An Analytical Study of Norms and Banach Spaces Induced by the Entropic Value-at-Risk

Amir Ahmadi-Javid*   Alois Pichler†

July 28, 2017

Abstract

This paper addresses the Entropic Value-at-Risk (EV@R), a recently introduced coherent risk measure. It is demonstrated that the norms induced by EV@R induce the same Banach spaces, irrespective of the confidence level. Three spaces, called the primal, dual, and bidual entropic spaces, corresponding with EV@R are fully studied. It is shown that these spaces equipped with the norms induced by EV@R are Banach spaces. The entropic spaces are then related to the L^p spaces, as well as specific Orlicz hearts and Orlicz spaces. This analysis indicates that the primal and bidual entropic spaces can be used as very flexible model spaces, larger than L^∞, over which all L^p-based risk measures are well-defined.

The dual EV@R norm and corresponding Hahn–Banach functionals are presented explicitly, which are not explicitly known for the Orlicz and Luxemburg norms that are equivalent to the EV@R norm. The duality relationships among the entropic spaces are investigated. The duality results are also used to develop an extended Donsker–Varadhan variational formula, and to explicitly provide the dual and Kusuoka representations of EV@R, as well as the corresponding maximizing densities in both representations.

Our results indicate that financial concepts can be successfully used to develop insightful tools for not only the theory of modern risk measurement but also other fields of stochastic analysis and modeling.

Keywords: Coherent Risk Measures, Dual Representation, Kusuoka Representation, Orlicz Hearts and Spaces, Orlicz and Luxemburg Norms, L^p Spaces, Moment and Cumulant Generating Functions, Donsker–Varadhan Variational Formula, Large Deviations, Relative Entropy, Kullback-Leibler Divergence

Classification: 90C15, 60B05, 62P05

*Industrial Engineering Department, Amirkabir University of Technology, Tehran, Iran
Contact: ahmadi_javid@aut.ac.ir

†Technische Universität Chemnitz, Fakultät für Mathematik, Chemnitz, Germany
Contact: alois.pichler@mathematik.tu-chemnitz.de, https://www.tu-chemnitz.de/mathematik/fima/

1 Introduction

The literature of risk measures has been extensively developed over the past two decades. A breakthrough was achieved when the seminal paper Artzner et al. (1999) used an axiomatic approach to define coherent risk measures.

We follow this axiomatic setting and consider a vector space S of $\mathbb{R}$-valued random variables on a reference probability space $(\Omega, \mathcal{F}, P)$. The set S is called the *model space*, which is used in this paper to represent a set of random losses. A risk measure $\rho : S \to \overline{\mathbb{R}}$ with $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ is called *coherent* if it satisfies the following four properties:

(P1) Translation invariance: $\rho(Y + c) = \rho(Y) + c$ for any $Y \in S$ and $c \in \mathbb{R}$
(P2) Subadditivity: $\rho(Y_1 + Y_2) \leq \rho(Y_1) + \rho(Y_2)$ for all $Y_1, Y_2 \in S$
(P3) Monotonicity: If $Y_1, Y_2 \in S$ and $Y_1 \leq Y_2$, then $\rho(Y_1) \leq \rho(Y_2)$
(P4) Positive homogeneity: $\rho(\lambda Y) = \lambda\rho(Y)$ for all $Y \in S$ and $\lambda > 0$

where the convention $(+\infty) + (-\infty) = (-\infty) + (+\infty) = +\infty$ is considered. The above setting is usually used in the operations research and insurance literature, while in the finance context, which mostly deals with monetary gains instead of losses, a functional $\varphi(\cdot)$ is called coherent if $\rho(Y) = \varphi(-Y)$ satisfies the above four properties.

This paper addresses the Entropic Value-at-Risk ($\mathsf{EV@R}$), a coherent risk measure introduced in Ahmadi-Javid (2011) and Ahmadi-Javid (2012a) (see Definition 2.1, and Remarks 2.7 and 4.11 for financial interpretations). This risk measure is the tightest upper bound that one can obtain from Chernoff's inequality for the Value-at-Risk ($\mathsf{V@R}$). In Ahmadi-Javid (2012a) and Ahmadi-Javid (2012b), it is shown that the dual representation of the $\mathsf{EV@R}$ can be given using the relative entropy (also known as Kullback-Leibler divergence) for bounded random variables and those random variables whose moment generating functions exist everywhere, respectively. The Kusuoka representation of the $\mathsf{EV@R}$ was recently obtained by Delbaen (2018) for bounded random variables.

