four partitions, and construct a random variable satisfying (3.22) on each of them. Let $B = \{a^-_{f^*} = a^+_{f^*}\}$. On $A_D^c \cap B$, we have $a^+_{f^*} \in \operatorname{dom} f$ while (3.21) tells us simply that $a^+_{f^*}Y > Z$. Thus $a^+_{f^*}$ satisfies (3.22) on $A_D^c \cap B$.

On $A_D^c \cap B^c \cap \{Y = 0\} =: C_0$, (3.21) reads as $Z < 0$, hence $\bar{X} \in \operatorname{dom}\mathcal{I}_f$ satisfies $\bar{X}Y - Df(\cdot,\bar{X}) = 0 > Z$.

On $A_D^c \cap B^c \cap \{Y > 0\} =: C_+$, (3.21) reads via (2.12) as $a_{f*}^+ > Z/Y$. Thus setting $X_+'' := \frac{1}{2}a_{f*}^+ + \frac{1}{2}\left(\frac{Z}{Y} \vee a_{f*}^-\right)$ on C_+ (and $X_+'' = \bar{X}$ on C_+^c), we have $a_{f*}^+ > X_+'' > a_{f*}^-$, hence $X_+'' \in \operatorname{dom} f$, and

$$X_+''Y = \frac{1}{2}a_{f*}^+Y + \frac{1}{2}Y\left(\frac{Z}{Y} \vee a_{f*}^-\right) \geq \frac{1}{2}a_{f*}^+Y + \frac{1}{2}Z > Z.$$

on C_+. Similarly, on $C_- := A_D^c \cap B^c \cap \{Y < 0\}$, (3.21) reads as $a_{f*}^- < Z/Y$, and putting $X_-'' := \frac{1}{2}a_{f*}^- + \frac{1}{2}\left(\frac{Z}{Y} \wedge a_{f*}^+\right)$, we have $a_{f*}^- < X_-'' < a_{f*}^+$ and

$$\begin{aligned} X_-''Y &= \frac{1}{2}a_{f*}^-Y + \frac{1}{2}Y\left(\frac{Z}{Y} \wedge a_{f*}^+\right) = \frac{1}{2}a_{f*}^-Y + \frac{1}{2}(-Y)\left(\frac{Z}{-Y} \vee (-a_{f*}^+)\right) \\ &\geq \frac{1}{2}a_{f*}^-Y + \frac{1}{2}Z > Z. \end{aligned}$$

Summing up,

$$\begin{aligned} \hat{X} : &= X'\mathbb{1}_{A_D} + a_{f*}^+\mathbb{1}_{A_D^c \cap B} + X_-''\mathbb{1}_{A_D^c \cap B^c \cap \{Y<0\}} \\ &\quad + \bar{X}\mathbb{1}_{A_D^c \cap B^c \cap \{Y=0\}} + X_+''\mathbb{1}_{A_D^c \cap B^c \cap \{Y>0\}} \end{aligned}$$

is a desired measurable selection of S. □

Proof of Lemma 3.14. Note first that there exists by definition an element $X_s \in \mathcal{D}_f$ with $\nu_s(X_s) > \beta$. Also, there exists an increasing sequence (A_n) in $\mathcal{F}$ such that $\mathbb{P}(A_n) \nearrow 1$ and $|\nu_s|(A_n) = 0$ for each n, by the singularity of ν_s. In particular, for any $X \in L^\infty$, $\nu_s(X\mathbb{1}_{A_n} + X_s\mathbb{1}_{A_n^c}) = \nu_s(X_s) > \beta$.

For the regular part, we first note that for each $P \in \mathcal{P}$, there exists a $Z_P \in L^1$ such that

$$\mathbb{E}[Z_P] > \alpha \text{ and } Z_P < \tilde{f}^*\left(\cdot, \frac{d\nu_r}{d\mathbb{P}}, \frac{dP}{d\mathbb{P}}\right) \text{ a.s.} \tag{3.23}$$

(even if $\Phi := \tilde{f}^*\left(\cdot, \frac{d\nu_r}{d\mathbb{P}}, \frac{dP}{d\mathbb{P}}\right) \notin L^1$). Indeed, since $\Phi^- \in L^1$, choosing $\varepsilon > 0$ so that $\mathbb{E}[\Phi] - \varepsilon > \alpha$, we have $\lim_N \mathbb{E}[(\Phi - \varepsilon) \wedge N] > \alpha$ by the monotone convergence theorem, hence we can take a big N_0 so that $(\Phi - \varepsilon) \wedge N_0$ do the job.

Given (3.23), Lemma 3.15 shows the existence of $X_P^0 \in L^0$ with

$$f(\cdot, X_P^0) < \infty \text{ and } X_P^0 d\nu_r/d\mathbb{P} - f(\cdot, X_P^0)dP/d\mathbb{P} \geq Z_P \text{ a.s.} \tag{3.24}$$

Note that this X_P^0 need not be in $\mathcal{D}_f$ (not even in L^∞) in general. So we approximate X_P^0 by elements of $\mathcal{D}_f$. Set $B_n := \{|X_P^0| \leq n\} \cap \{|f(\cdot, X_P^0)| \leq n\}$, and $C_n := A_n \cap B_n$, then $\mathbb{P}(C_n) \nearrow 1$ by the first part of (3.24), and $|\nu_s|(C_n) = 0$ for each n. Put

$$X_P^n := X_P^0\mathbb{1}_{C_n} + X_s\mathbb{1}_{C_n^c}.$$

Then $X_P^n \in \mathcal{D}_f$ for each n. Indeed, $X_P^0\mathbb{1}_{C_n}$ and $\mathbb{1}_{C_n}f(\cdot, X_P^0)$ are bounded and $f(\cdot, X_s)^+ \in L_u^1(\mathcal{P})$ by construction. Thus $X_P^n \in L^\infty$ and $f(\cdot, X_P^n) = \mathbb{1}_{C_n}f(\cdot, X_P^0) + \mathbb{1}_{C_n^c}f(\cdot, X_s) \in L_u^1(\mathcal{P})$ by the linearity and solidness of $L_u^1(\mathcal{P})$. On the other hand,

$$\begin{aligned} &\mathbb{E}[X_P^n d\nu_r/d\mathbb{P}] - E_P[f(\cdot, X_P^n)] \\ &\quad = \mathbb{E}\left[\mathbb{1}_{C_n}\left(X_P^0\frac{d\nu_r}{d\mathbb{P}} - \frac{dP}{d\mathbb{P}}f(\cdot, X_P^0)\right)\right] + \mathbb{E}\left[\mathbb{1}_{C_n^c}\left(X_s\frac{d\nu_r}{d\mathbb{P}} - \frac{dP}{d\mathbb{P}}f(\cdot, X_s)\right)\right] \\ &\quad \geq \mathbb{E}[\mathbb{1}_{C_n}Z_P] + E\left[\mathbb{1}_{C_n^c}\left(X_s\frac{d\nu_r}{d\mathbb{P}} - \frac{dP}{d\mathbb{P}}f(\cdot, X_s)\right)\right] \\ &\quad = \mathbb{E}[Z_P] + \mathbb{E}[\mathbb{1}_{C_n^c}\Xi_P], \end{aligned}$$

where $\Xi_P := X_s d\nu_r/d\mathbb{P} - f(\cdot, X_s)dP/d\mathbb{P} - Z_P \in L^1$. Since $\nu_s(X_P^n) = \nu_s(X_s) > \beta$,

$$\sup_{X\in\mathcal{D}_f} (\nu(X) - E_P[f(\cdot, X)]) \geq \mathbb{E}[X_P^n d\nu_r/d\mathbb{P}] - E_P[f(\cdot, X_P^n)] + \nu_s(X_P^n)$$
$$\geq \mathbb{E}[Z_P] + \nu_s(X_s) + \mathbb{E}[\mathbb{1}_{C_n^c}\Xi_P],$$

for each n. Since $\lim_n \mathbb{E}[\mathbb{1}_{C_n^c}\Xi_P] = 0$, we have

$$\sup_{X\in\mathcal{D}_f} (\nu(X) - E_P[f(\cdot, X)]) \geq \mathbb{E}[Z_P] + \nu_s(X_s) > \alpha + \beta,$$

as claimed. □

We now complete the proof of the first inequality in (3.2).

