

A Robust Version of Convex Integral Functionals

Keita Owari

Graduate School of Economics, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, 113-0033 Tokyo, Japan
owari@e.u-tokyo.ac.jp

Received: June 25, 2014

Accepted: September 13, 2014

We here study the pointwise supremum of convex integral functionals of the form $\mathcal{I}_{f,\gamma}(\xi) = \sup_Q \left(\int_{\Omega} f(\omega, \xi(\omega)) Q(d\omega) - \gamma(Q) \right)$ on $L^{\infty}(\Omega, \mathcal{F}, \mathbb{P})$ where $f : \Omega \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ is a proper normal convex integrand, γ is a proper convex function on the set of probability measures absolutely continuous w.r.t. $\mathbb{P}$, and the supremum is taken over all such measures. We give a pair of upper and lower bounds for the conjugate of $\mathcal{I}_{f,\gamma}$ as direct sums of a common regular part and respective singular parts; they coincide when $\text{dom}(\gamma) = \{\mathbb{P}\}$ as Rockafellar's classical result, while both inequalities can generally be strict. We then investigate when the conjugate eliminates the singular measures, which a fortiori yields the equality in bounds, and its relation to other finer regularity properties of the original functional and of the conjugate.

Keywords: Convex integral functionals, duality, robust stochastic optimization, financial mathematics

2010 Mathematics Subject Classification: 46N10, 46E30, 49N15, 52A41, 91G80

1. Introduction

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $f : \Omega \times \mathbb{R} \rightarrow (-\infty, \infty]$ a measurable mapping with $f(\omega, \cdot)$ being proper, convex, lsc for a.e. ω . Then $\xi \mapsto I_f(\xi) := \int_{\Omega} f(\omega, \xi(\omega)) \mathbb{P}(d\omega)$ defines a convex functional on $L^{\infty} := L^{\infty}(\Omega, \mathcal{F}, \mathbb{P})$, called the *convex integral functional*. Among many others, R. T. Rockafellar obtained in [22] that under mild integrability assumptions on f , the conjugate $I_f^*(\nu) = \sup_{\xi \in L^{\infty}} (\nu(\xi) - I_f(\xi))$ of I_f is expressed as the direct sum of regular and singular parts (which we call the *Rockafellar theorem*):

$$I_f^*(\nu) = I_{f^*}(d\nu_r/d\mathbb{P}) + \sup_{\xi \in \text{dom}(I_f)} \nu_s(\xi), \quad \forall \nu \in (L^{\infty})^*, \quad (1)$$

where $f^*(\omega, y) := \sup_x (xy - f(\omega, x))$ and ν_r (resp. ν_s) denotes the regular (resp. singular) part of $\nu \in (L^{\infty})^*$ in the Yosida-Hewitt decomposition. 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 latter property

implies the weak-compactness of all the sublevel sets of $I_f^*|_{L^1}$ and the continuity of I_f for the Mackey topology $\tau(L^\infty, L^1)$ and so on (e.g. [23, Th. 3K]).

This paper is concerned with a *robust version* of integral functionals of the form

$$\mathcal{I}_{f,\gamma}(\xi) := \sup_{Q \in \mathcal{Q}} \left(\int_{\Omega} f(\omega, \xi(\omega)) Q(d\omega) - \gamma(Q) \right), \quad \xi \in L^\infty.$$

where $\mathcal{Q}$ is the set of all probability measures Q absolutely continuous w.r.t. $\mathbb{P}$, and γ is a convex penalty function on $\mathcal{Q}$. (see Section 2 for precise formulation). As the pointwise supremum of convex functions, $\mathcal{I}_{f,\gamma}$ is convex, lower semicontinuous if so is each $I_{f,Q}(\xi) := \int_{\Omega} f(\omega, \xi(\omega)) Q(d\omega)$, and is even norm-continuous as soon as it is finite everywhere. On the other hand, it is less obvious what the convex conjugate $\mathcal{I}_{f,\gamma}^*$ is, when singular measures are eliminated, and whether the latter property yields finer regularity properties of $\mathcal{I}_{f,\gamma}$ and $\mathcal{I}_{f,\gamma}^*$.

Our main result (Theorem 3.1) is a robust version of the Rockafellar theorem which consists of a pair of upper and lower bounds (instead of a single equality) for $\mathcal{I}_{f,\gamma}^*$ on $(L^\infty)^*$. Both bounds are of forms analogous to (1) with a common regular part, but a difference appears in the singular parts. They coincide in the classical case $I_f = \mathcal{I}_{f,\delta_{\{\mathbb{P}\}}}$, while the subsequent Example 3.3 shows that both inequalities can be strict and one may not hope for sharper bounds in general. The same example shows also that the everywhere finiteness of $\mathcal{I}_{f,\gamma}$ is not enough for the property that $\mathcal{I}_{f,\gamma}^*$ *eliminates the singular measures*, while the lower bound in Theorem 3.1 provides us a simple sufficient condition in a form of “uniform integrability”. In Theorem 3.4 and its corollaries, it is shown that given $\text{dom}(\mathcal{I}_{f,\gamma}) = L^\infty$ (plus a technical assumption), the condition is even necessary, and is equivalent to other finer properties of $\mathcal{I}_{f,\gamma}$ and $\mathcal{I}_{f,\gamma}^*$, including the weak compactness of sublevels of $\mathcal{I}_{f,\gamma}^*|_{L^1}$, the Mackey continuity of $\mathcal{I}_{f,\gamma}$ on L^∞ among others, which are guaranteed solely by the finiteness in the classical case.

Certain class of *robust optimization* problems are formulated as (or reduced to) the minimization of a robust integral functional $\mathcal{I}_{f,\gamma}$ over a convex set $\mathcal{C} \subset L^\infty$. Our initial motivation of this work lies in the Fenchel duality for this type of problems with $\langle L^\infty, L^1 \rangle$ pairing:

$$\inf_{\xi \in \mathcal{C}} \mathcal{I}_{f,\gamma}(\xi) = - \min_{\eta \in L^1} \left(\mathcal{I}_{f,\gamma}^*(\eta) + \sup_{\xi \in \mathcal{C}} \langle -\xi, \eta \rangle \right), \quad (2)$$

The Rockafellar-type result provides us the precise form of $\mathcal{I}_{f,\gamma}^*|_{L^1}$, hence of the *dual problem*, while the classical Fenchel duality theorem tells us that a sufficient condition for (2) is the $\tau(L^\infty, L^1)$ -continuity of $\mathcal{I}_{f,\gamma}$ at some point $\xi_0 \in \mathcal{C} \cap \text{dom}(\mathcal{I}_{f,\gamma})$, or in another view, the finiteness implies (via the norm continuity) the $\langle L^\infty, (L^\infty)^* \rangle$ -duality which reduces to the $\langle L^\infty, L^1 \rangle$ -duality if $\mathcal{I}_{f,\gamma}^*$ eliminates the singular elements of $(L^\infty)^*$.

Our motivating example of minimization of $\mathcal{I}_{f,\gamma}$ is the *robust utility maximization*

$$\text{maximize } u(\xi) := \inf_{Q \in \mathcal{Q}} (\mathbb{E}_Q[U(\cdot, \xi)] + \gamma(Q)) \quad \text{over a convex cone } \mathcal{C} \subset L^\infty \quad (3)$$

where $U : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ is a *random utility function*. In this context, each $Q \in \mathcal{Q}$ is considered as a model used to evaluate the quality of control ξ via the expected utility $\mathbb{E}_Q[U(\cdot, \xi)] = \int_\Omega U(\omega, \xi(\omega))Q(d\omega)$, but we are not sure which model is true; so we optimize the worst case among all models Q penalized by $\gamma(Q)$ according to the likelihood. Note that (3) is equivalent to minimize $\mathcal{I}_{f_U, \gamma}$ with $f_U(\omega, x) = -U(\omega, -x)$ over the cone $-\mathcal{C}$. The corresponding Fenchel duality in $\langle L^\infty, L^1 \rangle$ constitute a half of what we call the *martingale duality* in mathematical finance; given the $\langle L^\infty, L^1 \rangle$ -duality with a “good” cone $\mathcal{C}$ having the polar generated by so-called *local martingale measures*, the theory of (semi)martingales takes care of the other half.

2. Preliminaries

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 random variables defined on it. As usual, we do not distinguish a random variable and the class it generates, and a constant $c \in \mathbb{R}$ is regarded as the random variable $c\mathbb{1}_\Omega$. Here $\mathbb{1}_A$ denotes the indicator of a set A in measure theoretic sense while $\delta_A := \infty\mathbb{1}_{A^c}$ is the one in convex analysis. The $\mathbb{P}$ -expectation of $\xi \in L^0$ is denoted by $\mathbb{E}[\xi] := \int_\Omega \xi(\omega)\mathbb{P}(d\omega)$ and we write $L^p := L^p(\Omega, \mathcal{F}, \mathbb{P})$ and $\|\cdot\|_p := \|\cdot\|_{L^p}$ for each $1 \leq p \leq \infty$. Any probabilistic notation without reference to a measure is to be understood with respect to $\mathbb{P}$. Especially, “a.s.” means “ $\mathbb{P}$ -a.s.”, and *identification of random variables is always made by $\mathbb{P}$* . For other probability measures Q absolutely continuous with respect to $\mathbb{P}$ ($Q \ll \mathbb{P}$), we write $\mathbb{E}_Q[\cdot]$ for Q -expectation, $L^p(Q) := \{\xi \in L^0 : \mathbb{E}_Q[|\xi|^p] < \infty\}$ (which is a set of $\mathbb{P}$ -classes!) etc, explicitly indicating the measure that involves. We write $Q \sim P$ to mean Q and P are equivalent ($Q \ll P$ and $P \ll Q$).

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 $\mathbb{P}(A) = 0$ (see [12, pp. 354–357]). The bilinear form of (L^∞, ba) is given by the (Radon) integral $\nu(\xi) = \int_\Omega \xi d\nu$ which coincides with the usual integral when ν is σ -additive. We regard any σ -additive $\nu \in ba$ as an element of L^1 via the mapping $\nu \mapsto d\nu/d\mathbb{P}$ which is an isometry from the subspace of such ν 's onto L^1 . In particular, the set

$$\mathcal{Q} := \{Q : \text{probability measures on } (\Omega, \mathcal{F}) \text{ with } Q \ll \mathbb{P}\}$$

is regarded as $\{\eta \in L^1 : \eta \geq 0, \mathbb{E}[\eta] = 1\}$. 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 , and we denote by ba^s the totality of such $\nu \in ba$. 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. [7, Th.

III.7.8]), thus $ba = L^1 \oplus ba^s$ with the above identification. We denote by ba_+ the set of *positive* elements of ba and $ba_+^s := ba_+ \cap ba^s$ etc.

2.1. Convex Penalty Function and Associated Orlicz Spaces

We make the following assumption on the *penalty function* γ :

Assumption 2.1. $\gamma : \mathcal{Q} \rightarrow \overline{\mathbb{R}}_+$ is a $\sigma(L^1, L^\infty)$ -lsc proper convex function such that

$$\inf_{Q \in \mathcal{Q}} \gamma(Q) = 0; \quad (4)$$

$$\exists Q_0 \in \mathcal{Q} \text{ with } Q_0 \sim \mathbb{P} \text{ and } \gamma(Q_0) < \infty; \quad (5)$$

$$\begin{aligned} \{Q \in \mathcal{Q} : \gamma(Q) \leq 1\} \text{ is } \sigma(L^1, L^\infty)\text{-compact} \\ (\Leftrightarrow \text{closed and uniformly integrable}). \end{aligned} \quad (6)$$

We denote $\mathcal{Q}_\gamma := \{Q \in \mathcal{Q} : \gamma(Q) < \infty\}$, the effective domain of γ .

Remark 2.2. (4) and (5) are normalizing assumptions and the latter one is equivalent to saying that $\mathcal{Q}_\gamma \sim \mathbb{P}$, i.e., for each $A \in \mathcal{F}$, $\mathbb{P}(A) = 0$ iff $Q(A) = 0$ for every $Q \in \mathcal{Q}_\gamma$. Given the lower semicontinuity of γ , (6) implies that the infimum in (4) is attained, and (6) is equivalent to apparently stronger

$$\{Q \in \mathcal{Q} : \gamma(Q) \leq c\} \text{ is } \sigma(L^1, L^\infty)\text{-compact, } \forall c > 0.$$

Indeed, for any $c > 1$, pick a $Q_c \in \mathcal{Q}_\gamma$ with $\gamma(Q_c) < 1/c$ (by (4)), then $\gamma(Q) \leq c$ implies $\gamma\left(\frac{1}{1+c}Q + \frac{c}{1+c}Q_c\right) \leq \frac{1}{1+c} \cdot c + \frac{c}{1+c} \cdot \frac{1}{c} = 1$, thus (6) implies the weak compactness of $\left\{\frac{1}{1+c}Q + \frac{c}{1+c}Q_c : \gamma(Q) \leq c\right\}$, hence of $\{Q \in \mathcal{Q}_\gamma : \gamma(Q) \leq c\}$.