Ahmadi-Javid (2011) and Ahmadi-Javid (2012a) also extended the $\mathsf{EV@R}$ to the class of φ-entropic risk measures (see (55)) by replacing the relative entropy in its dual representation with generalized entropies (also known as φ-divergences). This large class of risk measures includes the classical Average Value-at-Risk ($\mathsf{AV@R}$, also known as Conditional Value-at-Risk or expected shortfall). φ-entropic risk measures were examined further by Ahmadi-Javid (2012c), Breuer and Csiszár (2013), Breuer and Csiszár (2016), and Shapiro (2017) in developing various decision-making preferences with ambiguity aversion, measuring distribution model risks, and investigating distributionally robust counterparts.

There is a close relationship between norms and coherent risk measures defined on the same model space. Indeed, a coherent risk measure ρ can be used to define an order-preserving semi-norm on the model space S by

$$\|Y\|_\rho := \rho(|Y|),$$

whenever ρ is finite over S (see Pichler (2014); Pichler (2017) elaborate the reverse statement, as well as comparisons with other risk measures and norms).

This paper starts by investigating the norm induced by the EV@R, called the EV@R norm. The EV@R norms at the confidence levels $\alpha = 0$ and $\alpha = 1$ are identical to the L^1 and L^∞ norms, respectively. The EV@R norms at different confidence levels $0 < \alpha < 1$ are proven to be equivalent to each other, but they do not generate the L^p norms; the L^p norms are bounded by the EV@R norm while the converse does not hold true. The largest model space contained in the domain of EV@R is therefore strictly larger than L^∞, but smaller than every L^p space for $1 \leq p < \infty$ (see Kaina and Rüschendorf (2009) for a study of risk measures on L^p spaces).

One should recall that the natural topology defined on $\bigcap_{p\geq 1} L^p$, i.e., the smallest topology which contains each relative norm topology, cannot be defined by any norm (Arens (1946) and Bell (1977)). Therefore, instead of L^∞, one can equip the largest model space corresponding to EV@R with the norm induced by EV@R and include unbounded random variables in consideration. The new space is smaller than every L^p space, but includes all bounded random variables and those unbounded random variables whose moment generating functions exist around zero, such as those with normal or exponential distributions. This model space is sufficiently rich for most situations in practice.

The paper presents explicit expressions for the dual EV@R norm and corresponding Hahn–Banach functionals. This is an important achievement because there are no explicit formulas for the dual norm and Hahn–Banach functionals associated with the Orlicz and Luxemburg norms that are equivalent to the EV@R norm.

As side advantages of the duality analysis, we explicitly derive the dual and Kusuoka representations of EV@R, as well as their corresponding maximizing densities, for random variables whose moment generating functions are finite only around zero. Moreover, we prove an extended version of the Donsker–Varadhan variational formula and its dual.

It is shown that the EV@R norm, or its dual norm, is indeed equivalent to the Orlicz (or Luxemburg) norm on the associated Orlicz space. Thus, for the entropic spaces one has the general topological results that were developed and intensively studied for risk measures on Orlicz hearts and Orlicz spaces in the papers Biagini and Frittelli (2008), Cheridito and Li (2008), Cheridito and Li (2009), Svindland (2009), and Kupper and Svindland (2011); and more recently in Kiesel and Rüschendorf (2014), Farkas et al. (2015), and Delbaen and Owari (2016). The relation to Orlicz spaces further leads to results on risk measures based on quantiles, as addressed in Bellini and Rosazza Gianin (2012) and Bellini, Klar, et al. (2014). However, despite this equivalence to the Orlicz and Luxamburg norms, the EV@R norm and its dual norm seem favoured in practice, because they are both computationally tractable and their corresponding Hahn–Banach functionals are available explicitly.

Outline of the paper. Section 2 first introduces the EV@R, entropic spaces, and the associated norms; and then provides some basic results. Section 3 compares the entropic spaces with other spaces, particularly to the L^p and Orlicz spaces. Section 4 elaborates duality relations, and provides the dual norm and Hahn–Banach functionals explicitly. Section 5 concludes the paper by providing a summary and an outlook.

2 Entropic Value-at-Risk and entropic spaces

Throughout this paper, we consider the probability space $(\Omega, \mathcal{F}, P)$. To introduce the *Entropic Value-at-Risk* for an $\mathbb{R}$-valued random variable $Y \in L^0(\Omega, \mathcal{F}, P)$, we consider the moment-generating function $m_Y(t) := \mathbb{E}\, e^{tY}$ (cf. Laplace transform).