Proof of Theorem 3.1: the lower bound. It just suffices to take $\alpha = \mathcal{H}_{f^*}(\nu_r|\mathcal{P}) - \varepsilon/2$ and $\beta = \sup_{X\in\mathcal{D}_f} \nu_s(X) - \varepsilon/2$ for each $\varepsilon > 0$, then we have

$$\inf_{P\in\mathcal{P}} \sup_{X\in\mathcal{D}_f} (\nu(X) - E_P[f(\cdot, X)]) \geq (\mathcal{H}_{f^*}(\nu_r|\mathcal{P}) - \varepsilon/2) + (\sup_{X\in\mathcal{D}_f} \nu_s(X) - \varepsilon/2)$$
$$= \mathcal{H}_{f^*}(\nu_r|\mathcal{P}) + \sup_{X\in\mathcal{D}_f} \nu_s(X) - \varepsilon,$$

for all $\varepsilon > 0$, and the proof is complete. □

3.5. PROOF OF THEOREM 3.4

The implication (2) ⇒ (3) is trivial since $\mathbb{R} \subset L^\infty$, and (2) ⇒ (4) is an immediate consequence of (3.2) of Theorem 3.1. Thus it suffices to prove: (1) ⇒ (2), (3) ⇒ (2), (4) ⇒ (5) ⇔ (6), and (5) ⇒ (1). Note that the finiteness of $\mathcal{I}_f$ (i.e., $\mathrm{dom}\mathcal{I}_f = L^\infty$) is equivalent to $f(\cdot, X)^+ \in L^1(\mathcal{P})$, which implies through (2.2) that $f(\cdot, X) < \infty$ a.s. for all $X \in L^\infty$.

Proof of Theorem 3.4: (1) ⇒ (2). For each $X \in L^\infty$ fixed, we have to prove that

$$\lim_N \|f(\cdot, X)^+\mathbb{1}_{\{f(\cdot,X)^+\geq N\}}\|_{1,\mathcal{P}} = \lim_N \sup_{P\in\mathcal{P}} E_P[f(\cdot, X)^+\mathbb{1}_{\{f(\cdot,X)^+>N\}}] = 0. \tag{3.25}$$

Let $A_N := \{f(\cdot, X)^+ \geq N\} \in \mathcal{F}$ $(N \in \mathbb{N})$. Note that $\mathbb{P}(A_N) \to 0$ since $\mathcal{I}_f$ is finite. Picking an $X_0 \in \mathcal{D}_f$, put $X_N := (X - X_0)\mathbb{1}_{A_N}$. Then $X_N + X_0 = X\mathbb{1}_{A_N} + X_0\mathbb{1}_{A_N^c}$, thus

$$f(\cdot, X)^+\mathbb{1}_{A_N} = f(\cdot, X_N + X_0) - f(\cdot, X_0) + f(\cdot, X_0)\mathbb{1}_{A_N}$$
$$\leq \frac{1}{\lambda}\{f(\cdot, \lambda X_N + X_0) - f(\cdot, X_0)\} + f(\cdot, X_0)^+\mathbb{1}_{A_N}$$
$$\leq \frac{1}{\lambda} f(\cdot, \lambda X_N + X_0) + \frac{1}{\lambda} f(\cdot, X_0)^- + f(\cdot, X_0)^+\mathbb{1}_{A_N}, \quad \forall\lambda > 1,$$

where the second line follows from the fact that if φ is convex and *null at zero*, then $\lambda\varphi(x) \leq \lambda\varphi(x)$ for $\lambda \geq 1$, which we applied to $x \mapsto f(\cdot, x + X_0) - f(\cdot, X_0)$. Thus

$$\|f(\cdot, X)^+\mathbb{1}_{A_N}\|_{1,\mathcal{P}} \leq \frac{1}{\lambda}\mathcal{I}_f(\lambda X_N + X_0) + \frac{1}{\lambda}\|f(\cdot, X_0)^-\|_{1,\mathcal{P}} + \|f(\cdot, X_0)^+\mathbb{1}_{A_N}\|_{1,\mathcal{P}}$$
$$\leq \frac{|\mathcal{I}_f(\lambda X_N + X_0) - \mathcal{I}_f(X_0)|}{\lambda} + \frac{\mathcal{I}_f(X_0) + \|f(\cdot, X_0)^-\|_{1,\mathcal{P}}}{\lambda} + \|f(\cdot, X_0)^+\mathbb{1}_{A_N}\|_{1,\mathcal{P}}.$$

The third term tends to 0 as $N \to \infty$ regardless to λ since $f(\cdot, X_0)^+ \in L^1_u(\mathcal{P})$, while the second term tends to 0 as $\lambda \to \infty$ regardless to N. Next, observe that for each $\lambda > 1$, $\sup_N \|\lambda X_N + X_0\|_\infty \leq \lambda\|X\|_\infty + 2\|X_0\|_\infty < \infty$ and $\lambda X_N + X_0 \to X_0$ a.s. since $\mathbb{P}(A_N) \to 0$. Therefore, the Lebesgue property (3.7) shows that for each $\lambda > 0$, $\lim_N |\mathcal{I}_f(\lambda X_N + X_0) - \mathcal{I}_f(X_0)| = 0$. Now (3.25) follows from a diagonal argument. □

Proof of Theorem 3.4: (3) ⇒ (2). This follows from a simple observation:

$$a \le X \le b \text{ a.s., } a, b \in \mathbb{R} \;\Rightarrow\; f(\cdot, X) \le f(\cdot, a)^+ + f(\cdot, b)^+. \tag{3.26}$$

Indeed, the assumption implies that there exists a $[0,1]$-valued random variable α such that $X = \alpha a + (1-\alpha)b$ a.s. Since $x \mapsto f(\omega, x)$ is convex for a.e. ω, we see that $f(\omega, X(\omega)) \le \alpha(\omega) f(\omega, a) + (1-\alpha(\omega)) f(\omega, b) \le f(\omega, a)^+ + f(\omega, b)^+$ for a.e. $\omega \in \Omega$. □

Proof of Theorem 3.4: (4) ⇒ (5). Since $\mathcal{I}_f : L^\infty \to \mathbb{R}$ is norm-continuous (as a finite lower semicontinuous convex function on a Banach space), there exists a $\delta > 0$ such that $\mathcal{I}_f$ is bounded above by c' (say) on the ball $B_\delta := \{X \in L^\infty : \|X\|_\infty \le \delta\}$. Thus

$$\sup_{X \in B_\delta} \nu(X) \le \sup_{X \in B_\delta} \mathcal{I}_f(X) + \mathcal{I}_f^*(\nu) \le c' + \mathcal{I}_f^*(\nu), \quad \forall \nu \in ba.$$

This shows that $\Lambda_c := \{\nu \in ba : \mathcal{I}_f^*(\nu) \le c\}$ is contained in the $\sigma(ba, L^\infty)$-compact set $B^\circ_{\delta/(c+c')} = \{\nu \in ba : \sup_{X \in B_{\delta/(c+c')}} \nu(X) \le 1\} = \{\nu \in ba : \|\nu\| \le (c+c')/\delta\}$. Since Λ_c is $\sigma(ba, L^\infty)$-closed by the lower semicontinuity of $\mathcal{I}_f^*$ w.r.t. the same topology, we see that Λ_c is $\sigma(ba, L^\infty)$-compact. But (3.6) as well as Corollary 3.3 show that Λ_c is contained in L^1, and actually $\Lambda_c = \{Y \in L^1 : \mathcal{H}_{f^*}(Y|\mathcal{P}) \le c\}$. Thus Λ_c is compact for $\sigma(L^1, L^\infty)$ since the latter is the restriction of $\sigma(ba, L^\infty)$ to L^1. □

Lemma 3.16. *For any $a > 0$, there exists a $\beta(a)$ such that for all $X \in L^\infty$, $Y \in L^1$,*

$$\|X\|_\infty \le a,\ \mathbb{E}[XY] - \mathcal{H}_{f^*}(Y|\mathcal{P}) \ge \mathcal{I}_f(X) - 1 \;\Rightarrow\; \mathcal{H}_{f^*}(Y|\mathcal{P}) \le \beta(a).$$

Proof. Since $\mathbb{E}[XY] \le \frac{1}{2}(\mathcal{I}_f(2X) + \mathcal{H}_{f^*}(Y|\mathcal{P}))$, we see that $\mathbb{E}[XY] - \mathcal{H}_{f^*}(Y|\mathcal{P}) \ge c$ implies $\mathcal{H}_{f^*}(Y|\mathcal{P}) \le \mathcal{I}_f(2X) - 2c$, and putting $c = \mathcal{I}_f(X) - 1$,

$$\mathbb{E}[XY] - \mathcal{H}_{f^*}(Y|\mathcal{P}) \ge \mathcal{I}_f(X) - 1 \;\Rightarrow\; \mathcal{H}_{f^*}(Y|\mathcal{P}) \le \mathcal{I}_f(2X) - 2\mathcal{I}_f(X) + 2.$$