In general, any lower semicontinuous proper convex function γ on $\mathcal{Q}$ defines a function

$$\rho_\gamma(\xi) := \sup_{Q \in \mathcal{Q}_\gamma} (\mathbb{E}_Q[\xi] - \gamma(Q)) \quad \text{on } \left\{\xi \in L^0 : \xi^- \in \bigcap_{Q \in \mathcal{Q}_\gamma} L^1(Q)\right\} \supset L^\infty \cup L_+^0.$$

Regardless of Assumption 2.1, ρ_γ restricted to L^∞ is a $\sigma(L^\infty, L^1)$ -lsc finite-valued monotone convex function with $\rho_\gamma(\xi + c) = \rho_\gamma(\xi) + c$ if c is a constant, whose conjugate on L^1 is $\gamma(\eta)\mathbb{1}_{\mathcal{Q}}(\eta) + \infty\mathbb{1}_{L^1 \setminus \mathcal{Q}}(\eta)$. In financial mathematics, such a function is called a convex risk measure (up to a change of sign). Then (4) reads as $\rho_\gamma(0) = 0$, (5) as $\rho_\gamma(\varepsilon\mathbb{1}_A) = 0$ for some $\varepsilon > 0 \Rightarrow \mathbb{P}(A) = 0$, while (6) is equivalent to saying that ρ_γ has the *Lebesgue property on L^∞* (see [13, 5, 17]):

$$\sup_n \|\xi_n\|_\infty < \infty, \xi_n \rightarrow \xi \text{ a.s.} \Rightarrow \rho_\gamma(\xi) = \lim_n \rho_\gamma(\xi_n).$$

To a penalty function γ satisfying Assumption 2.1, we associate the gauge norm

$$\|\xi\|_{\rho_\gamma} = \inf \{\lambda > 0 : \rho_\gamma(|\xi|/\lambda) \leq 1\} = \sup_{Q \in \mathcal{Q}_\gamma} \frac{\mathbb{E}_Q[|\xi|]}{1 + \gamma(Q)}, \quad \xi \in L^0.$$

In view of (5), this is indeed a norm on the *Orlicz space*

$$L^{\rho_\gamma} = \{\xi \in L^0 : \exists \alpha > 0 \text{ with } \rho_\gamma(\alpha|\xi|) < \infty\} = \{\xi \in L^0 : \|\xi\|_{\rho_\gamma} < \infty\},$$

which is a *solid* subspace (lattice ideal) of L^0 (i.e., $\xi \in L^{\rho_\gamma}$ and $|\zeta| \leq |\xi|$ a.s. $\Rightarrow \zeta \in L^{\rho_\gamma}$), and $(L^{\rho_\gamma}, \|\cdot\|_{\rho_\gamma})$ is a Banach lattice. We consider the following subspaces of L^{ρ_γ} too:

$$\begin{aligned} M^{\rho_\gamma} &:= \{\xi \in L^0 : \forall \alpha > 0, \rho_\gamma(\alpha|\xi|) < \infty\}, \\ M_u^{\rho_\gamma} &:= \left\{ \xi \in L^0 : \forall \alpha > 0, \lim_N \rho_\gamma(\alpha|\xi| \mathbf{1}_{\{|\xi| > N\}}) = 0 \right\}. \end{aligned}$$

Both M^{ρ_γ} and $M_u^{\rho_\gamma}$ are solid as subspaces of L^0 . These Orlicz-type spaces are studied in [19] where γ corresponds to φ_0^* (more precisely $\varphi_0^* = \gamma \mathbf{1}_Q + \infty \mathbf{1}_{L^1 \setminus Q}$) and $\hat{\varphi} = \rho_\gamma$ in the notation of the current paper. In general under Assumption 2.1,

$$L^\infty \subset M_u^{\rho_\gamma} \subset M^{\rho_\gamma} \subset L^{\rho_\gamma} \subset \bigcap_{Q \in \mathcal{Q}_\gamma} L^1(Q),$$

while all inclusions can generally be strict ([19, Examples 3.6 and 3.7]). From the last inclusion, ρ_γ is well-defined on L^{ρ_γ} as a proper monotone convex function, and it is finite-valued on M^{ρ_γ} while not on L^{ρ_γ} (unless $M^{\rho_\gamma} = L^{\rho_\gamma}$), and $M_u^{\rho_\gamma}$ is the maximum solid subspace of L^0 to which ρ_γ retain the Lebesgue property (as the order-continuity; see [19, Theorem 3.5]), and is characterized by a uniform integrability property [19, Theorem 3.8]: for $\xi \in M^{\rho_\gamma}$,

$$\xi \in M_u^{\rho_\gamma} \Leftrightarrow \{\xi dQ/d\mathbb{P} : \gamma(Q) \leq c\} \text{ is uniformly integrable, } \forall c > 0. \quad (7)$$

Actually, by the same argument as in Remark 2.2, we have only to consider the case $c = 1$.

2.2. Robust Version of Integral Functionals

In the sequel, let $f : \Omega \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$ be a *proper normal convex integrand* on $\mathbb{R}$, i.e.,

$$\begin{aligned} f \text{ is } \mathcal{F} \otimes \mathcal{B}(\mathbb{R})\text{-measurable} \\ \text{and } f(\omega, \cdot) \text{ is an lsc proper convex function for a.e. } \omega. \end{aligned} \quad (8)$$

Since $\mathcal{F}$ is assumed $\mathbb{P}$ -complete, the first part is equivalent to the measurability of epigraphical mapping (see [24, Ch. 14] for a general reference). As immediate consequences of (8), $f(\cdot, \xi)$ is $\mathcal{F}$ -measurable for each $\xi \in L^0$, and f^* is also a proper normal convex integrand where

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

Given such f and γ , we define a robust analogue of convex integral functional

$$\mathcal{I}_{f,\gamma}(\xi) := \sup_{Q \in \mathcal{Q}_\gamma} (\mathbb{E}_Q[f(\cdot, \xi)] - \gamma(Q)) = \rho_\gamma(f(\cdot, \xi)), \quad \xi \in L^\infty. \quad (9)$$

This functional is well-defined as a proper convex functional on L^∞ as soon as:

$$\exists \xi_0 \in L^\infty \text{ s.t. } f(\cdot, \xi_0)^+ \in M^{\rho_\gamma}, \quad (10)$$

$$\exists \eta_0 \in L^{\rho_\gamma} \text{ s.t. } f^*(\cdot, \eta_0)^+ \in L^{\rho_\gamma}. \quad (11)$$

Indeed, since $f(\cdot, \xi) \geq \xi \eta_0 - f^*(\cdot, \eta_0)^+$, (11) implies $f(\cdot, \xi)^- \in L^{\rho_\gamma} \subset \bigcap_{Q \in \mathcal{Q}_\gamma} L^1(Q)$, so $\mathbb{E}_Q[f(\cdot, \xi)]$ is well-defined with values in $(-\infty, \infty]$ for each $Q \in \mathcal{Q}_\gamma$ and $\xi \in L^\infty$, thus so is $\mathcal{I}_{f,\gamma}(\xi) = \sup_{Q \in \mathcal{Q}_\gamma} (\mathbb{E}_Q[f(\cdot, \xi)] - \gamma(Q))$, while (10) shows that $\mathcal{I}_{f,\gamma}(\xi_0) < \infty$, thus $\mathcal{I}_{f,\gamma} \not\equiv \infty$. Also, $\mathcal{I}_{f,\gamma}$ is convex as the pointwise supremum of convex functions. Note also that

$$\{\xi \in L^\infty : f(\cdot, \xi)^+ \in M^{\rho_\gamma}\} \subset \text{dom}(\mathcal{I}_{f,\gamma}) \subset \{\xi \in L^\infty : f(\cdot, \xi)^+ \in L^{\rho_\gamma}\}. \quad (12)$$

where $\text{dom}(\mathcal{I}_{f,\gamma}) := \{\xi \in L^\infty : \mathcal{I}_{f,\gamma}(\xi) < \infty\}$. In particular, by (5),

$$\xi \in \text{dom}(\mathcal{I}_{f,\gamma}) \Rightarrow \xi(\omega) \in \text{dom} f(\omega, \cdot) \text{ for a.e. } \omega. \quad (13)$$

If $M^{\rho_\gamma} = L^{\rho_\gamma}$, all three sets in (12) coincide, so $\rho_\gamma(f(\cdot, \xi)^+) < \infty$ as soon as $\mathcal{I}_{f,\gamma}(\xi) < \infty$. In general, however, $\rho_\gamma(f(\cdot, \xi)^+) = \infty$ may happen even if $\xi \in \text{dom}(\mathcal{I}_{f,\gamma})$.

We next check that $\mathcal{I}_{f,\gamma}$ has a nice regularity on L^∞ .

Lemma 2.3. *Under Assumption 2.1, (10) and (11), $\mathcal{I}_{f,\gamma}$ is $\sigma(L^\infty, L^1)$ -lower semicontinuous, or equivalently $\mathcal{I}_{f,\gamma}$ has the Fatou property:*

$$\sup_n \|\xi_n\|_\infty < \infty, \xi_n \rightarrow \xi \text{ a.s.} \Rightarrow \mathcal{I}_{f,\gamma}(\xi) \leq \liminf_n \mathcal{I}_{f,\gamma}(\xi_n). \quad (14)$$

Proof. Suppose $a := \sup_n \|\xi_n\|_\infty < \infty$ and $\xi_n \rightarrow \xi$ a.s., then automatically $\xi \in L^\infty$. With η_0 as in (11), $f(\cdot, \xi_n) \geq -a|\eta_0| - f^*(\cdot, \eta_0)^+ \in L^{\rho_\gamma} \subset \bigcap_{Q \in \mathcal{Q}_\gamma} L^1(Q)$. Hence for each $Q \in \mathcal{Q}_\gamma$, Fatou's lemma shows that $\mathbb{E}_Q[f(\cdot, \xi)] \leq \liminf_n \mathbb{E}_Q[f(\cdot, \xi_n)]$, and we deduce (14) as

$$\begin{aligned} \sup_{Q \in \mathcal{Q}_\gamma} (\mathbb{E}_Q[f(\cdot, \xi)] - \gamma(Q)) &\leq \sup_{Q \in \mathcal{Q}_\gamma} \left(\liminf_n \mathbb{E}_Q[f(\cdot, \xi_n)] - \gamma(Q) \right) \\ &\leq \liminf_n \sup_{Q \in \mathcal{Q}_\gamma} (\mathbb{E}_Q[f(\cdot, \xi_n)] - \gamma(Q)). \end{aligned}$$

The equivalence between (14) and the weak*-lower semicontinuity follows from a well-known consequence of the Krein-Šmulian and Mackey-Arens theorems that (see e.g. [11]) a *convex* set $C \subset L^\infty$ is weak*-closed if and only if $C \cap \{\xi : \|\xi\|_\infty \leq a\}$ is L^0 -closed for all $a > 0$. $\square$

2.3. Robust f^* -divergence

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

$$\tilde{f}^*(\omega, y, z) := \sup_{x \in \text{dom} f(\omega, \cdot)} (xy - zf(\omega, x)), \quad \omega \in \Omega, (y, z) \in \mathbb{R} \times \mathbb{R}_+.$$

Noting that $(a_{f^*}^-, a_{f^*}^+) \subset \text{dom } f \subset [a_{f^*}^-, a_{f^*}^+]$ where $a_{f^*}^\pm := \lim_{k \rightarrow \pm} f^*(\cdot, k)/k$, we have

$$\tilde{f}^*(\omega, y, z) = \begin{cases} 0 & \text{if } y = z = 0, \\ y \cdot a_{f^*}^\pm(\omega) & \text{if } y \geq 0, z = 0, \\ z f^*(\omega, y/z) & \text{if } z > 0. \end{cases} \quad (15)$$

Lemma 2.4. *Suppose (8). Then $\tilde{f}^* : \Omega \times \mathbb{R} \times \mathbb{R}_+ \rightarrow \overline{\mathbb{R}}$ is a proper 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. $\omega \in \Omega$,*

$$xy \leq z f(\omega, x) + \tilde{f}^*(\omega, y, z), \quad \forall x \in \text{dom } f(\omega, \cdot), \forall y \in \mathbb{R}, \forall z \geq 0. \quad (16)$$

Proof. Since f is normal, there exists a sequence of measurable functions $(\xi_n)_{n \in \mathbb{N}} \subset L^0$ such that $\{\xi_n(\omega)\}_n \cap \text{dom } f(\omega, \cdot)$ is dense in $\text{dom } f(\omega, \cdot)$ ([23], Proposition 2D). Modifying the sequence as $\bar{\xi}_n := \xi_n \mathbb{1}_{\{f(\cdot, \xi_n) < \infty\}} + \xi_0 \mathbb{1}_{\{f(\cdot, \xi_n) = \infty\}}$ where $\xi_0 \in \text{dom } \mathcal{I}_{f, \gamma}$, we have for a.e. ω , $\bar{\xi}_n(\omega) \in \text{dom } f(\omega, \cdot)$ and $\{\bar{\xi}_n(\omega)\}_n$ is dense in $\text{dom } f(\omega, \cdot)$. Thus

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

Consequently, $\tilde{f}^*$ is a normal convex integrand as the countable supremum of affine integrands with $\tilde{f}^*(\cdot, 0, 0) = 0$. (16) is obvious from the definition. $\square$

Now we define $\mathcal{H}_{f^*}(\eta|Q) := \mathbb{E}[\tilde{f}^*(\cdot, \eta, dQ/d\mathbb{P})]$ for $\eta \in L^1$, $Q \in \mathcal{Q}_\gamma$, and

$$\mathcal{H}_{f^*, \gamma}(\eta) := \inf_{Q \in \mathcal{Q}_\gamma} (\mathcal{H}_{f^*}(\eta|Q) + \gamma(Q)), \quad \eta \in L^1. \quad (17)$$

In view of identification $\{\nu \in ba : \sigma\text{-additive}\} = L^1$, we define also for any σ -additive $\nu \in ba$,

$$\mathcal{H}_{f^*}(\nu|Q) := \mathcal{H}_{f^*}(d\nu/d\mathbb{P}|Q), \quad \mathcal{H}_{f^*, \gamma}(\nu) = \mathcal{H}_{f^*, \gamma}(d\nu/d\mathbb{P}).$$

Since $\tilde{f}^*(\cdot, y, 1) = f^*(\cdot, y)$, we recover $\mathcal{H}_{f^*}(\eta|\mathbb{P}) = \mathcal{H}_{f^*, \delta_{\{\mathbb{P}\}}}(\eta) = I_{f^*}(\eta)$ in the classical case. Under the above assumptions, $\mathcal{H}_{f^*}(\cdot|\cdot)$ and $\mathcal{H}_{f^*, \gamma}$ are well-defined.

Lemma 2.5. *Under Assumption 2.1, (8), (10) and (11), $\mathcal{H}_{f^*}(\cdot|\cdot)$ and $\mathcal{H}_{f^*, \gamma}$ are well-defined as proper convex functionals respectively on $L^1 \times \mathcal{Q}_\gamma$ and L^1 , and it holds*

$$\mathbb{E}[\eta \xi] \leq \mathcal{I}_{f, \gamma}(\xi) + \mathcal{H}_{f^*, \gamma}(\eta), \quad \forall \xi \in L^\infty, \forall \eta \in L^1. \quad (18)$$

Proof. Let $\psi_Q := dQ/d\mathbb{P}$ for each $Q \in \mathcal{Q}_\gamma$. In view of (13), we see that

$$\tilde{f}^*(\cdot, \eta, \psi_Q) \geq \xi \eta - \psi_Q f(\cdot, \xi) \geq \xi \eta - \psi_Q f(\cdot, \xi)^+ \in L^1, \quad (19)$$

for any $\xi \in \text{dom}(\mathcal{I}_{f,\gamma})$ ($\neq \emptyset$), $\eta \in L^1$ and $Q \in \mathcal{Q}_\gamma$, thus $\mathcal{H}_{f^*}(\eta|Q) = \mathbb{E}[\tilde{f}^*(\cdot, \eta, \psi_Q)] > -\infty$ is well-defined, convex on $L^1 \times \mathcal{Q}_\gamma$ (since $\tilde{f}^*$ is convex), and proper since $\mathcal{H}_{f^*}(\eta_0\psi_Q|Q) = \mathbb{E}_Q[f^*(\cdot, \eta_0)]$ with $\eta_0 \in L^{\rho_\gamma}$ as in (11) (then $\eta_0\psi_Q \in L^1$). Taking the expectation in (19)

$$\mathcal{H}(\eta|Q) + \gamma(Q) \geq \mathbb{E}[\xi\eta] - (\mathbb{E}_Q[f(\cdot, \xi)] - \gamma(Q)) \geq \mathbb{E}[\xi\eta] - \mathcal{I}_{f,\gamma}(\xi) > -\infty,$$

for any $\xi \in \text{dom}(\mathcal{I}_{f,\gamma})$, $\eta \in L^1$ and $Q \in \mathcal{Q}_\gamma$, so $\mathcal{H}_{f^*,\gamma}(\eta) = \inf_{Q \in \mathcal{Q}_\gamma} (\mathcal{H}_{f^*}(\eta|Q) + \gamma(Q)) > -\infty$ and we have (18) (which is trivially true when $\mathcal{I}_{f,\gamma}(\xi) = \infty$). The convexity of $\mathcal{H}_{f^*,\gamma}$ follows from that of $\mathcal{H}_{f^*}(\cdot|\cdot) + \gamma(\cdot)$ and of $\mathcal{Q}_\gamma$. $\square$

Remark 2.6. Under assumption (22) in the (main) Theorem 3.1 below, $\mathcal{H}_{f^*} + \gamma$ (resp. $\mathcal{H}_{f^*,\gamma}$) is weakly lower semicontinuous on $L^1 \times \mathcal{Q}_\gamma$ (resp. L^1), and for each $\eta \in L^1$, the infimum $\inf_{Q \in \mathcal{Q}_\gamma} (\mathcal{H}_{f^*}(\eta|Q) + \gamma(Q))$ is attained (see Appendix A). Though we could give direct proofs here, we will not use these properties, and the lower semicontinuity of $\mathcal{H}_{f^*,\gamma}$ will be obtained as Corollary 3.2 to the main theorem which does not internally use that property.