Definition 2.1 (Ahmadi-Javid (2011))**.** For an $\mathbb{R}$-valued random variable $Y \in L^0(\Omega, \mathcal{F}, P)$ with finite $m_Y(t_0)$ for some $t_0 > 0$, the *Entropic Value-at-Risk* ($\mathsf{EV@R}$) at confidence level $\alpha \in [0, 1]$ (or risk level $1 - \alpha$) is given by

$$\mathsf{EV@R}_\alpha(Y) := \inf_{t>0} \frac{1}{t} \log \frac{1}{1-\alpha} m_Y(t) \tag{1}$$

for $\alpha \in [0, 1)$, and is given by

$$\mathsf{EV@R}_1(Y) := \operatorname{ess\,sup}(Y) \tag{2}$$

for $\alpha = 1$. Further, the vector spaces

$$E := \left\{ Y \in L^0 : \mathbb{E}\, e^{t|Y|} < \infty \text{ for all } t > 0 \right\} \tag{3}$$

$$E^* := \left\{ Y \in L^0 : \mathbb{E}\, |Y| \log^+ |Y| < \infty \right\} \text{ [1]}$$

and

$$E^{**} := \left\{ Y \in L^0 : \mathbb{E}\, e^{t|Y|} < \infty \text{ for some } t > 0 \right\}$$

are called the *primal*, *dual*, and *bidual entropic spaces*, respectively.

Remark 2.2 (Entropic spaces). The names chosen here for the spaces E, E^*, and E^{**} are used to simplify working with these spaces. Moreover, these names clearly reveal the duality relationships among these spaces, that is, E^* and E^{**} are the norm dual and bidual spaces of E (see Section 4). The entropic spaces are also equivalent to specific Orlicz hearts and spaces (see Section 3.2). Some references denote the spaces E^* and E^{**} by $L \log L$ and $L_{\exp}$, which are related to the Zygmund and Lorentz spaces (see e.g., Section 4.6 of Bennett and Sharpley (1988), and Section 6.7 of Castillo and Rafeiro (2016)).

Remark 2.3 (Relation to insurance). $\mathsf{EV@R}$ generalizes the *exponential premium principle*, which is defined as

$$Y \mapsto \frac{1}{t} \log \mathbb{E}\, e^{tY} \tag{4}$$

for $t > 0$ fixed, see Kaas et al. (2008). The parameter t in (4) is also called the *risk aversion* parameter. Then, the objective function in (1) can be viewed as a *homogenization* of an affine perturbation of the exponential premium principle, considered in Shapiro et al. (2014, Section 6.3.2).

[1] $\log^+ z := \max\{0, \log z\}$

Remark 2.4. The entropic spaces E and E^{**} are defined in a related way. Distinguishing these two spaces turns out to be fundamental and essential in what follows. For example, it is possible to approximate random variables $Y \in E$ by simple functions, while random variables $Y \in E^{**} \setminus E$ cannot (see Section 3 below).

Remark 2.5. $\mathsf{EV@R}$ is well-defined on the following domain

$$E' := \left\{ Y \in L^0 : \mathbb{E}\, e^{tY} < \infty \text{ for some } t > 0 \right\},$$

and the entropic spaces E and E^{**} are vector spaces contained in E'.

The set E', however, is *not* a vector space, but it is a convex cone in L^0. To see this, consider, for example, $Y := -X$ where X follows the log-normal distribution. Recall that the moment generating function of the log-normal distribution is finite only on the negative half-axis, so Y belongs to E', while neither is Y in E nor in E^{**}.

Remark 2.6 (Cumulant-generating function). The logarithm of the moment-generating function,

$$K_Y(t) := \log \mathbb{E}\, e^{tY},$$

is also called the *cumulant-generating function* of the random variable Y in probability theory (and the coefficients κ_j in the Taylor series expansion $K_Y(t) = \sum_{j=1} \kappa_j \frac{t^j}{j!}$ are called *cumulants*). Hence, $\mathsf{EV@R}$ can be rewritten as

$$\mathsf{EV@R}_\alpha(Y) = \inf_{t>0} \frac{1}{t} \left(K_Y(t) + \log \frac{1}{1-\alpha} \right).$$

One should note that both $m_Y(t)$ and $K_Y(t)$ are convex in Y for any fixed t, and infinitely differentiable and convex in t for any fixed Y if they are finite.