We need to show that the right hand side is bounded above by a constant depending only on the norm $\|X\|_\infty$. By (3.26), the first term is dominated by $\|f(\cdot, -2\|X\|_\infty)^+\|_{1,\mathcal{P}} + \|f(\cdot, 2\|X\|_\infty)^+\|_{1,\mathcal{P}}$. Then picking a $Y_0 \in \mathrm{dom}\mathcal{H}_{f^*}(\cdot|\mathcal{P})$, we have $-2\mathcal{I}_f(X) \le \mathbb{E}[XY_0] + \mathcal{H}_{f^*}(Y_0|\mathcal{P}) \le \|X\|_\infty \|Y\|_1 + \mathcal{H}_{f^*}(Y|\mathcal{P})$. Summing up, the constant

$$\beta(a) := \|f(\cdot, -2a)^+\|_{1,\mathcal{P}} + \|f(\cdot, 2a)^+\|_{1,\mathcal{P}} + 2a\|Y_0\|_1 + 2\mathcal{H}_{f^*}(Y_0|\mathcal{P}) + 2 \tag{3.27}$$

do the job. □

Proof of Theorem 3.4: (5) ⇒ (6). By Lemma 3.16 (with β being as there),

$$\sup_{Y \in L^1} (\mathbb{E}[XY] - \mathcal{H}_{f^*}(Y|\mathcal{P})) = \sup\{\mathbb{E}[XY] - \mathcal{H}_{f^*}(Y|\mathcal{P}) : \mathcal{H}_{f^*}(Y|\mathcal{P}) \le \beta(\|X\|_\infty)\}.$$

The set $\{Y \in L^1 : \mathcal{H}_{f^*}(Y|\mathcal{P}) \le \beta(\|X\|_\infty)\}$ is weakly compact by (5), and $Y \mapsto \mathbb{E}[XY] - \mathcal{H}_{f^*}(Y|\mathcal{P})$ is weakly upper semicontinuous. Hence the supremum in the right hand side (hence the left hand side) is attained. □

Proof of Theorem 3.4: (5) ⇒ (1). Let $(X_n)_n \subset L^\infty$ be such that $\sup_n \|X_n\|_\infty =: a < \infty$ and $X_n \to X$ a.s. Then by Lemma 3.16, for all n, we can write

$$\mathcal{I}_f(X_n) = \sup_{Y \in \mathcal{C}} \left(\mathbb{E}[X_n Y] - \mathcal{H}_{f^*}(Y|\mathcal{P})\right)$$

where $\mathcal{C} := \{Y \in L^1 : \mathcal{H}_{f^*}(Y|\mathcal{P}) \leq \beta(a)\}$ does not depend on n, and is weakly compact (hence uniformly integrable) by (5). Hence denoting $Z_n = X - X_n$,

$$\begin{aligned}\left|\mathcal{I}_f(X) - \mathcal{I}_f(X_n)\right| &\leq \sup_{Y\in\mathcal{C}} \mathbb{E}[|Y||Z_n|] \\ &\leq \sup_{Y\in\mathcal{C}} \mathbb{E}[|Y|\mathbb{1}_{\{|Y|>N\}}|Z_n|] + \sup_{Y\in\mathcal{C}} \mathbb{E}[|Y|\mathbb{1}_{\{|Y|\leq N\}}|Z_n|] \\ &\leq a \sup_{Y\in\mathcal{C}} \mathbb{E}[|Y|\mathbb{1}_{\{|Y|>N\}}] + N\mathbb{E}[|Z_n|]\end{aligned}$$

The first term tends to zero as $N \to \infty$ by the uniform integrability regardless to n, while for fixed N, the second term tends to zero by the dominated convergence theorem. Thus a diagonal argument deduce the result. □

Proof of Theorem 3.4: (6) ⇒ (5). We appeal to a *perturbed James' theorem* [24, Th. 2], which states that if E is a real Banach space, $\varphi : E \to (-\infty, +\infty]$ is coercive, i.e., $\lim_{\|\xi\|\to\infty} \varphi(\xi)/\|\xi\| = +\infty$, and if the supremum $\sup_{\xi\in E}(\langle \xi, \eta\rangle - \varphi(\xi))$ is attained for all $\eta \in E^*$, all the level sets $\{\xi \in E : \varphi(\xi) \leq c\}$ are relatively $\sigma(E, E^*)$-compact. We apply this to $E = L^1$, $E^* = L^\infty$ and $\varphi = \mathcal{H}_{f^*}(\cdot|\mathcal{P})$. Since $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$ is $\sigma(L^1, L^\infty)$-lower semicontinuous, the level sets of $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$ are weakly closed, so we need only to check that $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$ is coercive. For each $Y \in L^1$, $\|Y\|_1 = \mathbb{E}[\mathrm{sgn}(Y)Y]$ where $\mathrm{sgn}(Y) = \mathbb{1}_{\{Y>0\}} - \mathbb{1}_{\{Y<0\}} \in L^\infty$, and $\mathcal{H}_{f^*}(Y|\mathcal{P}) = \sup_{X\in L^\infty}(\mathbb{E}[XY] - \mathcal{I}_f(X))$, thus

$$\begin{aligned}\mathcal{H}_{f^*}(Y|\mathcal{P}) &\geq \mathbb{E}[n\mathrm{sgn}(Y)Y] - \mathcal{I}_f(n\mathrm{sgn}(Y)) \\ &\geq n\|Y\|_1 - \|f(\cdot,-n)^+\|_{1,\mathcal{P}} - \|f(\cdot,n)^+\|_{1,\mathcal{P}},\end{aligned}$$

and note that the last two norms in the second line are finite for all n since $\mathcal{I}_f$ is supposed to be finite, hence $\frac{\mathcal{H}_{f^*}(Y|\mathcal{P})}{\|Y\|_1} \geq n - \frac{\|f(\cdot,-n)^+\|_{1,\mathcal{P}} + \|f(\cdot,n)^+\|_{1,\mathcal{P}}}{\|Y\|_1}$ for all n and $Y \in L^1$. This proves that $\mathcal{H}_{f^*}(\cdot|\mathcal{P})$ is coercive. □

4. Application: Duality in Robust Utility Maximization

In this section, we consider a key duality result in a robust optimal investment problem in mathematical finance, called the *robust utility maximization*, as an illustrative and motivating application. The basic problem is to maximize the *robust utility functional*

$$X \mapsto \inf_{P\in\mathcal{P}} E_P[U(X)]$$

over all admissible wealths X where U is a utility function and the set $\mathcal{P}$ of probabilities is understood as a set of candidate models in which one does not know the true one. See [12] for financial motivation of this problem. Generally speaking, the mathematics behind utility maximization problems (either robust or standard) is quite different depending on whether the utility function is finite on the entire real line or only on the positive half-line. Here we consider the former case (see [40], [39], [41] for the latter case), while we allow for the utility function itself to be random and for a claim (random endowment) to (present and) be unbounded. Also, we focus here only on the duality result in an abstract form, and some remarks on its consequences in a typical setting will be provided in Section 4.3.

4.1. Abstract Formulation

Again, a set $\mathcal{P}$ of probabilities $P \ll \mathbb{P}$ with (2.1) and (2.2) will be fixed throughout.