Note also that when f (hence f^* too) is non-random, all the integrability assumptions are trivialized, in which case $\mathcal{H}_{f^*}(\cdot|\cdot)$ is called the f^* -divergence while $\mathcal{H}_{f^*,\gamma}$ is a slight generalization of *robust f^* -divergence* (the latter is a special case with $\gamma(Q) = \delta_{\mathcal{P}}(Q)$ with $\mathcal{P} \subset \mathcal{Q}$), and the joint lower semicontinuity etc are found e.g. in [9, Lemma 2.7]. In fact, if f is finite ($\mathbb{P}(f(\cdot, x) < \infty, \forall x) = 1 \Leftrightarrow \lim_{|y| \rightarrow \infty} f^*(\cdot, y)/|y| = \infty$ a.s.), we have from (15) that

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

Here are some typical examples of penalty function γ and associated integral functionals.

Example 2.7 (Classical case). The “classical” integral functional $I_{f,\mathbb{P}}(\xi) := \mathbb{E}[f(\cdot, \xi)]$ corresponds to the penalty function $\gamma_{\mathbb{P}}(Q) := \delta_{\{\mathbb{P}\}}(Q)$ which clearly satisfies Assumption 2.1, and $\rho_{\gamma_{\mathbb{P}}}(\xi) = \mathbb{E}[\xi]$. Then $M_u^{\rho_\gamma} = M^{\rho_\gamma} = L^{\rho_\gamma} = L^1$, and the integrability assumptions (10) and (11) are identical to the ones in [22]:

$$\exists \xi_0 \in L^\infty \text{ with } f(\cdot, \xi_0)^+ \in L^1 \quad \text{and} \quad \exists \eta_0 \in L^1 \text{ with } f^*(\cdot, \eta_0)^+ \in L^1,$$

under which (1) holds true ([22], Th. 1), and since $I_{f^*,\mathbb{P}}(\eta) = \mathcal{H}_{f^*}(\eta|\mathbb{P})$, it reads as

$$I_{f,\mathbb{P}}^*(\nu) = \mathcal{H}_{f^*}(\nu_r|\mathbb{P}) + \sup_{\xi \in \text{dom}(I_{f,\mathbb{P}})} \nu_s(\xi), \quad \forall \nu \in ba.$$

We can also consider other probability $P \ll \mathbb{P}$ and $I_{f,P} := I_{f,\gamma_P}$ where $\gamma_P = \delta_{\{P\}}$, but we need a little care when $P \not\sim \mathbb{P}$ (then (5) is violated): here we consider $I_{f,P}$ as a functional on $L^\infty(\mathbb{P})$ rather than the space of P -equivalence classes of

P -essentially bounded random variables. Equivalently, $I_{f,P} = I_{g,\mathbb{P}}$ with $g(\cdot, x) = f(\cdot, x) \frac{dP}{d\mathbb{P}} \mathbf{1}_{\{dP/d\mathbb{P} > 0\}}$, and its conjugate is

$$I_{f,P}^*(\nu) = \mathcal{H}_{f^*}(\nu_r|P) + \infty \mathbf{1}_{\{\nu_r \not\ll P\}} + \sup_{\xi \in \text{dom}(I_{f,P})} \nu_s(\xi). \quad (20)$$

Example 2.8 (Homogeneous case). The formulation (9) covers the following form

$$\mathcal{I}_{f,\mathcal{P}}(\xi) := \sup_{Q \in \mathcal{P}} \mathbb{E}_Q[f(\cdot, \xi)] = \mathcal{I}_{f,\delta_{\mathcal{P}}}(\xi),$$

where $\mathcal{P} \subset \mathcal{Q}$ is a nonempty convex set. $\delta_{\mathcal{P}}$ clearly satisfies (4), and (6) (resp. (5)) is equivalent to the weak compactness of $\mathcal{P}$ itself (resp. $\exists Q \in \mathcal{P}$ with $Q \sim \mathbb{P}$), and $\rho_{\mathcal{P}}(\xi) := \rho_{\delta_{\mathcal{P}}}(\xi) = \sup_{Q \in \mathcal{P}} \mathbb{E}_Q[\xi]$ is a *positively homogeneous* monotone convex function, called *sublinear expectation or coherent risk measure* (modulo change of sign). In particular, $M^{\rho_{\mathcal{P}}} = L^{\rho_{\mathcal{P}}}$, while $M_u^{\rho_{\mathcal{P}}} \subsetneq M^{\rho_{\mathcal{P}}}$ is possible (see [19, Example 3.7]).

Example 2.9 (Polyhedral case). This is a special case of Example 2.8. Suppose we are given a finite number of probability measures $P_1, \dots, P_n \ll \mathbb{P}$ which generate a polyhedral convex set $\mathcal{P} = \text{conv}(P_1, \dots, P_n)$. This $\mathcal{P}$ is clearly (convex and) weakly compact in L^1 , (5) is equivalent to $\frac{1}{n}(P_1, \dots, P_n) \sim \mathbb{P}$, and we have $M_u^{\rho_{\mathcal{P}}} = M^{\rho_{\mathcal{P}}} = L^{\rho_{\mathcal{P}}} = \bigcap_{k \leq n} L^1(P_k)$ since

$$\frac{1}{n} \sum_{k \leq n} \mathbb{E}_{P_k}[\|\xi\|] \leq \|\xi\|_{\rho_{\mathcal{P}}} = \max_{1 \leq k \leq n} \mathbb{E}_{P_k}[\|\xi\|] \leq \sum_{k \leq n} \mathbb{E}_{P_k}[\|\xi\|].$$

In particular, noting that $I_{f,(\lambda_1 P_1 + \dots + \lambda_n P_n)} = \lambda_1 I_{f,P_1} + \dots + \lambda_n I_{f,P_n}$,

$$\mathcal{I}_{f,\mathcal{P}}(\xi) = \sup_{Q \in \mathcal{P}} I_{f,Q}(\xi) = \max_{1 \leq k \leq n} I_{f,P_k}(\xi), \quad \text{dom}(\mathcal{I}_{f,\mathcal{P}}) = \bigcap_{k \leq n} \text{dom}(I_{f,P_k}).$$

In [28, Cor. 2.8.11], the conjugate of pointwise maximum of *finitely many* convex functions is obtained, which reads in our context as (compare to (25) below): if $\text{dom}(\mathcal{I}_{f,\mathcal{P}}) \neq \emptyset$,

$$\begin{aligned} \mathcal{I}_{f,\mathcal{P}}^*(\nu) &= \min \{ \varphi^*(\nu) : \varphi \in \text{conv}(I_{f,P_1}, \dots, I_{f,P_n}) \} \\ &= \min_{Q \in \mathcal{P}} I_{f,Q}^*(\nu) \stackrel{(20)}{=} \min_{Q \in \mathcal{P}} \left\{ \mathcal{H}_{f^*}(\nu_r|Q) + \infty \mathbf{1}_{\{\nu_r \not\ll Q\}} + \sup_{\xi \in \text{dom}(I_{f,Q})} \nu_s(\xi) \right\}. \end{aligned} \quad (21)$$

Example 2.10 (Entropic penalty). Let $\gamma_{\text{ent}}(Q)$ be the relative entropy of Q w.r.t. $\mathbb{P}$:

$$\gamma_{\text{ent}}(Q) := \mathcal{H}_{x \log x}(Q|\mathbb{P}) := \mathbb{E} \left[\frac{dQ}{d\mathbb{P}} \log \frac{dQ}{d\mathbb{P}} \right].$$

This function satisfies Assumption 2.1: $\gamma_{\text{ent}}(\mathbb{P}) = 0$, hence (4) and (5), while (6) follows from the de la Vallée-Poussin theorem since $\lim_{x \rightarrow \infty} \frac{x \log x}{x} = \infty$. Let

$$L^{\Phi_{\text{exp}}} := \{\xi \in L^0 : \exists \alpha > 0, \mathbb{E}[\exp(\alpha|\xi|)] < \infty\} \quad (\text{exponential Orlicz space}),$$

$$M^{\Phi_{\text{exp}}} := \{\xi \in L^0 : \forall \alpha > 0, \mathbb{E}[\exp(\alpha|\xi|)] < \infty\} \quad (\text{Morse subspace}).$$

In this case, $\rho_{\text{ent}}(\xi) := \rho_{\gamma_{\text{ent}}}(\xi) = \log \mathbb{E}[\exp(\xi)]$ whenever $\xi^- \in L^{\Phi_{\text{exp}}}$. Hence, $M_u^{\rho_\gamma} = M^{\rho_\gamma} = M^{\Phi_{\text{exp}}} \subset L^{\Phi_{\text{exp}}} = L^{\rho_\gamma}$ and $M^{\Phi_{\text{exp}}} \subsetneq L^{\Phi_{\text{exp}}}$ if $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless (e.g. exponential random variable). The integrability assumptions (10) for $\xi_0 \in L^\infty$ and (11) for $\eta_0 \in L^{\Phi_{\text{exp}}}$ read as $\mathbb{E}[\exp(\alpha f(\cdot, \xi_0))] < \infty$ for all $\alpha > 0$ and $\mathbb{E}[\exp(\varepsilon f^*(\cdot, \eta_0))] < \infty$ for some $\varepsilon > 0$, respectively. Moreover, the corresponding integral functional is explicitly written as

$$\mathcal{I}_{f, \gamma_{\text{ent}}}(\xi) = \log \mathbb{E}[\exp(f(\cdot, \xi))], \quad \forall \xi \in L^\infty.$$

3. Statements of Main Results

3.1. A Rockafellar-Type Theorem for the Convex Conjugate

The following is a *robust analogue* of the Rockafellar theorem for the conjugate of $\mathcal{I}_{f, \gamma}$

$$\mathcal{I}_{f, \gamma}^*(\nu) := \sup_{\xi \in L^\infty} (\nu(\xi) - \mathcal{I}_{f, \gamma}(\xi)), \quad \nu \in ba = ba(\Omega, \mathcal{F}, \mathbb{P}).$$

Theorem 3.1. *Suppose Assumption 2.1, (8), (11) and*

$$\exists \xi_0 \in L^\infty \text{ such that } f(\cdot, \xi_0)^+ \in M_u^{\rho_\gamma}. \quad (22)$$

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

$$\mathcal{H}_{f^*, \gamma}(\nu_r) + \sup_{\xi \in \mathcal{D}_{f, \gamma}} \nu_s(\xi) \leq \mathcal{I}_{f, \gamma}^*(\nu) \leq \mathcal{H}_{f^*, \gamma}(\nu_r) + \sup_{\xi \in \text{dom}(\mathcal{I}_{f, \gamma})} \nu_s(\xi), \quad (23)$$

where $\mathcal{D}_{f, \gamma} := \{\xi \in L^\infty : f(\cdot, \xi)^+ \in M_u^{\rho_\gamma}\} \subset \text{dom}(\mathcal{I}_{f, \gamma})$ (by (12)).

A proof is given in Section 4.1. In contrast to the classical Rockafellar theorem (1), our robust version (23) consists of two inequalities instead of a single equality. But the possible difference appears only in the singular part, thus

Corollary 3.2 (Restriction to L^1). *Under the same assumptions as in Theorem 3.1,*

$$\mathcal{I}_{f, \gamma}^*(\eta) = \mathcal{H}_{f^*, \gamma}(\eta) = \inf_{Q \in \mathcal{Q}_\gamma} (\mathcal{H}_{f^*}(\eta|Q) + \gamma(Q)), \quad \forall \eta \in L^1.$$

In particular, $\mathcal{H}_{f^, \gamma}$ is weakly lower semicontinuous on L^1 , and*

$$\mathcal{I}_f(\xi) = \sup_{\eta \in L^1} (\mathbb{E}[\xi\eta] - \mathcal{H}_{f^*, \gamma}(\eta)), \quad \xi \in L^\infty. \quad (24)$$

Proof. The first assertion is clear from (23), by which the conjugate $\mathcal{H}_{f^*, \gamma}$ is lower semicontinuous for any topology consistent with the duality $\langle L^\infty, L^1 \rangle$, while (24) is a consequence of $\sigma(L^\infty, L^1)$ -lower semicontinuity of $\mathcal{I}_{f, \gamma}$ (Lemma 2.3) via the Fenchel-Moreau theorem. $\square$

In the classical case of Example 2.7, $\mathcal{D}_{f,\gamma} = \text{dom}(\mathcal{I}_{f,\gamma}) = \{\xi \in L^\infty : f(\cdot, \xi)^+ \in L^1\}$ (since $M_u^{\rho_\gamma} = L^{\rho_\gamma} = L^1$), hence (23) reduces to a single equality which is exactly (1) as in the Rockafellar theorem [22, Theorem 1]. The original version of [22] is slightly 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 $\xi \in L^\infty(\Omega, \mathcal{F}, \mu; \mathbb{R}^d)$. There are also some extensions replacing $L^\infty(\Omega, \mathcal{F}, \mu; \mathbb{R}^d)$ by some decomposable spaces of measurable functions taking values in a Banach space. See in this line [14], [4], and [23] for a general reference.

In the polyhedral case of Example 2.9 ($\gamma = \delta_{\mathcal{P}}$ and $\mathcal{P} = \text{conv}(P_1, \dots, P_n)$), we still have $\mathcal{D}_{f,\gamma} = \text{dom}(\mathcal{I}_{f,\gamma}) = \bigcap_{k \leq n} \text{dom}(I_{f,P_k})$. Thus (23) reduces to

$$\mathcal{I}_{f,\gamma}^*(\nu) = \min_{Q \in \mathcal{P}} \mathcal{H}_{f^*}(\nu_r|Q) + \sup \{ \nu_s(\xi) : \xi \in \bigcap_{k \leq n} \text{dom}(I_{f,P_k}) \}. \quad (25)$$

This is slightly sharper than (21) in the sense that regular and singular parts are separated.