Let $U : \Omega \times \mathbb{R} \to \mathbb{R}$ be a jointly measurable mapping with $x \mapsto U(\omega, x)$ being *concave and increasing*. We call such U a *random utility function*. Making a change of sign, $f_U(\omega, x) = -U(\omega, -x)$ is a monotone increasing normal convex integrand. We denote its conjugate f_U^* by V, or more explicitly $V(\omega, y) = \sup_x (U(\omega, x) - xy)$, and write $\mathcal{H}_V = \mathcal{H}_{f_U^*}$. Then the basic problem is to maximize the *robust utility functional*

$$u(X) := \inf_{P \in \mathcal{P}} E_P[U(\cdot, X)] = -\mathcal{I}_{f_U}(-X),$$

over a convex cone $\mathcal{C} \subset L^\infty$. Suppose that

$$U(\cdot, x) \in L_u^1(\mathcal{P}), \quad \forall x \in \mathbb{R}; \tag{4.1}$$

$$\exists Y_0 \in L^1(\mathcal{P}) \text{ such that } V(\cdot, Y_0) \in L^1(\mathcal{P}), \tag{4.2}$$

which mean that f_U satisfies (2.8) and $\mathbb{R} \subset \mathcal{D}_{f_U}$. Then the concave functional u on L^∞ is well-defined, finite and $\tau(L^\infty, L^1)$-continuous (Corollary 3.6). We suppose also the *Inada condition* in the sense: there exists a $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ such that

$$\omega \in \Omega_0 \;\Rightarrow \begin{cases} U(\omega, \cdot) \text{ is strictly concave, continuously differentiable,} \\ \lim_{x \downarrow -\infty} U'(\omega, x) = +\infty,\ \lim_{x \uparrow +\infty} U'(\omega, x) = 0. \end{cases} \tag{4.3}$$

This implies on the conjugate side that

$$\omega \in \Omega_0 \;\Rightarrow \begin{cases} V(\omega, \cdot) \text{ is finite, strictly convex, differentiable on } (0, \infty), \\ \lim_{y \downarrow 0} V'(\omega, y) = -\infty \text{ and } \lim_{y \uparrow +\infty} V'(\omega, y) = +\infty. \end{cases} \tag{4.4}$$

Finally, given a convex cone $\mathcal{C} \subset L^\infty$, let $\mathcal{C}^\circ := \{Y \in L^1 : \mathbb{E}[XY] \le 1,\ \forall X \in \mathcal{C}\}$ (the polar of $\mathcal{C}$ for the duality $\langle L^\infty, L^1 \rangle$ with $\langle X, Y \rangle = \mathbb{E}[XY]$). In a typical setup in finance which we shall briefly review in Section 4.3, $\mathcal{C}$ is the set of claims that the investor can superhedge with zero initial cost by trading underlying assets. Then $\mathcal{C}^\circ$ consists of positive multiples of *local martingale measures*, and the existence of *strictly positive* element of $\mathcal{C}^\circ$, which are (positive multiples of) *equivalent martingale measures*, is equivalent to the absence of *arbitrage* (more precisely *no free lunch with vanishing risks* (NFLVR)). But here we keep $\mathcal{C}$ abstract, and put $\mathcal{C}_V^\circ := \mathcal{C}^\circ \cap \operatorname{dom}\mathcal{H}_V(\cdot|\mathcal{P})$, then suppose

$$\mathcal{C}_V^{\circ,e} := \{Y \in \mathcal{C}_V^\circ : \mathbb{P}(Y > 0) = 1\} \ne \emptyset. \tag{4.5}$$

Note that since $x \to U(\cdot, x)$ is increasing, $V(\omega, y) = +\infty$ if $y < 0$, hence $\mathcal{H}_V(Y|\mathcal{P}) < \infty$ implies that $Y \ge 0$ a.s., thus $\mathcal{C}_V^\circ$ contains only positive elements, and $\mathcal{C}_V^{\circ,e} \subset \mathcal{C}_V^\circ \setminus \{0\}$.

Lemma 4.1. *With the above notation and assumptions (4.1), (4.2), (4.3) and (4.5),*

$$\inf_{Y \in \mathcal{C}_V^{\circ,e}} \mathcal{H}_V(Y|\mathcal{P}) < \mathcal{H}_V(0|\mathcal{P}). \tag{4.6}$$

Proof. Let $Y_1 \in \mathcal{C}_V^{\circ,e}$, then $\mathcal{H}_V(Y_1|\mathcal{P}) < \infty$, thus there is nothing to prove if the right hand side of (4.6) is infinite. So we suppose it is finite. Then there exist $P_1, P_0 \in \mathcal{P}$ such that $\mathcal{H}_V(Y_1|\mathcal{P}) = \mathcal{H}_V(Y_1|P_1)$ and $\mathcal{H}_V(0|\mathcal{P}) = \mathcal{H}_V(0|P_0) = E_{P_0}[V(\cdot, 0)]$. Then $V(\cdot, 0) \in L^1(P_0)$, and $P_1 \sim \mathbb{P}$. Let $Z_1 = dP_1/d\mathbb{P}$, $Z_0 = dP_0/d\mathbb{P}$ and

$$\varphi(\omega, \lambda) := \widetilde{V}(\omega, \lambda Y_1(\omega), \lambda Z_1(\omega) + (1-\lambda) Z_0(\omega)).$$

Then $\lambda \mapsto \varphi(\cdot, \lambda)$ is convex, hence $\frac{\varphi(\cdot,\lambda) - \varphi(\cdot,0)}{\lambda}$ decreases to Ξ (say) as $\lambda \downarrow 0$. On $\{Z_0 = 0\}$, $\varphi(\cdot, \lambda) = \lambda Z_1 V(\cdot, Y_1/Z_1)$, hence $\Xi = Z_1 V(\cdot, Y_1/Z_1)$. On the set $\{Z_0 > 0\}$, we have $V(\cdot, 0) = U(\cdot, \infty) < \infty$ a.s. since $V(\cdot, 0) \in L^1(P_0)$, while for a.e. ω, $x \mapsto V(\omega, x)$ is differentiable and $V'(\omega, 0) = \lim_{y \downarrow 0} V'(\omega, 0) = -\infty$ by (4.4). Using the identity $U(\cdot, -V'(\cdot, y)) = V(\cdot, y) - yV'(\cdot, y)$, and $\xi_\lambda := \lambda Y_1/(\lambda Z_1 + (1-\lambda)Z_0) \downarrow 0$, we have

$$\begin{aligned} \frac{d\varphi(\cdot, \lambda)}{d\lambda} &= U\left(\cdot, -V'(\cdot, \xi_\lambda)\right)(Z_1 - Z_0) + V'(\cdot, \xi_\lambda) Y_1 \\ &\downarrow U(\cdot, \infty)(Z_1 - Z_0) + V'(\cdot, 0) Y_1 = -\infty \quad \text{a.s. on } \{Z_0 > 0\}. \end{aligned}$$

Consequently, $\Xi = -\infty$ a.s. on the $\{Z_0 > 0\}$ of positive probability. On the other hand, $\frac{\varphi(\cdot,1)-\varphi(\cdot,0)}{1} \in L^1$, hence the monotone convergence theorem shows that

$$\frac{\mathcal{H}_V(\lambda Y_0|\lambda Z_1 + (1-\lambda)Z_0) - \mathcal{H}_V(0|\mathcal{P})}{\lambda} = \mathbb{E}\left[\frac{\varphi(\cdot,\lambda)-\varphi(\cdot,0)}{\lambda}\right] \downarrow -\infty.$$

Thus $\mathcal{H}_V(\lambda_0 Y_1|\mathcal{P}) \le \mathcal{H}_V(\lambda Y_0|\lambda_0 Z_1 + (1-\lambda_0)Z_0) < \mathcal{H}_V(0|\mathcal{P})$ for a small $\lambda_0 > 0$. $\square$

4.2. Duality and Indifference Valuation

The next one is the basic duality result.

Proposition 4.2. *Suppose (4.1), (4.2), (4.3) and (4.5). Then*

$$\sup_{X\in\mathcal{C}} u(X) = \min_{Y\in\mathcal{C}_V^\circ\setminus\{0\}} \mathcal{H}_V(Y|\mathcal{P}) = \inf_{Y\in\mathcal{C}_V^{\circ,e}} \mathcal{H}_V(Y|\mathcal{P}). \tag{4.7}$$

(The infimum over $\mathcal{C}_V^{\circ,e}$ is not generally attained). If we suppose additionally

$$\sup_{X\in\mathcal{C}} E_P[U(\cdot,X)] < E_P[U(\cdot,\infty)], \quad \forall P \in \mathcal{P}^e, \tag{4.8}$$

then u is well-defined with values in $[-\infty,\infty)$ on the convex cone

$$\bar{\mathcal{C}}^V := \{X \in L^0 : XY \in L^1 \text{ and } \mathbb{E}[XY] \le 0,\ \forall Y \in \mathcal{C}^\circ\}, \tag{4.9}$$

and we have

$$\sup_{X\in\bar{\mathcal{C}}^V} u(X) = \sup_{X\in\mathcal{C}} u(X) = \min_{Y\in\mathcal{C}^\circ\setminus\{0\}} \mathcal{H}_V(Y|\mathcal{P}). \tag{4.10}$$