To the best of our knowledge, Rockafellar-type result for the robust form (9) of convex integral functionals (including the homogeneous case of Example 2.8) is new. A possible complaint would be the difference between singular parts in the upper and lower bounds in (23). In the full generality of Theorem 3.1, however, *both* inequalities can really be strict and one can not hope for sharper bounds as the next example illustrates (see Appendix B for details).

Example 3.3 (Badly Behaving Integrand). Let $(\Omega, \mathcal{F}) := (\mathbb{N}, 2^{\mathbb{N}})$ with $\mathbb{P}$ given by $\mathbb{P}(\{n\}) = 2^{-n}$, and $(P_n)_n$ a sequence of probability measures on $2^{\mathbb{N}}$ specified by $P_1(\{1\}) = 1$; $P_n(\{1\}) = 1 - 1/n$, $P_n(\{n\}) = 1/n$. Then $\mathcal{P} = \overline{\text{conv}}(P_n; n \in \mathbb{N})$ is weakly compact in $L^1(\mathbb{N}, 2^{\mathbb{N}}, \mathbb{P})$, thus $\gamma = \delta_{\mathcal{P}}$ is a penalty function satisfying Assumption 2.1 and $\rho_\gamma(\xi) = \sup_n \mathbb{E}_{P_n}[\xi]$ if $\xi \in L_+^0$. In this case, L^∞ is regarded as the sequence space ℓ^∞ with the norm $\|\xi\|_\infty = \sup_n |\xi(n)|$, and $\nu \in ba_+^s(\mathbb{N}, 2^{\mathbb{N}}, \mathbb{P})$ if and only if ν vanishes on any finite set, or equivalently, for any $\nu \in ba_+$,

$$\nu \in ba_+^s \Leftrightarrow \|\nu\| \cdot \liminf_n \xi(n) \leq \nu(\xi) \leq \|\nu\| \cdot \limsup_n \xi(n), \quad \forall \xi \in \ell^\infty = L^\infty. \quad (26)$$

(Such $\nu \neq 0$ exists, thus $ba_+^s \setminus \{0\} \neq \emptyset$; see [1, Lemmas 16.29 and 16.30]). Now we set

$$f(n, x) = nx^+ e^x, \quad n \in \mathbb{N} = \Omega, \quad x \in \mathbb{R}. \quad (27)$$

Then $\mathcal{I}_{f,\gamma}(\xi) = \sup_n ((1 - \frac{1}{n})\xi(1)^+ e^{\xi(1)} + \xi(n)^+ e^{\xi(n)^+}) \leq 2\|\xi\|_\infty e^{\|\xi\|_\infty}$, so $\text{dom}(\mathcal{I}_{f,\gamma}) = L^\infty$, and $\lim_{N \rightarrow \infty} \sup_n \mathbb{E}_{P_n}[f(\cdot, \xi) \mathbf{1}_{\{f(\cdot, \xi) \geq N\}}] = \limsup_n \xi(n)^+ e^{\xi(n)^+}$ (Lemma B.1), thus

$$0 \in \mathcal{D}_{f,\gamma} = \{\xi \in \ell^\infty : \limsup_n \xi(n) \leq 0\} \subsetneq \text{dom}(\mathcal{I}_{f,\gamma}) = L^\infty. \quad (28)$$

As for $\mathcal{I}_{f,\gamma}^*$, $\text{dom}(\mathcal{I}_{f,\gamma}^*) \subset ba_+$ since $\mathcal{I}_{f,\gamma}$ is increasing, and $\mathcal{H}_{f^*,\gamma}(0) = 0$ since $f^*(\cdot, 0) = \inf_x f(\cdot, x) = 0$, thus (23) reads as $\sup_{\xi \in \mathcal{D}_{f,\gamma}} \nu(\xi) \leq \mathcal{I}_{f,\gamma}^*(\nu) \leq$

$\sup_{\xi \in \text{dom}(\mathcal{I}_{f,\gamma})} \nu(\xi)$ on ba_+^s . On the other hand, for $\nu \in ba_+^s$, $\sup_{\xi \in \mathcal{D}_{f,\gamma}} \nu(\xi) = 0$, $\sup_{\xi \in \text{dom} \mathcal{I}_f} \nu(\xi) = +\infty$, and (Lemma B.2):

$$\mathcal{I}_{f,\gamma}^*(\nu) = \sup_{x \geq 0} x(\|\nu_s\| - e^x), \quad \forall \nu \in ba_+^s.$$

In particular, $\mathcal{I}_{f,\gamma}^*(\nu) = 0$ if $\nu \in U_+^s := \{\nu \in ba_+^s : \|\nu\| = \nu(\mathbb{N}) \leq 1\}$, $0 < \mathcal{I}_{f,\gamma}^*(\nu) < \infty$ if $\nu \in ba_+^s \setminus U_+^s := \{\nu \in ba_+^s : \|\nu\| > 1\}$, and $\lim_{\|\nu\| \rightarrow \infty, \nu \in ba_+^s \setminus U_+^s} \mathcal{I}_{f,\gamma}^*(\nu) = \infty$. In summary,

- (a) $\mathcal{I}_{f,\gamma}^*$ coincides with the lower bound $\sup_{\xi \in \mathcal{D}_{f,\gamma}} \nu(\xi) = 0$ on U_+^s , while
- (b) on $ba_+^s \setminus U_+^s$, $\mathcal{I}_{f,\gamma}^*$ is strictly between the upper and lower bounds and it runs through the whole interval of these bounds (in this specific case, $[0, \infty]$).

3.2. Finer Properties in the Finite-Valued Case

We now consider the regularities of $\mathcal{I}_{f,\gamma}$ and $\mathcal{H}_{f^*,\gamma}$ in terms of the dual pairing $\langle L^\infty, L^1 \rangle$. In the classical case of Example 2.7, the singular part of I_f^* in (1) is trivialized (i.e., $\delta_{\{0\}}$) as soon as $I_f := \mathcal{I}_{f,\{\mathbb{P}\}}$ is finite-valued, then I_f^* reduces entirely to I_{f^*} . It implies that all the sublevels of I_{f^*} are $\sigma(L^1, L^\infty)$ -compact (see [23, Th. 3K]), which is equivalent to the continuity of I_f for the Mackey topology $\tau(L^\infty, L^1)$, and I_f admits a σ -additive subgradient at every point (weak* subdifferentiable). Consequently, we can work entirely with the dual pair $\langle L^\infty, L^1 \rangle$.

In the robust case, the “triviality of singular part of $\mathcal{I}_{f,\gamma}^*$ ” should be understood as

$$\forall \nu \in ba, \mathcal{I}_{f,\gamma}^*(\nu) < \infty \Rightarrow \nu \text{ is } \sigma\text{-additive}, \quad (29)$$

i.e. that $\mathcal{I}_{f,\gamma}^*$ eliminates the singular measures, which still makes sense even though $\mathcal{I}_{f,\gamma}^*$ itself need not be the direct sum of regular and singular parts. This guarantees in particular that (23) reduces to a single equality (of course). In the case of Example 3.3, $\mathcal{D}_{f,\gamma} \subsetneq \text{dom}(\mathcal{I}_{f,\gamma}) = L^\infty$, and $\mathcal{I}_{f,\gamma}^*(\nu_s) < \infty$ as long as $\nu_s \geq 0$. Thus $\text{dom}(\mathcal{I}_{f,\gamma}) = L^\infty$ is not enough for (29), while from (23), $\mathcal{D}_{f,\gamma} = L^\infty$ is clearly sufficient. In fact, given the finiteness (plus a technical assumption), $\mathcal{D}_{f,\gamma} = L^\infty$ is also necessary for (29), and equivalent to other basic $\langle L^\infty, L^1 \rangle$ -regularities of $\mathcal{I}_{f,\gamma}$ and $\mathcal{I}_{f,\gamma}^*$ which follow solely from the finiteness in the classical case.

Theorem 3.4. *In addition to the assumptions of Theorem 3.1, suppose $\text{dom}(\mathcal{I}_{f,\gamma}) = L^\infty$ and*

$$\exists \xi'_0 \in L^\infty \text{ with } f(\cdot, \xi'_0)^- \in M^{\rho_\gamma}. \quad (30)$$

Then the following are equivalent:

- (i) $\mathcal{D}_{f,\gamma} = L^\infty$, i.e., $f(\cdot, \xi)^+ \in M_u^{\rho_\gamma}$ for all $\xi \in L^\infty$;
- (ii) $\mathbb{R} \subset \mathcal{D}_{f,\gamma}$, i.e., $f(\cdot, x)^+ \in M_u^{\rho_\gamma}$ for all $x \in \mathbb{R}$;
- (iii) $\mathcal{I}_{f,\gamma}^*$ eliminates singular measures in the sense of (29);
- (iv) $\{\eta \in L^1 : \mathcal{H}_{f^*,\gamma}(\eta) \leq c\}$ is $\sigma(L^1, L^\infty)$ -compact for all $c \in \mathbb{R}$;

- (v) $\mathcal{I}_{f,\gamma}$ is continuous for the Mackey topology $\tau(L^\infty, L^1)$;
- (vi) $\mathcal{I}_{f,\gamma}(\xi) = \lim_n \mathcal{I}_{f,\gamma}(\xi_n)$ if $\sup_n \|\xi_n\|_\infty < \infty$ and $\xi_n \rightarrow \xi$ a.s. (the Lebesgue property);
- (vii) $\mathcal{I}_{f,\gamma}(\xi) = \max_{\eta \in L^1} (\mathbb{E}[\xi\eta] - \mathcal{H}_{f^*,\gamma}(\eta))$, i.e., the supremum is attained in (24), $\forall \xi \in L^\infty$.

Here implications (i) $\Leftrightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Leftrightarrow$ (v) $\Rightarrow$ (vi) and (v) $\Rightarrow$ (vii) are true without (30).

A proof will be given in Section 4.2.

Remark 3.5. The finiteness of $\mathcal{I}_{f,\gamma}$ already implies that $f(\cdot, \xi)^+ \in L^{\rho_\gamma}$, $\forall \xi \in L^\infty$ (in particular f is finite-valued), while the additional assumption (30) is made to guarantee $f(\cdot, \xi)^+ \in M^{\rho_\gamma}$ for all $\xi \in L^\infty$ (see (49) and the subsequent paragraph). These coincide in the homogeneous case (Example 2.8; including the classical case) since then $L^{\rho_\gamma} = M^{\rho_\gamma}$, but in general, $M^{\rho_\gamma} \subsetneq L^{\rho_\gamma}$ is possible. A sufficient condition for (30) is that

$$\exists \eta_0 \in M^{\rho_\gamma} \text{ with } f^*(\cdot, \eta_0)^+ \in M^{\rho_\gamma}, \quad (31)$$

(then $f(\cdot, \xi)^- \in M^{\rho_\gamma}$ for all $\xi \in L^\infty$), which is identical to (11) (contained in the standing assumptions) in the homogeneous case.

Remark 3.6. The equivalence between (iv), (vi) and (vii) for *convex risk measures* on L^∞ , i.e., for $\rho_\gamma|_{L^\infty}$ with penalty function γ as in Assumption 2.1 is known ([13] and [5]) followed by some generalizations: [16, 17] for convex risk measures on Orlicz spaces, and [19, 20] for *finite-valued* monotone convex functions on solid spaces of measurable functions among others.

From (ii) $\Rightarrow$ (v), we derive a simple criterion in terms of integrability of f for the $\langle L^\infty, L^1 \rangle$ -Fenchel duality for the minimization of robust integral functional $\mathcal{I}_{f,\gamma}$. Here we recall Fenchel's duality theorem (see [21, Th. 1]): if $\langle E, E' \rangle$ is a dual pair, φ, ψ are proper convex functions on E , and if either φ or ψ is $\tau(E, E')$ -continuous at some $x \in \text{dom}(\varphi) \cap \text{dom}(\psi)$, then

$$\inf_{x \in E} (\varphi(x) + \psi(x)) = - \min_{x' \in E'} (\varphi^*(x') + \psi^*(-x')).$$

Putting $E = L^\infty$, $E' = L^1$, $\varphi = \mathcal{I}_{f,\gamma}$ and $\psi = \delta_{\mathcal{C}}$ with $\mathcal{C} \subset L^\infty$ convex, (ii) $\Rightarrow$ (v) tells us that

Corollary 3.7 (Fenchel Duality). *Let γ be a penalty function satisfying Assumption 2.1 and f a proper normal convex integrand. If $f(\cdot, x)^+ \in M_u^{\rho_\gamma}$ for all $x \in \mathbb{R}$, and $f^*(\cdot, \eta_0)^+ \in L^{\rho_\gamma}$ for some $\eta_0 \in L^{\rho_\gamma}$, then for any convex set $\mathcal{C} \subset L^\infty$,*

$$\inf_{\xi \in \mathcal{C}} \mathcal{I}_{f,\gamma}(\xi) = - \min_{\eta \in L^1} \left(\mathcal{H}_{f^*,\gamma}(-\eta) + \sup_{\xi' \in \mathcal{C}} \mathbb{E}[\xi'\eta] \right).$$

If in addition $\mathcal{C}$ is a convex cone, the right hand side is equal to $-\min_{\eta \in \mathcal{C}^\circ} \mathcal{H}_{f^,\gamma}(-\eta)$, where $\mathcal{C}^\circ = \{\eta \in L^1 : \mathbb{E}[\xi\eta] \leq 1, \forall \xi \in \mathcal{C}\}$ (the one-sided polar of $\mathcal{C}$ in $\langle L^\infty, L^1 \rangle$).*

The *subdifferential* of $\mathcal{I}_{f,\gamma}$ at $\xi \in L^\infty$ is the following set of $\nu \in (L^\infty)^*$ called subgradients:

$$\partial\mathcal{I}_{f,\gamma}(\xi) := \{\nu \in (L^\infty)^* : \nu(\xi) - \mathcal{I}_{f,\gamma}(\xi) \geq \nu(\xi') - \mathcal{I}_{f,\gamma}(\xi'), \forall \xi' \in L^\infty\}.$$

We say that $\mathcal{I}_{f,\gamma}$ is subdifferentiable at ξ if $\partial\mathcal{I}_{f,\gamma}(\xi) \neq \emptyset$. In view of (24), $\eta \in \partial\mathcal{I}_{f,\gamma}(\xi) \cap L^1$ (then η is called a σ -additive subgradient of $\mathcal{I}_{f,\gamma}$ at ξ) if and only if it maximizes $\eta' \mapsto \mathbb{E}[\xi\eta'] - \mathcal{H}_{f^*,\gamma}(\eta')$, thus (vii) is equivalent to saying that for every $\xi \in L^\infty$, $\partial\mathcal{I}_{f,\gamma}(\xi) \cap L^1 \neq \emptyset$. Note also that $\partial\mathcal{I}_{f,\gamma}(\xi) \subset \text{dom}(\mathcal{I}_{f,\gamma}^*)$ since $\mathcal{I}_{f,\gamma}^*(\nu) = \sup_{\xi' \in L^\infty} (\nu(\xi') - \mathcal{I}_{f,\gamma}(\xi')) \leq \nu(\xi) - \mathcal{I}_{f,\gamma}(\xi) < \infty$ if $\nu \in \partial\mathcal{I}_{f,\gamma}(\xi)$. Thus (iii) implies $\partial\mathcal{I}_{f,\gamma}(\xi) \subset L^1$. Summing up,

Corollary 3.8. *Under the assumptions of Theorem 3.4, (i)–(vii) are equivalent also to*

(viii) $\emptyset \neq \partial\mathcal{I}_{f,\gamma}(\xi) \subset L^1$ for every $\xi \in L^\infty$.

The weak compactness of the sublevels of $\mathcal{H}_{f^*,\gamma}$ can be viewed as a generalization of the *de la Valle-Poussin theorem* which asserts that a set $\mathcal{C} \subset L^1$ is uniformly integrable if and only if there exists a function $g : \mathbb{R} \rightarrow (-\infty, \infty]$ which is *coercive*: $\lim_{|y| \rightarrow \infty} g(y)/y = \infty$ and $\sup_{\eta \in \mathcal{C}} \mathbb{E}[g(|\eta|)] = \sup_{\eta \in \mathcal{C}} \mathcal{H}_{g,\delta_{\{\mathbb{P}\}}}(|\eta|) < \infty$ (e.g. [6, Th. II.22]). The *coercivity condition* is equivalent to saying that $\text{dom}(g^*) = \mathbb{R}$. Now we have as a consequence of (ii) $\Leftrightarrow$ (iv):

Corollary 3.9 (cf. [9] when g is non-random). *A set $\mathcal{C} \subset L^1$ is uniformly integrable if and only if there exists a convex penalty function γ on $\mathcal{Q}$ satisfying Assumption 2.1 as well as a proper normal convex integrand g with $g^*(\cdot, x) \in M_u^{\rho_\gamma}$, $\forall x \in \mathbb{R}$, such that $\sup_{\eta \in \mathcal{C}} \mathcal{H}_{g,\gamma}(\eta) < \infty$.*

Proof. Let $f = g^*$, then $f^* = g^{**} = g$ (since normal). Then $\mathcal{D}_{f,\gamma} = L^\infty$ by assumption and $\mathcal{H}_{g,\gamma} = \mathcal{H}_{f^*,\gamma}$ is well-defined while $\sup_{\eta \in \mathcal{C}} \mathcal{H}_{g,\gamma}(\eta) < \infty$ guarantees (11) as well. Now the sufficiency is nothing but (ii) $\Rightarrow$ (iv), while the necessity is clear from the above paragraph. $\square$

3.3. Examples of “Nice” Integrands and Robust Utility Maximization

When f is *non-random* and finite, $\mathbb{R} \subset \mathcal{D}_{f,\gamma}$ is automatic, while $f^*(y) \in M_u^{\rho_\gamma}$ for any $y \in \text{dom } f^* \neq \emptyset$ (since constant). Here are some ways to generate “nice” random integrands.

Example 3.10 (Random scaling). Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be a (non-random) finite convex function $\not\equiv 0$, and $W \in L^0$ be strictly positive (i.e., $\mathbb{P}(W > 0) = 1$). Then put

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

In this case, $f^*(\omega, y) = g^*(y/W(\omega))$ and $\mathbb{R} \subset \mathcal{D}_{f,\gamma}$ is true if

$$\exists \delta > 0, p > 1 \quad \text{such that } g(-\delta W^p)^+ \vee g(\delta W^p)^+ \in M_u^{\rho_\gamma}. \quad (32)$$

Note that $|Wx| = \frac{\delta}{2}|W(2x/\delta)| \leq \frac{1}{2}(\delta W^p + \frac{2^q}{\delta^{q-1}}|x|^q)$ where $\frac{1}{p} + \frac{1}{q} = 1$. Applying the (quasi) convexity of g twice, (32) implies for each $x \in \mathbb{R}$

$$g(Wx) \leq g(-\delta W^p)^+ \vee g(\delta W^p)^+ \vee g\left(-\frac{2^q}{\delta^{q-1}}|x|\right)^+ \vee g\left(\frac{2^q}{\delta^{q-1}}|x|\right)^+ \in M_u^{\rho_\gamma}.$$

Also, since $g \not\equiv 0$, $\text{dom } g^* \setminus \{0\} \neq \emptyset$. If $y \in \text{dom } g^*$ and $y > 0$ (resp. $y < 0$),

$$0 \leq W \leq 1 + W^p \leq 1 + \frac{g(\delta W^p)^+ + g^*(y)}{y\delta}; \quad \text{resp.} \quad \leq 1 + \frac{g(-\delta W^p)^+ + g^*(y)}{-y\delta}.$$

In both cases, (32) implies $W \in M^{\rho_\gamma}$, and consequently, $\eta_y = yW \in M^{\rho_\gamma}$ and $f^*(\cdot, \eta_y) = g^*(y) \in L^\infty$. Thus (31) ($\Rightarrow$ (11)) follows from (32) as well. If in addition g is monotone increasing, $g(-W^p)^+ \leq g(0)^+$, thus the half of (32) is automatically true.

Example 3.11 (Random parallel shift). Let f be a finite normal convex integrand satisfying (11) and $\mathbb{R} \subset \mathcal{D}_f$, and $B \in L^0$. Then put

$$f_B(\omega, x) = f(\omega, x + B(\omega)), \quad (\omega, x) \in \Omega \times \mathbb{R}. \quad (33)$$

By convexity of f , $f(\cdot, x+B) \leq \frac{\varepsilon}{1+\varepsilon}f(\cdot, \frac{1+\varepsilon}{\varepsilon}x) + \frac{1}{1+\varepsilon}f(\cdot, (1+\varepsilon)B)$ and $f(\cdot, \frac{\varepsilon}{1+\varepsilon}x) \leq \frac{\varepsilon}{1+\varepsilon}f(\cdot, x+B) + \frac{1}{1+\varepsilon}f(\cdot, -\varepsilon B)$, thus putting $\Gamma_\alpha(x) = f(\cdot, \alpha x)^+/\alpha$,

$$\frac{1+\varepsilon}{\varepsilon}f\left(\cdot, \frac{\varepsilon}{1+\varepsilon}x\right) - \Gamma_\varepsilon(-B) \leq f_B(\cdot, x) \leq \frac{\varepsilon}{1+\varepsilon}f\left(\cdot, \frac{1+\varepsilon}{\varepsilon}x\right) + \Gamma_{1+\varepsilon}(B), \quad (34)$$

$$f^*(\cdot, y) - \Gamma_{1+\varepsilon}(B) \leq f_B^*(\cdot, y) \leq f^*(\cdot, y) + \Gamma_\varepsilon(-B), \quad (35)$$

where $f_B^*(\cdot, y) = f^*(\cdot, y) - yB$, and (35) follows from (34) by taking conjugates. Thus if

$$\exists \varepsilon > 0 \text{ such that } \Gamma_{1+\varepsilon}(B) \in M_u^{\rho_\gamma} \text{ and } \Gamma_\varepsilon(-B) \in L^{\rho_\gamma}, \quad (36)$$

then $\mathbb{R} \subset \mathcal{D}_{f_B, \gamma}$ and f_B^* satisfies (11) ((31) if f^* does). Moreover, (35) implies in this case

$$\mathcal{H}_{f_B^*, \gamma}(\eta) < \infty \Leftrightarrow \mathcal{H}_{f^*, \gamma}(\eta) < \infty \Rightarrow \eta B \in L^1,$$

and $\mathcal{H}_{f_B^*, \gamma}$ is explicitly given in terms of $\mathcal{H}_{f^*, \gamma}$ as

$$\mathcal{H}_{f_B^*, \gamma}(\eta) = \mathcal{H}_{f^*, \gamma}(\eta) - \mathbb{E}[\eta B], \quad \forall \eta \in \text{dom}(\mathcal{H}_{f_B^*, \gamma}) = \text{dom}(\mathcal{H}_{f^*, \gamma}). \quad (37)$$

We can combine the preceding two examples:

Example 3.12. Let $g : \mathbb{R} \rightarrow \mathbb{R}$, 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 (11) and $\mathbb{R} \subset \mathcal{D}_{h, \gamma}$ if (g, W) satisfies (32) and (36) holds with $f = g$. Note that if we apply Example 3.11 to $f(\cdot, x) = g(Wx)$ and B/W , then $h(\cdot, x) = f(\cdot, x + B/W)$ and e.g. $f(\cdot, (1+\varepsilon)B/W) = g((1+\varepsilon)B)$.

Our initial motivation was a duality method for *robust utility maximization* of the general form (3) with *random utility function* $U : \Omega \times \mathbb{R} \rightarrow \overline{\mathbb{R}}$. See [10] for the financial background of the problem. A motivational example of random utility is of the type $U_{D,B}(\cdot, x) = U(D^{-1}x + B)$ where $U : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ is a proper concave increasing function and $D, B \in L^0$ (with $D > 0$ a.s.) correspond respectively to the discount factor and a payoff of a claim. Then the problem is to

$$\text{maximize } u_{D,B,\gamma}(\xi) := \inf_{Q \in \mathcal{Q}} (\mathbb{E}_Q[U(D^{-1}\xi + B)] + \gamma(Q)) = -\mathcal{I}_{f_{D,B,\gamma}}(-\xi) \quad (38)$$

over a convex cone $\mathcal{C} \subset L^\infty$ where $f_{D,B}(\cdot, x) = -U(-D^{-1}x + B)$ is a proper normal convex integrand of the form in Example 3.12. The full detail of this problem in more concrete financial setup will be given in a separate future paper together with an application to a robust version of *utility indifference valuation*. Here we just give a criterion in terms of “integrabilities” of D and B for the duality without singular term as well as its explicit form *when U is finite on $\mathbb{R}$* . It constitutes a half of what we call the martingale duality method (see e.g. [2, 25, 3] for the other half in the classical case and [18]¹ for a partial result in the robust case). For the case $\text{dom}(U) = \mathbb{R}_+$, see [26] when D, B are constants; [27] with bounded B , and [9] for $\text{dom}(U) = \mathbb{R}$ with constant D, B ; see also [10] for more thorough references. The following is an immediate consequence of Corollary 3.7 and Example 3.12.

Corollary 3.13. *In the above notation, suppose U is finite on $\mathbb{R}$, and*

$$\exists \delta, \varepsilon > 0 \text{ with } U(-\delta D^{-(1+\varepsilon)})^- \in M_u^{\rho_\gamma}, U(-(1+\varepsilon)B)^- \in M_u^{\rho_\gamma}, U(\varepsilon B)^- \in L^{\rho_\gamma}. \quad (39)$$

Then for any nonempty convex cone $\mathcal{C} \subset L^\infty$, it holds that

$$\sup_{\xi \in \mathcal{C}} u_{B,D,\gamma}(\xi) = \min_{\eta \in \mathcal{C}_V^\circ} (\mathcal{H}_{V,\gamma}(D\eta) + \mathbb{E}[DB\eta]), \quad (40)$$

where $V(y) := \sup_{x \in \mathbb{R}} (U(x) - xy)$ and $\mathcal{C}_V^\circ := \{\eta \in \mathcal{C}^\circ : \mathcal{H}_{V,\gamma}(D\eta) < \infty\}$.

Here $\mathcal{C}$ can be any convex cone and the possibility of both sides being ∞ is not excluded; it does not happen iff $\mathcal{C}_V^\circ \neq \emptyset$. If in addition $\mathcal{C}_V^{\circ,\varepsilon} := \{\eta \in \mathcal{C}_V^\circ : \eta > 0 \text{ a.s.}\} \neq \emptyset$, we can replace “ $\min_{\eta \in \mathcal{C}_V^\circ}$ ” by “ $\inf_{\eta \in \mathcal{C}_V^{\circ,\varepsilon}}$ ” etc with a little more effort and certain regularities of U . Choosing a “good” cone $\mathcal{C}$, these conditions as well as the dual problem have clear financial interpretations and consequences (see [2] for a good exposition in the classical case).

A couple of features deserve attention: (i) We directly invoke Fenchel’s theorem to the functional $u_{D,B,\gamma} = -\mathcal{I}_{f_{D,B,\gamma}}(-\cdot)$ by means of our main theorem, instead of interchanging “ $\sup_{\xi \in \mathcal{C}}$ ” and “ $\inf_{Q \in \mathcal{Q}}$ ”, and invoke a *classical* duality (see [2, 3] and its references) under each Q , so we do not need to mind what happens under “extreme” $Q \in \mathcal{Q}_\gamma$; (ii) embedding the randomness D, B to the utility function

¹There an earlier version of this paper (still available as arXiv:1101.2968) was used.

U instead of transforming the domain $\mathcal{C}$ to $D^{-1}\mathcal{C} + B$, we retain the “good form” of $\mathcal{C}$ (which is essential for the probabilistic techniques to work well), and obtain a criterion for the duality in terms solely of B and D (when U is finite). Those integrability conditions are weak even in the classical case where $\gamma = \delta_{\{\mathbb{P}\}}$ and $D \equiv 1$ (then (39) reads as $U(-(1+\varepsilon)B)^-, U(\varepsilon B)^- \in L^1$ for *some* $\varepsilon > 0$ complementing the result of [3]).