Proof. Under the assumptions of the proposition, u is concave, finite and $\tau(L^\infty, L^1)$-continuous on the whole L^∞ by Corollary 3.6. Then a version of the Fenchel duality theorem ([29], Th. 1) applied to the dual pair $\langle L^\infty, L^1\rangle$ and the pair of concave/convex functions $(u, \delta_\mathcal{C})$ shows that

$$\sup_{X\in\mathcal{C}} u(X) = \sup_{X\in L^\infty} (u(X) - \delta_\mathcal{C}(X)) = \min_{Y\in L^1} \left(\delta_\mathcal{C}^*(Y) - u_*(Y)\right),$$

where $u_*(Y) := \inf_{X\in L^\infty}(\mathbb{E}[XY] - u(X)) = -\mathcal{I}_{f_U}^*(Y)$ is the *concave* conjugate, while $\mathcal{I}_{f_U}^*(Y) = \mathcal{H}_V(Y|\mathcal{P})$ on L^1 by Corollary 3.3, and $\delta_\mathcal{C}^*(Y) = \sup_{X\in\mathcal{C}} \mathbb{E}[XY] = \delta_{\mathcal{C}^\circ}(Y)$ since $\mathcal{C}$ is a cone. Summing up, we see that $\sup_{X\in\mathcal{C}} u(X) = \min_{Y\in\mathcal{C}^\circ} \mathcal{H}_V(Y|\mathcal{P}) = \min_{Y\in\mathcal{C}_V^\circ} \mathcal{H}_V(Y|\mathcal{P})$, and Lemma 4.1 shows that 0 must not be a minimizer, which establishes the first equality in (4.7). Next, let $\hat{Y} \in \mathcal{C}_V^\circ$ be a minimizer and pick a $Y \in \mathcal{C}_V^{\circ,e}$. Then $\lambda Y + (1-\lambda)\hat{Y} \in \mathcal{C}_V^{\circ,e}$ for $\lambda \in (0,1]$ and the function $\lambda \mapsto \mathcal{H}_V(\lambda Y + (1-\lambda)\hat{Y}|\mathcal{P})$ is upper semicontinuous on $[0,1]$ as a finite convex function on the line. Hence

$$\begin{aligned}\inf_{Y'\in\mathcal{C}_V^{\circ,e}} \mathcal{H}_V(Y'|\mathcal{P}) &\le \inf_{\lambda\in(0,1]} \mathcal{H}_V(\lambda Y + (1-\lambda)\hat{Y}|\mathcal{P}) \\ &\le \limsup_{\lambda\downarrow 0} \mathcal{H}_V(\lambda Y + (1-\lambda)\hat{Y}|\mathcal{P}) \le \mathcal{H}_V(\hat{Y}|\mathcal{P}).\end{aligned}$$

This proves the second equality in (4.7).

The assumption (4.8) implies that

$$\forall P \in \mathcal{P}^e,\ \exists Y \in \mathcal{C}^\circ \text{ such that } \tilde{V}(\cdot, Y, dP/d\mathbb{P}) \in L^1. \tag{4.11}$$

To see this, apply the first part of this proposition replacing $\mathcal{P}$ by $\{P\}$ (note that the Fenchel duality is valid without the finiteness of the supremum). Then (4.11) together with Young's inequality $zU(\cdot,x) \le \tilde{V}(\cdot,y,z) + xy$ (see (2.13)) shows that $U(\cdot,X)^+ \in L^1(P)$ for all $P \in \mathcal{P}^e$ whenever $X \in \bar{\mathcal{C}}^V$, and this easily extends to all $P \in \mathcal{P}$. So, u is well-defined on $\bar{\mathcal{C}}^V$ with values in $[-\infty,\infty)$, and $u(X) = \inf_{P\in\mathcal{P}} E_P[U(\cdot,X)] \le$

$\inf_{P\in\mathcal{P}}(\mathcal{H}_V(Y|P)+\mathbb{E}[XY])\leq\mathcal{H}_V(Y|\mathcal{P})$ for any $X\in\overline{\mathcal{C}}^V$ and $Y\in\mathcal{C}^\circ$ again by Young's inequality. Consequently,

$$\sup_{X\in\mathcal{C}}u(X)\leq\sup_{X\in\overline{\mathcal{C}}^V}u(X)\leq\inf_{Y\in\mathcal{C}^\circ}\mathcal{H}_V(Y|\mathcal{P})=\sup_{X\in\mathcal{C}}u(X).$$

This establishes (4.10). □

In addition to the random utility function U verifying (4.1) – (4.3), we now suppose we are given a random variable $B\in L^0$ which is considered as the payoff of a claim (or an option). Then consider the new random utility function

$$U_B(\omega,x)=U(\omega,x+B(\omega)). \tag{4.12}$$

Writing $f_U(\cdot,x)=-U(\cdot,-x)$, we observe $f_{U_B}(\cdot,x)=f_U(\cdot,x-B)$ (or $U_B(\cdot,x)=-f_U(\cdot,-x-B)$) in the notation of Example 3.11, and assumption (3.16) there reads as

$$\exists\varepsilon>0\text{ such that }U(-(1+\varepsilon)B^-)\in L^1_u(\mathcal{P})\text{ and }U(-\varepsilon B^+)\in L^1(\mathcal{P}) \tag{4.13}$$

With this assumption, the new random utility function U_B still satisfies (4.1) and (4.2), while (4.3) is invariant under translation. In view of Example 3.11 (with changes of signs) we have also that $\mathcal{H}_{V_B}(Y|\mathcal{P})<\infty\Leftrightarrow\mathcal{H}_V(Y|\mathcal{P})<\infty\Rightarrow YB\in L^1$, and

$$\mathcal{H}_{V_B}(Y|\mathcal{P})=\begin{cases}\mathcal{H}_V(Y|\mathcal{P})+\mathbb{E}[YB] & \text{if }\mathcal{H}_V(Y|\mathcal{P})<\infty,\\ +\infty & \text{otherwise,}\end{cases} \tag{4.14}$$

where $V_B(\cdot,y)=\sup_x(U_B(\cdot,x)-xy)=V(\cdot,y)+yB$. In particular, assumption (4.8) imposed on U (not U_B) still implies through (4.11) that $U_B(\cdot,X)^+\in L^1(P)$ for all $P\in\mathcal{P}$, thus u_B is also well-defined on $\overline{\mathcal{C}}^V$ as long as the original U satisfies (4.8).

We introduce a couple of more sets:

$$\begin{aligned}\mathcal{M}_V&:=\{Y\in\mathcal{C}^\circ:\mathbb{E}[Y]=1,\ \mathcal{H}_V(yY|\mathcal{P})<\infty,\ \exists y>0\},\\ \mathcal{M}_V^e&:=\{Y\in\mathcal{M}_V:\ Y>0\text{ a.s.}\}.\end{aligned}$$

Each $Y\in\mathcal{C}_V^\circ\setminus\{0\}$ is expressed as $Y=yY'$ with $y>0$ and $Y'\in\mathcal{M}_V$ (just put $y=\mathbb{E}[Y]$ and $Y'=\frac{1}{y}Y$), while for any $y>0$ and $Y'\in\mathcal{M}_V$, $yY'\in\mathcal{C}_V^\circ\setminus\{0\}$ if and only if $\mathcal{H}_V(yY'|\mathcal{P})<\infty$, and then $Y'B\in L^1$ and $\mathcal{H}_{V_B}(yY'|\mathcal{P})=\mathcal{H}_V(yY'|\mathcal{P})+yE[Y'B]$. Note also that each element of $\mathcal{M}_V$ is a probability measure (with the identification of a measure Q and its density $dQ/d\mathbb{P}$), and $\mathcal{M}_V^e$ is the set of $Q\in\mathcal{M}_V$ which are equivalent to $\mathbb{P}$. Finally, observe that addition of a constant to B does not affect (4.13) (just change the constant $\varepsilon>0$), and $u_B(x+X)=u_{x+B}(X)$. These arguments prove the following:

Corollary 4.3. *Suppose (4.1), (4.2), (4.3), (4.5) and (4.13). Then it holds that*

$$\begin{aligned}\sup_{X\in\mathcal{C}}u_B(x+X)&=\inf_{Y\in\mathcal{C}_V^{\circ,e}}(\mathcal{H}_V(Y|\mathcal{P})+\mathbb{E}[Y(x+B)])\\ &=\inf_{y>0}\inf_{Q\in\mathcal{M}_V^e}\big(\mathcal{H}_V(yQ|\mathcal{P})+yE_Q[B]+xy\big)\end{aligned} \tag{4.15}$$

The infimums are attained if we remove the superscript "^{e}". Under (4.8), we have also

$$\sup_{X\in\overline{\mathcal{C}}^V}u_B(x+X)=\sup_{X\in\overline{\mathcal{C}}^V}u_B(x+X).$$

Note that assumptions (4.5) and (4.8) as well as sets $\overline{\mathcal{C}}^V$, $\mathcal{C}_V^\circ$, $\mathcal{C}_V^{\circ,e}$, $\mathcal{M}_V$ and $\mathcal{M}_V^e$ do not depend on B (and x). Also, when U is non-random, assumptions (4.1) and (4.2) just say that U is finite on the whole $\mathbb{R}$.