4. Proofs

4.1. Proof of Theorem 3.1

The upper bound is simply a consequence of Young’s inequality (18) and $\nu = \nu_r + \nu_s$:

$$\mathcal{I}_{f,\gamma}^*(\nu) = \sup_{\xi \in \text{dom}(\mathcal{I}_{f,\gamma})} (\nu_r(\xi) - \mathcal{I}_{f,\gamma}(\xi) + \nu_s(\xi)) \stackrel{(18)}{\leq} \mathcal{H}_{f^*,\gamma}(\nu_r) + \sup_{\xi \in \text{dom}(\mathcal{I}_{f,\gamma})} \nu_s(\xi).$$

The lower bound is more involved. First, fix $\nu \in ba$ and define

$$L_\nu(Q, \xi) := \nu(\xi) - \mathbb{E}_Q[f(\cdot, \xi)] + \gamma(Q), \quad Q \in \mathcal{Q}_\gamma, \xi \in \mathcal{D}_{f,\gamma}.$$

It is finite-valued on $\mathcal{Q}_\gamma \times \mathcal{D}_{f,\gamma}$, convex in $Q \in \mathcal{Q}_\gamma$ and concave in $\xi \in \mathcal{D}_{f,\gamma}$. Moreover,

Lemma 4.1. *With the notation above, it holds that*

$$\inf_{Q \in \mathcal{Q}_\gamma} \sup_{\xi \in \mathcal{D}_{f,\gamma}} L_\nu(Q, \xi) = \sup_{\xi \in \mathcal{D}_{f,\gamma}} \inf_{Q \in \mathcal{Q}_\gamma} L_\nu(Q, \xi). \quad (41)$$

Proof. We claim that for any $\xi \in \mathcal{D}_{f,\gamma}$ and $c > 0$, $\Lambda_c(\xi) := \{Q \in \mathcal{Q}_\gamma : L_\nu(Q, \xi) \leq c\}$ is weakly compact in L^1 . Again let $\psi_Q := dQ/d\mathbb{P}$ for each $Q \in \mathcal{Q}_\gamma$. For the uniform integrability of $\Lambda_c(\xi)$, pick a $Q_0 \in \mathcal{Q}_\gamma$ with $\gamma(Q_0) = 0$ (see Remark 2.2) and observe that

$$\mathbb{E}_Q[f(\cdot, \xi)] \leq \mathbb{E} \left[2f(\cdot, \xi)^+ \frac{\psi_Q + \psi_{Q_0}}{2} \right] \leq \rho_\gamma(2f(\cdot, \xi)^+) + \frac{1}{2}\gamma(Q) < \infty, \quad (42)$$

since $\xi \in \mathcal{D}_{f,\gamma}$. Thus $\Lambda_c(\xi) \subset \{Q \in \mathcal{Q}_\gamma : \gamma(Q) \leq 2(c - \nu(\xi) + \rho_\gamma(2f(\cdot, \xi)^+))\}$, hence $\Lambda_c(\xi)$ is uniformly integrable by (6) and so is $\Lambda'_c(\xi) := \{f(\cdot, \xi)^+ \psi_Q : L_\nu(Q, \xi) \leq c\}$ by (7) since $\xi \in \mathcal{D}_{f,\gamma}$.

To see that $\Lambda_c(\xi)$ is weakly closed, it suffices to show that it is norm-closed since convex. So let $(Q_n)_n$ be a sequence in Λ_c converging in L^1 to some Q , then passing to a subsequence, we can suppose without loss that $\psi_{Q_n} \rightarrow \psi_Q$ a.s. too. From the previous paragraph, $(f(\cdot, \xi)^+ \psi_{Q_n})_n$ is uniformly integrable, hence by the (reverse) Fatou lemma, we have $\mathbb{E}_Q[f(\cdot, \xi)] \geq \limsup_n \mathbb{E}_{Q_n}[f(\cdot, \xi)]$ and consequently,

$$\begin{aligned} L_\nu(Q, \xi) &= \nu(\xi) - \mathbb{E}_Q[f(\cdot, \xi)] + \gamma(Q) \\ &\leq \nu(\xi) + \liminf_n -\mathbb{E}_{Q_n}[f(\cdot, \xi)] + \liminf_n \gamma(Q_n) \\ &\leq \liminf_n (\nu(\xi) - \mathbb{E}_{Q_n}[f(\cdot, \xi)] + \gamma(Q_n)) \leq c. \end{aligned}$$

We deduce that $\Lambda_c(\xi)$ is weakly closed, hence weakly compact. Now (41) follows from a minimax theorem (see [19, Appendix A]). $\square$

Noting that $\mathcal{I}_{f,\gamma}^*(\nu) \geq \sup_{\xi \in \mathcal{D}_{f,\gamma}} (\nu(\xi) - \mathcal{I}_{f,\gamma}(\xi)) = \sup_{\xi \in \mathcal{D}_{f,\gamma}} \inf_{Q \in \mathcal{Q}_\gamma} L_\nu(Q, \xi)$, we deduce from Lemma 4.1 that for any $\nu \in ba$,

$$\mathcal{I}_{f,\gamma}^*(\nu) \geq \inf_{Q \in \mathcal{Q}_\gamma} \sup_{\xi \in \mathcal{D}_{f,\gamma}} L_\nu(Q, \xi) = \inf_{Q \in \mathcal{Q}_\gamma} \sup_{\xi \in \mathcal{D}_{f,\gamma}} (\nu(\xi) - \mathbb{E}_Q[f(\cdot, \xi)] + \gamma(Q)). \quad (43)$$

Lemma 4.2. *Let $\eta, \zeta, \psi \in L^1$ such that $\psi \geq 0$ a.s. and*

$$\tilde{f}^*(\cdot, \eta, \psi) = \sup_{x \in \text{dom } f} (x\eta - \psi f(\cdot, x)) > \zeta \text{ a.s.} \quad (44)$$

Then there exists a $\hat{\xi} \in L^0$ such that

$$f(\cdot, \hat{\xi}) < \infty \text{ a.s. and } \hat{\xi}\eta - \psi(\cdot, \hat{\xi}) \geq \zeta \text{ a.s.} \quad (45)$$

Proof. This amounts to proving that the multifunction

$$S(\omega) := \{x \in \text{dom } f(\omega, \cdot) : x\eta(\omega) - \psi(\omega)f(\omega, x) \geq \zeta(\omega)\}$$

admits a measurable selection. S is nonempty valued by (44), and measurable since $g(\omega, x) := \psi(\omega)f(\omega, x) - x\eta(\omega)$ (with the convention $0 \cdot \infty = 0$) is a normal convex integrand (see [24, Prop. 14.44, Cor. 14.46]), and $S(\omega) = \text{dom } f(\omega, \cdot) \cap \{x : g(\omega, x) \leq -\zeta(\omega)\}$. On $\{\psi > 0\}$, we have simply $S = \{x : f(\cdot, x) - x\frac{\eta}{\psi} \leq -\frac{\zeta}{\psi}\}$ which is *closed* since f is normal. Thus

$$S'(\omega) = \begin{cases} S(\omega) & \text{if } \omega \in \{\psi > 0\}, \\ \emptyset & \text{if } \omega \in \{\psi = 0\}, \end{cases}$$

is a closed-valued measurable multifunction with $\text{dom } S' = \{\omega : S'(\omega) \neq \emptyset\} = \{\psi > 0\}$, thus the standard measurable selection theorem (see [24], Cor. 14.6) shows the existence of $\xi' \in L^0$ such that $\xi'(\omega) \in S'(\omega) = S(\omega)$ for $\omega \in \{\psi > 0\}$.

On $\{\psi = 0\}$, the multifunction S need not be closed-valued. So we explicitly construct a selector. First, on $\{\psi = \eta = 0\}$, we can take any $\xi_0 \in \text{dom } \mathcal{I}_{f,\gamma} (\neq \emptyset$ by assumption), which satisfies $\xi(\omega) \in \text{dom } f(\omega, \cdot)$ for a.e. ω by (13), and $\xi_0\eta - \psi f(\cdot, \xi_0) = 0 > \zeta$ on $\{\psi = \eta = 0\}$ since (44) reads as $\zeta < 0$ when $\psi = \eta = 0$.

Next, we put

$$\xi'' := \frac{1}{2} \left(\left(a_f^+ \wedge \frac{\zeta}{\eta} \right) + \left(a_f^- \vee \frac{\zeta}{\eta} \right) \right) \text{ on } \{\psi = 0, \eta \neq 0\},$$

where $a_{f*}^\pm(\omega) = \lim_{x \rightarrow \pm\infty} f^*(\omega, x)/x$ as in (15). Recalling that $\text{dom } f(\omega, \cdot) = \{a_{f*}^+(\omega)\}$ if $a_{f*}^-(\omega) = a_{f*}^+(\omega)$, and otherwise $(a_{f*}^-(\omega), a_{f*}^+(\omega)) \subset \text{dom } f(\omega, \cdot) \subset$

$[a_{f*}^-(\omega), a_{f*}^+(\omega)]$, we have $\xi''(\omega) \in \text{dom } f(\omega, \cdot)$ for $\omega \in \{\psi = 0, \eta \neq 0\}$. Also, (44) reads as $a_f^+\eta > \zeta$ when $\psi = 0$ and $\eta > 0$, hence

$$\xi''\eta - \psi f(\cdot, \xi'') = \xi''\eta = \frac{1}{2}\zeta + \frac{1}{2}(a_f^- \vee \zeta) \geq \zeta \quad \text{on } \{\psi = 0, \eta > 0\}.$$

Similarly, (44) reads as $a_f^-\eta > \zeta$ on $\{\psi = 0, \eta < 0\}$, hence

$$\xi''\eta = \frac{1}{2}((a_f^+\eta \vee \zeta) + (a_f^-\eta \wedge \zeta)) = \frac{1}{2}(a_f^+ \vee \zeta) + \frac{1}{2}\zeta \geq \zeta \quad \text{on } \{\psi = 0, \eta < 0\}.$$

Now $\hat{\xi} := \xi' \mathbf{1}_{\{\psi > 0\}} + \xi_0 \mathbf{1}_{\{\psi = 0, \eta = 0\}} + \frac{1}{2}\xi'' \mathbf{1}_{\{\psi = 0, \eta \neq 0\}}$ is a desired measurable selection. $\square$

Proof of Theorem 3.1. The upper bound is already established at the beginning of this section. For the lower bound, it suffices to show that for any $\nu = \nu_r + \nu_s \in ba$,

$$\alpha < \mathcal{H}_{f*, \gamma}(\nu_r), \beta < \sup_{\xi \in \mathcal{D}_{f, \gamma}} \nu_s(\xi) \Rightarrow \alpha + \beta < \mathcal{I}_{f, \gamma}^*(\nu). \quad (46)$$

In the sequel, we fix an arbitrary $\nu = \nu_r + \nu_s \in ba$ and α, β as in (46), and we denote

$$\eta := \frac{d\nu_r}{d\mathbb{P}} \quad \text{and} \quad \psi_Q := \frac{dQ}{d\mathbb{P}}, \quad \forall Q \in \mathcal{Q}_\gamma.$$

By the assumption on β , there exists a $\xi_s \in \mathcal{D}_{f, \gamma}$ with $\nu_s(\xi_s) > \beta$, and by the singularity of ν_s , there exists an increasing sequence $(A_n)_n$ in $\mathcal{F}$ with $\mathbb{P}(A_n) \uparrow 1$ and $|\nu_s|(A_n) = 0$ for all n , so that $\nu_s(\xi_s \mathbf{1}_{A_n^c}) = \nu_s(\xi_s) > \beta$. On the other hand, the assumption on α implies that

$$\alpha < \mathcal{H}_{f*}(\nu_r | Q) + \gamma(Q) = \mathbb{E} \left[\tilde{f}^*(\cdot, \eta, \psi_Q) \right] + \gamma(Q), \quad \forall Q \in \mathcal{Q}_\gamma.$$

Then for each $Q \in \mathcal{Q}_\gamma$, there exists a $\zeta_Q \in L^1$ with

$$\mathbb{E}[\zeta_Q] > \alpha - \gamma(Q) \quad \text{and} \quad \zeta_Q < \tilde{f}^*(\cdot, \eta, \psi_Q) \text{ a.s.}$$

(even if $\Phi := \tilde{f}^*(\cdot, \eta, \psi_Q) \notin L^1$: 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, so $(\Phi - \varepsilon) \wedge N_0$ with a big N_0 does the job.) Therefore, Lemma 4.2 implies that there exists a $\xi_Q^0 \in L^0$ such that

$$f(\cdot, \xi_Q^0) < \infty \quad \text{and} \quad \xi_Q^0 \eta - \psi_Q f(\cdot, \xi_Q^0) \geq \zeta_Q \text{ a.s.} \quad (47)$$

Note that this does not guarantee that ξ_Q^0 is in $\mathcal{D}_{f, \gamma}$ (if it was, there would be nothing to prove anymore). So we approximate ξ_Q^0 by elements of $\mathcal{D}_{f, \gamma}$. Let

$B_n := \{|\xi_Q^0| \leq n\} \cap \{|f(\cdot, \xi_Q^0)| \leq n\}$ and $C_n := A_n \cap B_n$, then $\mathbb{P}(C_n) \uparrow 1$ since $f(\cdot, \xi_Q^0) < \infty$ a.s. Put

$$\xi_Q^n := \xi_Q^0 \mathbf{1}_{C_n} + \xi_s \mathbf{1}_{C_n^c}, \quad \forall n \in \mathbb{N}.$$

Then for each n , $\xi_Q^n \in \mathcal{D}_{f,\gamma}$ since $\|\xi_Q^0\|_\infty \leq n + \|\xi_s\|_\infty < \infty$ and

$$f(\cdot, \xi_Q^n)^+ = f(\cdot, \xi_Q^0)^+ \mathbf{1}_{C_n} + f(\cdot, \xi_s)^+ \mathbf{1}_{C_n^c} \leq n + f(\cdot, \xi_s)^+ \in M_u^{\rho_\gamma}$$

by the solidness of the space. On the other hand,

$$\begin{aligned} \nu_r(\xi_Q^n) - \mathbb{E}_Q[f(\cdot, \xi_Q^n)] &= \mathbb{E}[\xi_Q^n \eta - \psi_Q f(\cdot, \xi_Q^n)] \\ &= \mathbb{E}[\mathbf{1}_{C_n}(\xi_Q^0 \eta - \psi_Q f(\cdot, \xi_Q^0))] + \mathbb{E}[\mathbf{1}_{C_n^c}(\xi_s \eta - \psi_Q f(\cdot, \xi_s))] \\ &\geq \mathbb{E}[\zeta_Q \mathbf{1}_{C_n}] + \mathbb{E}[\mathbf{1}_{C_n^c}(\xi_s \eta - \psi_Q f(\cdot, \xi_s))] \\ &= \mathbb{E}[\zeta_Q] + \mathbb{E}[\underbrace{\mathbf{1}_{C_n^c}(\xi_s \eta - \psi_Q f(\cdot, \xi_s) - \zeta_Q)}_{=:\Xi_Q}]. \end{aligned}$$

Note that $\Xi_Q \in L^1$ since $f(\cdot, \xi_s) \in \bigcap_{Q \in \mathcal{Q}_\gamma} L^1(Q)$ (by $\xi_s \in \mathcal{D}_{f,\gamma}$), thus $\lim_n \mathbb{E}[\mathbf{1}_{C_n^c} \Xi_Q] = 0$. Therefore, noting that $\nu_s(\xi_Q^n) = \nu_s(\mathbf{1}_{C_n^c} \xi_Q^n) = \nu_s(\mathbf{1}_{C_n^c} \xi_s) = \nu_s(\xi_s)$,

$$\begin{aligned} \sup_{\xi \in \mathcal{D}_{f,\gamma}} (\nu(\xi) - \mathbb{E}_Q[f(\cdot, \xi)]) &\geq \sup_n (\nu_r(\xi_Q^n) - \mathbb{E}_Q[f(\cdot, \xi_Q^n)] + \nu_s(\xi_Q^n)) \\ &\geq \limsup_n (\mathbb{E}[\zeta_Q] + \mathbb{E}[\mathbf{1}_{C_n^c} \Xi_Q] + \nu_s(\xi_s)) \\ &= \mathbb{E}[\zeta_Q] + \nu_s(\xi_s) > \alpha - \gamma(Q) + \beta. \end{aligned}$$

In view of (43), we have

$$\mathcal{I}_{f,\gamma}^*(\nu) \geq \inf_{Q \in \mathcal{Q}_\gamma} \sup_{\xi \in \mathcal{D}_{f,\gamma}} (\nu(\xi) - \mathbb{E}_Q[f(\cdot, \xi)] + \gamma(Q)) \geq \alpha + \beta.$$

Since $\alpha < \mathcal{H}_{f^*,\gamma}(\nu_r)$ and $\beta < \sup_{\xi \in \mathcal{D}_{f,\gamma}} \nu_s(\xi)$ are arbitrary, this completes the proof. $\square$

4.2. Proof of Theorem 3.4

In the sequel, the assumptions of Theorem 3.4 **excepting** (30) are supposed without notice. The implication (i) $\Rightarrow$ (ii) is trivial since $\mathbb{R} \subset L^\infty$, and (i) $\Rightarrow$ (iii) is clear from (23) of Theorem 3.1, while (ii) $\Rightarrow$ (i) follows from (cf. [22, 23])

$$a \leq \xi \leq b \text{ a.s., } a, b \in \mathbb{R} \Rightarrow f(\cdot, \xi) \leq f(\cdot, a)^+ + f(\cdot, b)^+. \quad (48)$$

Indeed, the assumption implies the existence of a $[0, 1]$ -valued random variable α such that $\xi = \alpha a + (1 - \alpha)b$ a.s. Since $f(\omega, \cdot)$ is convex for a.e. ω , we see that $f(\omega, \xi(\omega)) \leq \alpha(\omega)f(\omega, a) + (1 - \alpha(\omega))f(\omega, b) \leq f(\omega, a)^+ + f(\omega, b)^+$ for a.e. $\omega \in \Omega$.

Given $\text{dom}(\mathcal{I}_{f,\gamma}) = L^\infty$ and $\sigma(L^\infty, L^1)$ -lsc of $\mathcal{I}_{f,\gamma}$ (Lemma 2.3) as well as $\mathcal{I}_{f,\gamma}^*|_{L^1} = \mathcal{H}_{f^*,\gamma}$ (Corollary 3.2), $(iv) \Leftrightarrow (v)$ is a special case of the following (with $E = L^\infty$ and $E' = L^1$):

Lemma 4.3 ([15], Propositions 1 and 2). *Let $\langle E, E' \rangle$ be a dual pair and φ a $\sigma(E, E')$ -lsc finite convex function on E with the conjugate φ^* on E' . Then φ is $\tau(E, E')$ -continuous on E if and only if $\{x' \in E' : \varphi^*(x') \leq c\}$ is $\sigma(E', E)$ -compact for each $c \in \mathbb{R}$.*

Proof of $(iii) \Rightarrow (iv)$. Put $E = L^\infty$, $E' = ba$, then $\tau(L^\infty, ba)$ is the norm-topology while the finite lsc convex function $\mathcal{I}_{f,\gamma}$ on the Banach space L^∞ is norm-continuous (see [8, Ch. 1, Cor. 2.5]). Thus $\Lambda_c := \{\nu \in ba : \mathcal{I}_{f,\gamma}(\nu) \leq c\}$ is $\sigma(ba, L^\infty)$ -compact from Lemma 4.3, a fortiori $\{\eta \in L^1 : \mathcal{H}_{f^*,\gamma}(\eta) \leq c\} = \Lambda_c \cap L^1 = \Lambda_c$ (by Corollary 3.2 and (iii)) is $\sigma(L^1, L^\infty)$ -compact since $\sigma(L^1, L^\infty) = \sigma(ba, L^\infty)|_{L^1}$. $\square$

Proof of $(v) \Rightarrow (vi)$. This follows from the observation that

$$\sup_n \|\xi_n\|_\infty < \infty \text{ and } \xi_n \rightarrow \xi \text{ a.s.} \Rightarrow \xi_n \rightarrow \xi \text{ for } \tau(L^\infty, L^1).$$

Indeed, for any weakly compact ($\Rightarrow$ uniformly integrable) subset $\mathcal{C} \subset L^1$,

$$\begin{aligned} \sup_{\eta \in \mathcal{C}} \mathbb{E}[|\xi - \xi_n||\eta|] &\leq \sup_{\eta \in \mathcal{C}} \mathbb{E}[|\xi - \xi_n||\eta| \mathbf{1}_{\{|\eta| > N\}}] + \sup_{\eta \in \mathcal{C}} \mathbb{E}[|\xi - \xi_n||\eta| \mathbf{1}_{\{|\eta| \leq N\}}] \\ &\leq 2 \sup_n \|\xi_n\|_\infty \sup_{\eta \in \mathcal{C}} \mathbb{E}[\eta \mathbf{1}_{\{|\eta| > N\}}] + N \mathbb{E}[|\xi - \xi_n|]. \end{aligned}$$

Taking a diagonal, we see $q_{\mathcal{C}}(\xi - \xi_n) := \sup_{\eta \in \mathcal{C}} |\mathbb{E}[(\xi - \xi_n)\eta]| \rightarrow 0$, while $q_{\mathcal{C}}$ with $\mathcal{C}$ running through (convex, circled) weakly compact subsets of L^1 generates $\tau(L^\infty, L^1)$. $\square$

Proof of $(iv) \Rightarrow (vii)$. Fix an arbitrary $\eta_0 \in \text{dom}(\mathcal{H}_{f^*,\gamma})$ ($\neq \emptyset$) and $\xi \in L^\infty$. Observe that

$$\mathbb{E}[\xi(\eta + \eta_0)] = \mathbb{E}\left[2\xi \frac{\eta + \eta_0}{2}\right] \leq \mathcal{I}_{f,\gamma}(2\xi) + \frac{1}{2}(\mathcal{H}_{f^*,\gamma}(\eta) + \mathcal{H}_{f^*,\gamma}(\eta_0)), \quad \forall \eta \in L^1.$$

Hence $\mathbb{E}[\xi\eta] - \mathcal{H}_{f^*,\gamma}(\eta) \leq \mathcal{I}_{f,\gamma}(2\xi) + \|\xi\|_\infty \|\eta_0\|_1 + \frac{1}{2}\mathcal{H}_{f^*,\gamma}(\eta_0) - \frac{1}{2}\mathcal{H}_{f^*,\gamma}(\eta)$. Putting $C_{c,\xi} := 2(c + \mathcal{I}_{f,\gamma}(2\xi) + \|\xi\|_\infty \|\eta_0\|_1) + \mathcal{H}_{f^*,\gamma}(\eta_0)$ which does not depend on η , we see that

$$\mathbb{E}[\xi\eta] - \mathcal{H}_{f^*,\gamma}(\eta) \geq c \Rightarrow \mathcal{H}_{f^*,\gamma}(\eta) \leq C_{c,\xi}.$$

Consequently, $\{\eta \in L^1 : \mathbb{E}[\xi\eta] - \mathcal{H}_{f^*,\gamma}(\eta) \geq c\}$ is weakly compact for each $c > 0$, since it is contained in a weakly compact set $\{\eta \in L^1 : \mathcal{H}_{f^*,\gamma}(\eta) \leq C_{c,\xi}\}$ and $\eta \mapsto \mathbb{E}[\xi\eta] - \mathcal{H}_{f^*,\gamma}(\eta)$ is weakly upper semicontinuous. Therefore, $\sup_{\eta \in L^1} (\mathbb{E}[\xi\eta] - \mathcal{H}_{f^*,\gamma}(\eta))$ is attained. $\square$

From now on, we assume (30) (thus all the assumptions of Theorem 3.4) which implies that

$$f(\cdot, \xi)^+ \in M^{\rho_\gamma}, \quad \forall \xi \in L^\infty. \quad (49)$$

Indeed, by $\text{dom}(\mathcal{I}_{f,\gamma}) = L^\infty$ and $\exists \xi_0 \in \mathcal{D}_{f,\gamma}$, we have for all $\xi \in L^\infty$ that

$$\begin{aligned} \rho_\gamma(\lambda f(\cdot, \xi)) &\leq \frac{1}{2} \rho_\gamma(f(\cdot, 2\lambda\xi - (2\lambda - 1)\xi_0)) \\ &\quad + \frac{1}{2} \rho_\gamma((2\lambda - 1)f(\cdot, \xi_0)^+) < \infty, \quad \lambda > 1. \end{aligned}$$

Here the first term in the right hand side is $\frac{1}{2}\mathcal{I}_{f,\gamma}(2\lambda\xi - (2\lambda - 1)\xi_0) < \infty$. On the other hand, if $f(\cdot, \xi_0)^- \in M^{\rho_\gamma}$ as in (30), putting $A = \{f(\cdot, \xi) \geq 0\}$ with an arbitrary $\xi \in L^\infty$, we see that $f(\cdot, \xi \mathbf{1}_A + \xi_0' \mathbf{1}_{A^c}) = f(\cdot, \xi)^+ + f(\cdot, \xi_0') \mathbf{1}_{A^c}$, hence $f(\cdot, \xi)^+ \leq f(\cdot, \xi \mathbf{1}_A + \xi_0' \mathbf{1}_{A^c}) + f(\cdot, \xi_0')^-$. Therefore, for all $\lambda > 1$,

$$\rho_\gamma(\lambda f(\cdot, \xi)^+) \leq \frac{1}{2} \rho_\gamma(2\lambda f(\cdot, \xi \mathbf{1}_A + \xi_0' \mathbf{1}_{A^c})) + \frac{1}{2} \rho_\gamma(2\lambda f(\cdot, \xi_0')^-).$$

Proof of (vi) $\Rightarrow$ (i). In view of (7) and (49), it suffices to show that for each $\xi \in L^\infty$,

$$\lim_N \sup_{\gamma(Q) \leq 1} \mathbb{E}_Q[f(\cdot, \xi)^+ \mathbf{1}_{\{f(\cdot, \xi)^+ \geq N\}}] = 0, \quad \forall c > 0.$$

Let $A_N := \{f(\cdot, \xi) \geq N\} \in \mathcal{F}$ ($N \in \mathbb{N}$), then $\mathbb{P}(A_N) \rightarrow 0$ since $\mathcal{I}_{f,\gamma}$ is finite. Pick a $\xi_0 \in \mathcal{D}_{f,\gamma}$ and put $\xi_N := (\xi - \xi_0) \mathbf{1}_{A_N}$ so that $\xi_N + \xi_0 = \xi \mathbf{1}_{A_N} + \xi_0 \mathbf{1}_{A_N^c}$. Then $f(\cdot, \xi_N + \xi_0) = f(\cdot, \xi) \mathbf{1}_{A_N} + f(\cdot, \xi_0) \mathbf{1}_{A_N^c}$ while for $\lambda > 1$

$$f(\cdot, \xi_N + \xi_0) \leq \frac{1}{\lambda} f(\cdot, \lambda \xi_N + \xi_0) + \frac{\lambda - 1}{\lambda} f(\cdot, \xi_0),$$

hence

$$f(\cdot, \xi)^+ \mathbf{1}_{A_N} = f(\cdot, \xi) \mathbf{1}_{A_N} \leq \frac{1}{\lambda} f(\cdot, \lambda \xi_N + \xi_0) + \frac{1}{\lambda} f(\cdot, \xi_0)^- + f(\cdot, \xi_0)^+ \mathbf{1}_{A_N}.$$

Note that since $\xi_0 \in \mathcal{D}_{f,\gamma}$ and since $f(\cdot, \xi_0)^- \in L^{\rho_\gamma}$,

$$C_1 := \sup_{\gamma(Q) \leq c} \mathbb{E}_Q[f(\cdot, \xi_0)^-] < \infty \quad \text{and} \quad \lim_N \sup_{\gamma(Q) \leq c} \mathbb{E}_Q[f(\cdot, \xi_0)^+ \mathbf{1}_{A_N}] = 0.$$

Also, $\mathbb{E}_Q[f(\cdot, \lambda \xi_N + \xi_0)] \leq \rho_\gamma(f(\cdot, \lambda \xi_N + \xi_0)) + \gamma(Q) = \mathcal{I}_{f,\gamma}(\lambda \xi_N + \xi_0) + \gamma(Q)$, while for each $\lambda > 1$, we have $\sup_N \|\lambda \xi_N + \xi_0\|_\infty \leq \lambda \|\xi\|_\infty + (\lambda - 1)\|\xi_0\|_\infty < \infty$ and $\lambda \xi_N + \xi_0 \rightarrow \xi_0$ a.s., thus $\mathcal{I}_{f,\gamma}(\lambda \xi_N + \xi_0) \rightarrow \mathcal{I}_{f,\gamma}(\xi_0)$ by (vi). Hence for large enough $N > N_\lambda$ (say), $\mathcal{I}_{f,\gamma}(\lambda \xi_N + \xi_0) \leq \mathcal{I}_{f,\gamma}(\xi_0) + 1 =: C_2$, so

$$\sup_{\gamma(Q) \leq c} \mathbb{E}_Q[f(\cdot, \xi)^+ \mathbf{1}_{A_N}] \leq \frac{C_2 + c + C_1}{\lambda} + \sup_{\gamma(Q) \leq c} \mathbb{E}_Q[f(\cdot, \xi_0)^+ \mathbf{1}_{A_N}].$$

Now a diagonal argument yields that $\sup_{\gamma(Q) \leq c} \mathbb{E}_Q[f(\cdot, \xi)^+ \mathbf{1}_{A_N}] \rightarrow 0$. $\square$

Proof of (vii) $\Rightarrow$ (iv). Under (49), $\lim_{\|\eta\|_1 \rightarrow \infty} \mathcal{H}_{f^*,\gamma}(\eta)/\|\eta\|_1 = \infty$, i.e., $\mathcal{H}_{f^*,\gamma}$ is coercive on L^1 . For, since $\|\eta\|_1 = \mathbb{E}[\text{sgn}(\eta)\eta]$ where $\text{sgn}(\eta) = \mathbb{1}_{\{\eta \geq 0\}} - \mathbb{1}_{\{\eta < 0\}} \in L^\infty$, $f(\cdot, n \text{sgn}(\eta)) \leq f(\cdot, -n)^+ + f(\cdot, n)^+ \in M^{\rho_\gamma}$ by (48), (49), and $\mathcal{H}_{f^*,\gamma}(\eta) = \sup_{\xi \in L^\infty} (\mathbb{E}[\xi\eta] - \mathcal{I}_{f,\gamma}(\xi))$, we have

$$\begin{aligned} \mathcal{H}_{f^*,\gamma}(\eta) &\geq \mathbb{E}[n \text{sgn}(\eta)\eta] - \mathcal{I}_{f,\gamma}(n \text{sgn}(\eta)) \\ &\geq n\|\eta\|_1 - \frac{\rho_\gamma(2f(\cdot, -n)^+) - \rho_\gamma(2f(\cdot, n)^+)}{2} > -\infty, \quad \forall n. \end{aligned}$$

Then the *coercive James' theorem* [16, Th. 2] applied to the coercive function $\mathcal{H}_{f^*,\gamma}$ on the Banach space $E = L^1$ implies the relative weak compactness of all the sublevel sets $\{\eta \in L^1 : \mathcal{H}_{f^*,\gamma}(\eta) \leq c\}$ which are weakly closed since $\mathcal{H}_{f^*,\gamma}$ is weakly lower semicontinuous. $\square$

A. Lower semicontinuity of $\mathcal{H}_{f^*} + \gamma$

Lemma A.1. *Under the assumptions of Theorem 3.1, $\mathcal{H}_{f^*} + \gamma$ is jointly weakly lower semicontinuous on $L^1 \times \mathcal{Q}_\gamma$, and $\inf_{Q \in \mathcal{Q}_\gamma} (\mathcal{H}_{f^*}(\eta|Q) + \gamma(Q))$ is attained for every $\eta \in L^1$.*

Proof. Let ξ_0 be as in (22) (so $f(\cdot, \xi_0)^+ \in M_u^{\rho_\gamma}$), $\psi_Q = dQ/d\mathbb{P}$ for each $Q \in \mathcal{Q}_\gamma$, and put $K(c, a) := 2(c + \rho_\gamma(2f(\cdot, \xi_0)^+) + a\|\xi_0\|_\infty) < \infty$. Then by (42) in the proof of Lemma 4.1,

$$\|\eta\|_1 \leq a \quad \text{and} \quad \mathcal{H}_{f^*}(\nu|Q) + \gamma(Q) \leq c \Rightarrow \gamma(Q) \leq K(c, a). \quad (50)$$

Also, from (19) with ξ_0 above and (7), we see that whenever $(\eta_n)_n$ is uniformly integrable and $\sup_n \gamma(Q_n) < \infty$, $\{\tilde{f}^*(\cdot, \eta_n, \psi_{Q_n})^-\}_n$ is uniformly integrable.