Remark 4.4. When U is non-random and *bounded above* (i.e., $\sup_x U(x) < \infty$), the same duality result as (4.15) is obtained in [26] and see also [11] for the case without B. When U is (non-random) and finite only on $\mathbb{R}_+$, [40] obtained a similar result with $B = 0$, extended by [41] to the case with bounded B, and [39] consider the case of more general form of robust utility in a *penalized form*. The common approach to the duality in those papers is roughly to interchange the "$\sup_{X\in\mathcal{C}}$" and "$\inf_{P\in\mathcal{P}}$" by a minimax argument, and then to apply a suitable duality available for each *fixed* P in the particular setups ([2], [3], [20, 21], [37], [5]). Instead, we directly analyzed the robust utility functional u by means of our robust version of Rockafellar theorem, which gave us the precise information on the regularity of u in terms of the random utility function U as well as a criterion in terms of the integrability of B for u_B to retain the regularity of plain robust utility functional u. Consequently, we could obtain the duality in a considerably more general setting with a much simpler proof as long as the utility function is finite on the whole $\mathbb{R}$.

Even if we restrict ourselves to the classical case with $\mathcal{P} = \{\mathbb{P}\}$, our duality result *without singular term* is still quite general. In this case (with non-random U), [3] obtained a similar duality which is stated in an Orlicz space framework (more general than L^∞), while their duality generally has a *singular term* and the condition for removing the singular term is that $U(\varepsilon B^-) \in L^1$ *for all* $\varepsilon > 0$ in our notation (compare to the first half of (4.13)). See also [25] for a complement to [3] regarding this point in the classical case.

Finally, we consider a robust version of the *buyer's utility indifference price* of a claim B which is defined to be the real number

$$\pi(B) := \sup\{x \in \mathbb{R} : \sup_{X\in\mathcal{C}} u_B(-x + X) \geq \sup_{X\in\mathcal{C}} u(X)\}.$$

The interpretation of this quantity as a price is as follows. Consider the two alternative strategies: One is to buy the claim at now at the price x which yields the payoff B at the maturity, so the terminal *net* gain from the investment $X \in \mathcal{C}$ is $-x + X + B$, and the maximum possible robust utility in this case is $\sup_{X\in\mathcal{C}} u_B(x + X)$. On the other hand, if one does not buy the claim and just invest in $X \in \mathcal{C}$, then the $\sup_{X\in\mathcal{C}} u(X)$ is the maximum possible robust utility. For this investor, it is better to buy the claim as long as $\sup_{X\in\mathcal{C}} u_B(-x + X) > \sup_{X\in\mathcal{C}} u(X)$, while it is not if the converse (strict) inequality holds. In this sense $\pi(B)$ is the maximum acceptable price of B for the investor.

From (4.15), a straightforward computation yields the following:

Corollary 4.5 (cf. [36], [3], [22] when $\mathcal{P} = \{\mathbb{P}\}$). *Under (4.1–4.3), (4.5) and (4.13)*

$$\pi(B) = \inf_{Q\in\mathcal{M}_V^e} \left(E_Q[B] - \gamma(Q)\right),$$

where $\gamma(Q) = \inf_{y>0} \frac{1}{y}\left(\mathcal{H}_V(yQ|\mathcal{P}) - \inf_{y'>0, Q'\in\mathcal{M}_V} \mathcal{H}_V(y'Q'|\mathcal{P})\right)$.

4.3. A Typical Setup in Finance

Here we briefly review a typical financial setup and explain how it is reduced to the abstract framework discussed above. Let S be an $\mathbb{R}^d$-valued *locally bounded* semimartingale on a *filtered* probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, which describes the evolution of discounted prices of underlying assets (see [16] for the notation of stochastic calculus). If we interpret $\mathbb{P}$ as the reference model, the change of probability from $\mathbb{P}$ to another $P \ll \mathbb{P}$ typically corresponds to the change of the "drift" of S, so $\mathcal{P}$ corresponds to the "confidence region" of the drift (see [14] for concrete examples). A *self-financing* strategy and its *discounted* gain are modeled respectively by a d-dimensional predictable process θ and its stochastic integral $\theta \cdot S = \int_0^{\cdot} \theta dS$ w.r.t. S. We usually consider only those θ which are "admissible" in a suitable sense. A common choice of admissible class is:

$$\Theta_{bb} := \{\theta : (S, \mathbb{P})\text{-integrable}, \ \theta_0 = 0, \ \theta \cdot S \text{ is uniformly bounded below}\},$$

Then a typical choice of $\mathcal{C}$ is the set of *claims superreplicable with zero initial costs*:

$$\mathcal{C}_{bb} := \{X \in L^\infty : X \leq \theta \cdot S_T \text{ for some } \theta \in \Theta_{bb}\}.$$

Then the polar $\mathcal{C}_{bb}^\circ$ consists of positive multiples of probability measures $Q \ll \mathbb{P}$ under which S is a local martingale (*local martingale measures*) and the assumption $\mathcal{C}_{bb,V}^{\circ,e} \neq \emptyset$ reads as the existence of a local martingale measure Q_0 *equivalent to* $\mathbb{P}$ with finite robust divergence $\mathcal{H}_V(yQ_0|\mathcal{P}) < \infty$ for some $y > 0$, which implies a robust version of *no-arbitrage* (NA): if $\theta \in \Theta_{bb}$ with $P(\theta \cdot S_T \geq 0) = 1$ for all $P \in \mathcal{P}$, then $P(\theta \cdot S_T > 0) = 0$ for all $P \in \mathcal{P}$. However, each $P \in \mathcal{P}$ considered as a model may admit an arbitrage.

It would be more natural to consider the maximization of u over the set of admissible wealths $\mathcal{K}_{bb} := \{\theta \cdot S_T : \theta \in \Theta_{bb}\} (\not\subset L^\infty)$, rather than $\mathcal{C}_{bb}$. In fact, the formulation with $\mathcal{C}_{bb}$ is a clever reduction of the former. By definition, we see easily that $\sup_{X \in \mathcal{K}_{bb}} u(X) \geq \sup_{X \in \mathcal{C}_{bb}} u(X)$. On the other hand, if $\theta \in \Theta_{bb}$, then $\theta \cdot S$ is a supermartingale under all local martingale measures, hence $\theta \cdot S_T \in L^1(Q)$ and $E_Q[\theta \cdot S_T] \leq 0$ for all $Q \in \mathcal{M}_V$. Consequently, $\mathcal{K}_{bb} \subset \overline{\mathcal{C}}_{bb}^V$. The second part of Proposition 4.2 then tells us that

$$\sup_{X \in \mathcal{C}_{bb}} u(X) = \sup_{X \in \mathcal{K}_{bb}} u(X) = \sup_{X \in \overline{\mathcal{C}}_{bb}^V} u(X).$$

Thus the maximization over $\mathcal{C}_{bb}$ and that over $\mathcal{K}_{bb}$ are quantitatively equivalent, and we can further enlarge the domain to the "closure" $\overline{\mathcal{C}}_{bb}^V$ of $\mathcal{C}_{bb}$ without changing the optimal value. The last part is important since neither $\sup_{X \in \mathcal{C}_{bb}} u(X)$ nor $\sup_{X \in \mathcal{K}_{bb}} u(X)$ are attained (excepting trivial cases; see e.g. [38]). On the other hand, with a few more regularity assumptions on U, the supremum over $\overline{\mathcal{C}}_{bb}^V$ is indeed attained and the maximizer is explicitly obtained in terms of the dual optimizer. Moreover under a sort of "time-consistency" of $\mathcal{P}$, the optimal $\hat{X} \in \overline{\mathcal{C}}_{bb}^V$ admits a stochastic integral representation $\hat{\theta} \cdot S_T$ where $\hat{\theta} \cdot S$ need not be bounded below but it is a *supermartingale* under all $Q \in \mathcal{M}_V$. See [27] for detail where a result of an earlier version of this paper was used.