To see that $\mathcal{H}_{f^*} + \gamma$ is weakly lsc, it suffices by convexity that for each $c \in \mathbb{R}$, $\Lambda_c := \{(\eta, \psi_Q) \in L^1 \times L^1 : Q \in \mathcal{Q}_\gamma, \mathcal{H}_{f^*}(\eta|Q) + \gamma(Q) \leq c\}$ is norm-closed. Let $(\eta_n, \psi_{Q_n})_n \subset \Lambda_c$ be norm convergent to (η, ψ_Q) . Passing to a subsequence, we can assume a.s. convergence too. By the above paragraph and the norm convergence, $(\eta_n)_n$ and $\{\tilde{f}^*(\cdot, \eta_n, \psi_{Q_n})^-\}_n$ are uniformly integrable, while $\gamma(Q) \leq \sup_n \gamma(Q_n) < \infty$ by (50) and lsc of γ . Thus by Fatou's lemma,

$$\begin{aligned} \mathcal{H}_{f^*}(\eta|Q) + \gamma(Q) &= \mathbb{E}[\tilde{f}^*(\eta, \psi_Q)] + \gamma(Q) \\ &\leq \liminf_n \mathbb{E}[\tilde{f}^*(\eta_n, \psi_{Q_n})] + \liminf_n \gamma(Q_n) \\ &\leq \liminf_n \left(\mathbb{E}[\tilde{f}^*(\eta_n, \psi_{Q_n})] + \gamma(Q_n) \right) \leq c, \end{aligned}$$

hence $(\eta, \psi_Q) \in \Lambda_c$, obtaining the lower semicontinuity of $\mathcal{H}_{f^*} + \gamma$. In particular, $\mathcal{H}_{f^*}(\eta|\cdot) + \gamma(\cdot)$ is weakly lower semicontinuous on $\mathcal{Q}_\gamma$, so another application of (50) as well as (6) imply that the infimum $\inf_{Q \in \mathcal{Q}_\gamma} (\mathcal{H}_{f^*}(\eta|Q) + \gamma(Q))$ is attained for each $\eta \in L^1$. $\square$

B. Some Details of Example 3.3

Let $(\mathbb{N}, 2^{\mathbb{N}}, \mathbb{P})$ and $(P_n)_n$ be as in Example 3.3 with $\mathcal{P} = \overline{\text{conv}}(P_n, n \in \mathbb{N})$ (closed convex hull). The weak compactness of $\mathcal{P}$ follows from $\sup_n P_n(\{k, k+1, k+2, \dots\}) = \sup\{1/n : n \geq k\} = 1/k \rightarrow 0$ as $k \rightarrow \infty$. Also $\rho_\gamma(\xi) = \sup_{P \in \mathcal{P}} \mathbb{E}_P[\xi] = \sup_n \mathbb{E}_{P_n}[\xi]$ if $\xi \geq 0$ since $P \mapsto \mathbb{E}_P[\xi \wedge N]$ is continuous for any $N \in \mathbb{N}$, so $\sup_{P \in \mathcal{P}} \mathbb{E}_P[\xi] = \sup_N \sup_{P \in \mathcal{P}} \mathbb{E}_P[\xi \wedge N] = \sup_N \sup_{P \in \text{conv}(P_n; n \in \mathbb{N})} \mathbb{E}_P[\xi \wedge N]$, while if $P = \alpha_1 P_{n_1} + \dots + \alpha_l P_{n_l}$, then $\mathbb{E}_P[\xi \wedge N] = \alpha_1 \mathbb{E}_{P_{n_1}}[\xi \wedge N] + \dots + \alpha_l \mathbb{E}_{P_{n_l}}[\xi \wedge N] \leq \max_{1 \leq i \leq l} \mathbb{E}_{P_{n_i}}[\xi \wedge N] \leq \sup_n \mathbb{E}_{P_n}[\xi \wedge N]$.

Lemma B.1. *Let f be given by (27) in Example 3.3. Then we have*

$$\lim_{N \rightarrow \infty} \sup_n \mathbb{E}_{P_n}[f(\cdot, \xi) \mathbf{1}_{\{f(\cdot, \xi) \geq N\}}] = \limsup_n \xi(n)^+ e^{\xi(n)^+}. \quad (51)$$

Proof. Let $h(x) := x^+ e^x$ and fix $\xi \in \ell^\infty$. If $\|\xi\|_\infty = 0$, then both sides of (51) are 0, thus we assume $\|\xi\|_\infty > 0$ ($\Leftrightarrow h(\|\xi\|_\infty) > 0$). Note that for any $N, n \in \mathbb{N}$, we have (by definition)

$$\mathbb{E}_{P_n}[f(\cdot, \xi) \mathbf{1}_{\{f(\cdot, \xi) \geq N\}}] = \left(1 - \frac{1}{n}\right) h(\xi(1)) \mathbf{1}_{\{h(\xi(1)) \geq N\}} + h(\xi(n)) \mathbf{1}_{\{nh(\xi(n)) \geq N\}}.$$

In particular, $\mathbb{E}_{P_n}[f(\cdot, \xi) \mathbf{1}_{\{f(\cdot, \xi) \geq N\}}] \leq h(\xi(1)) \mathbf{1}_{\{h(\xi(1)) \geq N\}} + h(\xi(n)) \mathbf{1}_{\{nh(\xi(n)) \geq N\}}$, thus

$$\lim_{N \rightarrow \infty} \sup_n \mathbb{E}_{P_n}[f(\cdot, \xi) \mathbf{1}_{\{f(\cdot, \xi) \geq N\}}] \leq \lim_{N \rightarrow \infty} \sup_{n \geq N/h(\|\xi\|_\infty)} h(\xi(n)) = \limsup_n h(\xi(n)) =: \alpha.$$

This is “ $\leq$ ” in (51). If $\alpha = 0$, we are done; otherwise, for any $\varepsilon > 0$ and $N \in \mathbb{N}$, there exists some $n_N^\varepsilon > N/(\alpha - \varepsilon) > 0$ with $h(\xi(n_N^\varepsilon)) > \alpha - \varepsilon$. In particular, $n_N^\varepsilon h(\xi(n_N^\varepsilon)) > N$, hence

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

This proves “ $\geq$ ” in (51). □

Since $h(x) = x^+ e^x$ is increasing, continuous and $h(0) = 0$, $\limsup_n h(\xi(n)) = 0 \Leftrightarrow \limsup_n \xi(n) = 0$, hence (28). Recall that $\text{dom}(\mathcal{I}_{f,\gamma}^*) \subset ba_+$ (since $\mathcal{I}_{f,\gamma}$ is increasing). Finally,

Lemma B.2. *The conjugate $\mathcal{I}_{f,\gamma}^*$ is explicitly given on ba_+^s as:*

$$\mathcal{I}_{f,\gamma}^*(\nu) = \sup_{x \geq 0} x(\|\nu\| - e^x), \quad \forall \nu \in ba_+^s. \quad (52)$$

Proof. Let $\nu \in ba_+^s$. Since $\mathbb{E}_{P_n}[f(\cdot, \xi)] = \left(1 - \frac{1}{n}\right) h(\xi(1)) + h(\xi(n))$, we have

$$h(\xi(n)) \stackrel{(*)}{\leq} \mathbb{E}_{P_n}[f(\cdot, \xi)] \stackrel{(**)}{\leq} h(\xi(1)) + h(\xi(n)).$$

From (*) and (26), $\nu(\xi) - \mathcal{I}_f(\xi) \leq \|\nu\| \limsup_n \xi(n) - \sup_n h(\xi(n)) \leq \|\xi^+\|_\infty (\|\nu\| - e^{\|\xi^+\|_\infty}) \leq \sup_{x \geq 0} x(\|\nu\| - e^x)$ which shows “ $\leq$ ” in (52). Considering $\bar{x}^0 := (0, x, x, \dots) \in \ell^\infty$ with $x \geq 0$ (then $\|\bar{x}^0\|_\infty = x$ and $\nu(\bar{x}^0) = x\|\nu\|$ since ν vanishes on finite sets), we deduce from (*) that

$$\mathcal{I}_{f,\gamma}^*(\nu) \geq \sup_{\xi \in \ell^\infty} (\nu(\xi) - h(\xi(1)) - h(\|\xi\|_\infty)) \geq \sup_{x \geq 0} (x\|\nu\| - xe^x). \quad \square$$

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

References

- [1] C. D. Aliprantis, K. C. Border: Infinite Dimensional Analysis: A Hitchhiker’s Guide, 3rd Ed., Springer, Berlin (2006).
- [2] F. Bellini, M. Frittelli: On the existence of minimax martingale measures, *Math. Finance* 12 (2002) 1–21.
- [3] S. Biagini, M. Frittelli, M. Grasselli: Indifference price with general semimartingales, *Math. Finance* 21 (2011) 423–446.
- [4] C. Castaing, M. Valadier: Convex Analysis and Measurable Multifunctions, *Lecture Notes in Math.* 580, Springer, Berlin (1977).
- [5] F. Delbaen: Differentiability properties of utility functions, in: *Optimality and Risk – Modern Trends in Mathematical Finance*, F. Delbaen et al. (ed.), Springer, Berlin (2009) 39–48.
- [6] C. Dellacherie, P.-A. Meyer: Probabilities and Potential, *North-Holland Mathematics Studies* 29, North-Holland, Amsterdam (1978).
- [7] N. Dunford, J. T. Schwartz: Linear Operators. Part I: General Theory, *Wiley Classics Library*, John Wiley & Sons, New York (1988).
- [8] I. Ekeland, R. Temam: Convex Analysis and Variational Problems, *Studies in Mathematics and its Applications* 1, North-Holland, Amsterdam (1976).
- [9] H. Föllmer, A. Gundel: Robust projections in the class of martingale measures, *Ill. J. Math.* 50 (2006) 439–472.
- [10] H. Föllmer, A. Schied, S. Weber: Robust preferences and robust portfolio choice, in: *Mathematical Modelling and Numerical Methods in Finance*, A. Bensoussan, Q. Zhang, P. G. Ciarlet (eds.), *Handbook of Numerical Analysis* 15, North-Holland, Amsterdam (2009) 29–88.
- [11] A. Grothendieck: Topological Vector Spaces, *Notes on Mathematics and its Applications* Gordon and Breach, New York (1973).
- [12] E. Hewitt, K. Stromberg: Real and Abstract Analysis: A Modern Treatment of the Theory of Functions of a Real Variable, *Graduate Texts in Mathematics* 25, Springer, New York (1975).

- [13] E. Jouini, W. Schachermayer, N. Touzi: Law invariant risk measures have the Fatou property, in: *Advances in Mathematical Economics* 9, S. Kusuoka et al. (ed.), Springer, Tokyo (2006) 49–71.
- [14] A. Kozek: Convex integral functionals on Orlicz spaces, *Commentat. Math.* 21 (1980) 109–135.
- [15] J.-J. Moreau: Sur la fonction polaire d’une fonction semi-continue supérieurement, *C. R. Acad. Sci., Paris* 258 (1964) 1128–1130.
- [16] J. Orihuela, M. Ruiz Galán: A coercive James’s weak compactness theorem and nonlinear variational problems, *Nonlinear Anal.* 75 (2012) 598–611.
- [17] J. Orihuela, M. Ruiz Galán: Lebesgue property for convex risk measures on Orlicz spaces, *Math. Financ. Econ.* 6 (2012) 15–35.
- [18] K. Owari: On admissible strategies in robust utility maximization, *Math. Financ. Econ.* 6 (2012) 77–92.
- [19] K. Owari: Maximum Lebesgue extension of monotone convex functions, *J. Funct. Anal.* 266 (2014) 3572–3611.
- [20] K. Owari: On the Lebesgue property of monotone convex functions, *Math. Financ. Econ.* 8 (2014) 159–167.
- [21] R. T. Rockafellar: Extension of Fenchel’s duality theorem for convex functions, *Duke Math. J.* 33 (1966) 81–89.
- [22] R. T. Rockafellar: Integrals which are convex functionals. II, *Pacific J. Math.* 39 (1971) 439–469.
- [23] R. T. Rockafellar: Integral functionals, normal integrands and measurable selections, in: *Nonlinear Operators and the Calculus of Variations* (Brussels, 1975), J. P. Gossez et al. (ed.), *Lecture Notes in Math.* 543, Springer, Berlin (1976) 157–207.
- [24] R. T. Rockafellar, R. J.-B. Wets: *Variational Analysis*, *Grundlehren der Mathematischen Wissenschaften* 317, Springer, Berlin (1998).
- [25] W. Schachermayer: A super-martingale property of the optimal portfolio process, *Finance Stoch.* 7 (2003) 433–456.
- [26] A. Schied: Optimal investments for risk- and ambiguity-averse preferences: a duality approach, *Finance Stoch.* 11 (2007) 107–129.
- [27] W. Wittmüss: Robust optimization of consumption with random endowment, *Stochastics* 80 (2008) 459–475.
- [28] C. Zălinescu: *Convex Analysis in General Vector Spaces*, World Scientific, River Edge (2002).