In the last part, the convex duality technique collaborates with another duality technique from the theory of martingales, where the key is (in quite rough terms) the duality between the set of all stochastic integrals w.r.t. a fixed semimartingale S and the set of all probability measures which makes S a local martingale. Thus it is essential there that the *discounted gains* are expressed as stochastic integrals. In contrast, nominal (or non-discounted) gains are not of this form. A merit of our framework where the utility function itself is allowed to be random is that one can still work with *discounted gains* even if one is interested in the maximization over the nominal gains. Indeed, the nominal gain is by definition the discounted gain divided by the discount factor. Then we can embed the discount factor into the utility function and estimate the random utility thus obtained as in Example 3.10.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the financial support of the Center for Advanced Research in Finance (CARF) at the Graduate School of Economics of the University of Tokyo.

APPENDIX A. OMITTED PROOFS

A.1. ON EXAMPLE 3.2

Here we give some computational detail of Example 3.2 omitted in the main text. Let the probability space be $(\mathbb{N}, 2^{\mathbb{N}}, \mathbb{P})$ with $\mathbb{P}(\{n\}) = 2^{-n}$, and $\mathcal{P} = \overline{\mathrm{conv}}(P_n, n \in \mathbb{N})$ (closed convex hull) of the sequence $(P_n)_n$ given by $P_1(\{1\}) = 1$, $P_n(\{1\}) = 1 - 1/n$ and $P_n(\{n\}) = 1/n$. To see the weak compactness of $\mathcal{P}$, it suffices to note that $\sup_n P_n(\{k, k+1, k+2, ...\}) = \sup\{1/n : n \geq k\} = 1/k \to 0$ as $k \to \infty$. Also, $\sup_{P \in \mathcal{P}} E_P[X] = \sup_n E_{P_n}[X]$ if $X \geq 0$ since $P \mapsto E_P[X \wedge N]$ is continuous for any $N \in \mathbb{N}$, hence

$$\sup_{P \in \mathcal{P}} E_P[X] = \sup_N \sup_{P \in \mathcal{P}} E_P[X \wedge N] = \sup_N \sup_{P \in \mathrm{conv}(P_n; n \in \mathbb{N})} E_P[X \wedge N],$$

while if P is of the form $P = \alpha_1 P_{n_1} + \cdots \alpha_l P_{n_l}$, then $E_P[X \wedge N] = \alpha_1 E_{P_{n_1}}[X \wedge N] + \cdots + \alpha_l E_{P_{n_l}}[X \wedge N] \leq \max_{1 \leq i \leq l} E_{P_{n_i}}[X \wedge N] \leq \sup_n E_{P_n}[X \wedge N]$.

Lemma A.1. *Let f be given by (3.4) in Example 3.2. Then we have*

$$\lim_{N \to \infty} \sup_n E_{P_n}[f(\cdot, X) \mathbb{1}_{\{f(\cdot, X) \geq N\}}] = \limsup_n X(n)^+ e^{X(n)^+}. \tag{A.1}$$

Proof. Let $h(x) = x^+ e^x$. For any $X \in L^\infty = l^\infty$ and $N \in \mathbb{N}$,

$$E_{P_n}[f(\cdot, X) \mathbb{1}_{\{f(\cdot, X) \geq N\}}] = \left(1 - \frac{1}{n}\right) h(X(1)) \mathbb{1}_{\{h(X(1)) \geq N\}} + h(X(n)) \mathbb{1}_{\{n h(X(n)) \geq N\}}$$

Let $\alpha := \limsup_n h(X(n))$. Then for any $\varepsilon > 0$ and $N \in \mathbb{N}$, there exists some $n_N^\varepsilon > N/(\alpha - \varepsilon)$ such that $h(X(n_N^\varepsilon)) > \alpha - \varepsilon$. In particular, $n_N^\varepsilon h(X(n_N^\varepsilon)) > N$, hence

$$\sup_n E_{P_n}[f(\cdot, X) \mathbb{1}_{\{f(\cdot, X) \geq N\}}] \geq \sup_n h(X(n)) \mathbb{1}_{\{n h(X(n)) \geq N\}} \geq h(X(n_N^\varepsilon)) > \alpha - \varepsilon.$$

Thus we have "$\geq$" in (A.1). On the other hand, putting $\beta := h(\|X\|_\infty)$,

$$E_{P_n}[f(\cdot, X) \mathbb{1}_{\{f(\cdot, X) \geq N\}}] \leq h(X(1)) \mathbb{1}_{\{h(X(1)) \geq N\}} + h(X(n)) \mathbb{1}_{\{n\beta \geq N\}}.$$

If $\beta = 0$ ($\Leftrightarrow \|X\|_\infty = 0$), both sides of (A.1) are zero, while if $\beta > 0$, we have

$$\lim_{N \to \infty} \sup_n E_{P_n}[f(\cdot, X) \mathbb{1}_{\{f(\cdot, X) \geq N\}}] \leq \lim_{N \to \infty} \sup_{n \geq N/\beta} h(X(n)) = \limsup_n h(X(n))$$

as claimed. □

Since $h(x) = x^+ e^x$ is increasing, continuous and $h(0) = 0$, $\limsup_n h(X(n)) = 0$ if and only if $\limsup_n X(n) = 0$, and $h(X(n)) \leq h(\|X\|_\infty)$, so consequently

$$\mathcal{D}_f = \{X \in L^\infty : \limsup_n X(n) \leq 0\} \quad \text{and} \quad \mathrm{dom}(\mathcal{I}_f) = L^\infty. \tag{A.2}$$

On the other hand,

Lemma A.2. *The conjugate $\mathcal{I}_f^*$ is explicitly computed on ba^s as:*

$$(\mathcal{I}_f)^*(\nu) = \begin{cases} \sup_{x \in \mathbb{R}} x(\|\nu_s\| - e^x) & \text{if } \nu \in ba_+^s, \\ +\infty & \text{if } \nu \in ba^s \setminus ba_+^s. \end{cases} \tag{A.3}$$

Proof. Since $\mathcal{I}_f$ is monotone increasing, $\mathcal{I}_f^*(\nu)$ is finite only if $\nu \in ba_+$. For $\nu \in ba_+^s$, Since $E_{P_n}[f(\cdot, X)] = \left(1 - \frac{1}{n}\right) h(X(1)) + h(X(n))$ where $h(x) = x^+ e^x$, we have

$$h(X(n)) \leq E_{P_n}[f(\cdot, X)] \leq h(X(1)) + h(X(n)). \tag{A.4}$$

From the first inequality, we have

$$\nu_s(X) - \mathcal{I}_f(X) \leq \|\nu_s\| \limsup_n X(n) - \sup_n h(X(n)) \leq \|X^+\|_\infty (\|\nu_s\| - e^{\|X^+\|_\infty}).$$

This shows $(\mathcal{I}_f)^*(\nu_s) \le \sup_{X\in L^\infty} \|X^+\|_\infty(\|\nu_s\| - e^{\|X^+\|_\infty}) = \sup_{x\ge 0} x(\|\nu_s\| - e^x)$.

On the other hand, the second inequality in (A.4) shows that

$$(\mathcal{I}_f)^*(\nu_s) \ge \sup_{X\in L^\infty} (\nu_s(X) - h(X(1)) - h(\|X\|_\infty))$$
$$\ge \sup_{X\in L^\infty, X(1)=0} (\nu_s(X) - \|X\|_\infty e^{\|X\|_\infty}) \ge \sup_{x\ge 0}(x\nu(1) - xe^x).$$

Since $\nu(1) = \|\nu\|$ if $\nu \in ba^s_+$, this implies $(\mathcal{I}_f)^*(\nu_s) = \sup_{x\ge 0} x(\|\nu_s\| - e^x)$. □

A.2. ON EXAMPLES 3.10 AND 3.11

Lemma A.3. *Let f be given as (3.11) with $g : \mathbb{R} \to \mathbb{R}$ being (deterministic and) convex and W being a strictly positive random variable. Then (3.12) implies that $\mathbb{R} \subset \mathcal{D}_f$.*

Proof. Note that $|Wx| = \frac{\delta}{2}|W(2x/\delta)| \le \frac{\delta}{2}\left(W^p + \frac{2^{q-1}}{\delta^{q-1}}|x|^q\right) = \frac{1}{2}\left(\delta W^p + \frac{2^{q-1}}{\delta^{q-2}}|x|^q\right)$ where $\frac{1}{p} + \frac{1}{q} = 1$. Then applying the quasi-convexity of g ($\Leftarrow$ convexity) twice, we have

$$g(Wx) \le g(-\delta W^p) \vee g(\delta W^p) \vee g\left(-\frac{2^{q-1}}{\delta^{q-2}}|x|\right) \vee g\left(\frac{2^{q-1}}{\delta^{q-2}}|x|\right)$$

The last two components are finite constants for each $x \in \mathbb{R}$, hence do not matter, and we see that (3.12) implies through this inequality that $g(xW)^+ \in L^1_u(\mathcal{P})$ for each $x \in \mathbb{R}$. □

Proof of (3.14) and (3.15). By the convexity of f and $f_B(\cdot, x) = f(\cdot, x + B)$ (definition),

$$f_B(\cdot, x) = f\left(\cdot, \frac{\varepsilon}{1+\varepsilon}\frac{1+\varepsilon}{\varepsilon}x + \frac{1+\varepsilon}{1+\varepsilon}B\right)$$
$$\le \frac{\varepsilon}{1+\varepsilon} f\left(\cdot, \frac{1+\varepsilon}{\varepsilon}x\right) + \frac{1}{1+\varepsilon} f(\cdot, (1+\varepsilon)B).$$

This is the upper bound in (3.14). Taking the conjugate, we have also the lower bound in (3.15). Recall that $f^*_B(\cdot, y) = f^*(\cdot, y) - yB$. Then noting $-\varepsilon yB \le f(\cdot, -\varepsilon B) + f^*(\cdot, y)$,

$$f^*_B(\cdot, y) \le f^*(\cdot, y) + \frac{1}{\varepsilon} f(\cdot, -\varepsilon B) + \frac{1}{\varepsilon} f^*(\cdot, y) = \frac{1+\varepsilon}{\varepsilon} f^*(\cdot, y) + \frac{1}{\varepsilon} f(\cdot, -\varepsilon B).$$

This is the upper bound in (3.15) which implies also the lower bound in (3.14) by taking the conjugate. □

REFERENCES

[1] Aliprantis, C. D. and K. C. Border (2006): Infinite dimensional analysis: A hitchhiker's guide. 3rd edn. Springer, Berlin.

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

[3] Biagini, S., M. Frittelli and M. Grasselli (2011): Indifference price with general semimartingales. *Math. Finance* **21**, 423–446.

[4] Castaing, C. and M. Valadier (1977): Convex analysis and measurable multifunctions, *Lecture Notes in Math.*, vol. 580. Springer-Verlag, Berlin.

[5] Cvitanić, J., W. Schachermayer and H. Wang (2001): Utility maximization in incomplete markets with random endowment. *Finance Stoch.* **5**, 259–272.

[6] Delbaen, F. (2009): Differentiability properties of utility functions. In: Optimality and risk—modern trends in mathematical finance. Springer, Berlin, pp. 39–48.

[7] Dellacherie, C. and P.-A. Meyer (1978): Probabilities and potential, *North-Holland Mathematics Studies*, vol. 29. North-Holland Publishing Co., Amsterdam.

[8] Denis, L., M. Hu and S. Peng (2011): Function spaces and capacity related to a sublinear expectation: Application to G-brownian motion paths. *Potential Analysis* **34**, 139–161.

[9] Dunford, N. and J. T. Schwartz (1988): Linear operators. Part I. General theory. Wiley Classics Library. John Wiley & Sons Inc., New York. With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[10] Ekeland, I. and R. Temam (1976): Convex analysis and variational problems. North-Holland Publishing Co., Amsterdam. Translated from the French, Studies in Mathematics and its Applications, Vol. 1.

[11] Föllmer, H. and A. Gundel (2006): Robust projections in the class of martingale measures. *Illinois J. Math.* **50**, 439–472 (electronic).

[12] Föllmer, H., A. Schied and S. Weber (2009): Robust preferences and robust portfolio choice. In: Bensoussan, A., Q. Zhang and P. G. Ciarlet (eds.), Mathematical Modelling and Numerical Methods in Finance, *Handbook of Numerical Analysis*, vol. 15. North-Holland, pp. 29–88.

[13] Grothendieck, A. (1973): Topological vector spaces. Gordon and Breach Science Publishers, New York. Translated from the French by Orlando Chaljub, Notes on Mathematics and its Applications.

[14] Hernández-Hernández, D. and A. Schied (2006): Robust utility maximization in a stochastic factor model. *Statist. Decisions* **24**, 109–125.

[15] Hewitt, E. and K. Stromberg (1975): Real and abstract analysis, *Graduate Texts in Mathematics*, vol. 25. Springer-Verlag, New York. A modern treatment of the theory of functions of a real variable, Third printing.

[16] Jacod, J. and A. N. Shiryaev (2003): Limit Theorems for Stochastic Processes, *Grundlehren Math. Wiss.*, vol. 288. 2nd edn. Springer-Verlag.

[17] Jouini, E., W. Schachermayer and N. Touzi (2006): Law invariant risk measures have the fatou property. *Advances in Mathematical Economics* **9**, 49–71.

[18] Kozek, A. (1976/77): Orlicz spaces of functions with values in Banach spaces. *Comment. Math. Prace Mat.* **19**, 259–288.

[19] Kozek, A. (1980): Convex integral functionals on Orlicz spaces. *Comment. Math. Prace Mat.* **21**, 109–135.

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

[21] Kramkov, D. and W. Schachermayer (2003): Necessary and sufficient conditions in the problem of optimal investment in incomplete markets. *Ann. Appl. Probab.* **13**, 1504–1516.

[22] Mania, M. and M. Schweizer (2005): Dynamic exponential utility indifference valuation. *Ann. Appl. Probab.* **15**, 2113–2143.

[23] Moreau, J.-J. (1964): Sur la fonction polaire d'une fonction semi-continue supérieurement. *C. R. Acad. Sci. Paris* **258**, 1128–1130.

[24] Orihuela, J. and M. Ruiz Galán (2012): A coercive James's weak compactness theorem and nonlinear variational problems. *Nonlinear Anal., Theory Methods Appl., Ser. A, Theory Methods* **75**, 598–611.

[25] Owari, K. (2011): A note on utility maximization with unbounded random endowment. *Asia-Pacific Financial Markets* **18**, 89–103.

[26] Owari, K. (2011): Robust utility maximization with unbounded random endowment. *Adv. Math. Econ.* **14**, 147–181.

[27] Owari, K. (2012): On admissible strategies in robust utility maximization. *Math. Financ. Econ.* **6**, 77–92.

[28] Pennanen, T. (2011): Convex duality in stochastic optimization and mathematical finance. *Math. Oper. Res.* **36**, 340–362.

[29] Rockafellar, R. T. (1966): Extension of Fenchel's duality theorem for convex functions. *Duke Math. J.* **33**, 81–89.

[30] Rockafellar, R. T. (1968): Integrals which are convex functionals. *Pacific J. Math.* **24**, 525–539.

[31] Rockafellar, R. T. (1971): Convex integral functionals and duality. In: Contributions to nonlinear functional analysis (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1971). Academic Press, New York, pp. 215–236.

[32] Rockafellar, R. T. (1971): Integrals which are convex functionals. II. *Pacific J. Math.* **39**, 439–469.

[33] Rockafellar, R. T. (1974): Conjugate duality and optimization. No. 16 in CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, Pa.

[34] Rockafellar, R. T. (1976): Integral functionals, normal integrands and measurable selections. In: Nonlinear operators and the calculus of variations (Summer School, Univ. Libre Bruxelles, Brussels, 1975), *Lecture Notes in Math.*, vol. 543. Springer, Berlin, pp. 157–207.

[35] Rockafellar, R. T. and R. J.-B. Wets (1998): Variational analysis, *Grundlehren der Mathematischen Wissenschaften*, vol. 317. Springer-Verlag, Berlin.

[36] Rouge, R. and N. El Karoui (2000): Pricing via utility maximization and entropy. *Math. Finance* **10**, 259–276.

[37] Schachermayer, W. (2001): Optimal investment in incomplete markets when wealth may become negative. *Ann. Appl. Probab.* **11**, 694–734.

[38] Schachermayer, W. (2003): A super-martingale property of the optimal portfolio process. *Finance Stoch.* **7**, 433–456.

[39] Schied, A. (2007): Optimal investments for risk- and ambiguity-averse preferences: a duality approach. *Finance Stoch.* **11**, 107–129.

[40] Schied, A. and C.-T. Wu (2005): Duality theory for optimal investments under model uncertainty. *Statist. Decisions* **23**, 199–217.

[41] Wittmüss, W. (2008): Robust optimization of consumption with random endowment. *Stochastics* **80**, 459–475.

[42] Yosida, K. (1980): Functional analysis, *Grundlehren der Mathematischen Wissenschaften*, vol. 123. Sixth edn. Springer-Verlag, Berlin.