Remark 2.7 (Lower and upper bounds). We have the following bounds for the $\mathsf{EV@R}$:

$$\mathbb{E}\, Y \le \mathsf{EV@R}_\alpha(Y) \le \operatorname{ess\,sup} Y. \tag{5}$$

The right inequality follows from monotonicity of $\mathsf{EV@R}$. To prove the left one, note that $\log(\cdot)$ is a concave function, and it follows from Jensen's inequality that

$$\frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY} \ge \frac{1}{t} \log \frac{1}{1-\alpha} + \frac{1}{t} \mathbb{E} \log e^{tY} \ge \mathbb{E}\, Y,$$

and hence

$$\mathbb{E}\, Y \le \mathsf{EV@R}_\alpha(Y)$$

for all $0 \le \alpha \le 1$. Moreover, the above lower bound can be improved by

$$\mathsf{AV@R}_\alpha(Y) \le \mathsf{EV@R}_\alpha(Y), \tag{6}$$

where $\mathsf{AV@R}_\alpha$ is the *Average Value-at-Risk* at confidence level $\alpha \in [0, 1)$, given by

$$\mathsf{AV@R}_\alpha(Y) := \inf_{t\in\mathbb{R}} \left\{ t + \frac{1}{1-\alpha}\,\mathbb{E}\max\left(Y-t,0\right)\right\} = \frac{1}{1-\alpha}\int_\alpha^1 \mathsf{V@R}_u(Y)\mathrm{d}u \tag{7}$$

and $\mathsf{AV@R}_1(Y) := \operatorname{ess\,sup}(Y)$ for $\alpha = 1$. $\mathsf{V@R}_\alpha$ stands for the *Value-at-Risk* (or generalized inverse cumulative distribution function) at confidence level $\alpha \in [0, 1]$, i.e.,

$$\mathsf{V@R}_\alpha(Y) := \inf\left\{y : \; P(Y \le y) \ge \alpha\right\}. \tag{8}$$

From a financial perspective, the inequality (6) indicates that the user of $\mathsf{EV@R}$ is more conservative than who prefers $\mathsf{AV@R}$ to quantify the risk Y. The reason is that $\mathsf{EV@R}_\alpha(Y)$ generally incorporates all $\mathsf{V@R}_t(Y)$ with $0 \le t \le 1$ (see (54)), while $\mathsf{AV@R}_\alpha(Y)$ only depends on $\mathsf{V@R}_t(Y)$ with $\alpha \le t \le 1$ (see (7)). Actually, under mild conditions, it can be shown that the additional value

$$\mathsf{EV@R}_\alpha(Y) - \mathsf{AV@R}_\alpha(Y)$$

tends to zero whenever all $\mathsf{V@R}_t(Y)$ with $0 \le t < \alpha$ tend to $-\infty$, which is the best possible outcome for a risk position.

Remark 2.8 (The special case $\alpha = 0$). For $\alpha = 0$, it holds that $\mathsf{EV@R}_0(Y) = \mathbb{E}\,Y$ for $Y \in E^{**}$. Indeed, we have $|e^y - 1 - y| \le \frac{y^2}{2} e^{|y|}$ by Taylor's theorem. It follows from Hölder's inequality that

$$|\,\mathbb{E}\,e^{tY} - 1 - t\;\mathbb{E}\,Y| \le \frac{t^2}{2}\,\mathbb{E}\,Y^2 e^{t|Y|} \le \frac{t^2}{2}\left(\mathbb{E}\,Y^4\right)^{1/2}\left(\mathbb{E}\,e^{2t|Y|}\right)^{1/2} \tag{9}$$

whenever $Y \in E^{**}$ and t is small enough. From this, we conclude that the limit for $t \to 0$ exists in (9), further that $\mathbb{E}\,e^{tY} = 1 + t\cdot\mathbb{E}\,Y + O(t^2)$. Using (5), we conclude that $\mathsf{EV@R}_0(Y) = \mathbb{E}\,Y$ whenever $Y \in E^{**}$.

Remark 2.9 (Convexification of function in (1)). The objective function in (1) can be rewritten as

$$z\log\frac{1}{1-\alpha}\,\mathbb{E}\,e^{\frac{Y}{z}}.$$

The above function is finite and jointly convex in $(z, Y) \in \mathbb{R}_{>0} \times E'$. Indeed, it is the perspective function associated with the convex function $\log\frac{1}{1-\alpha}\,\mathbb{E}\,e^Y$. The convexity of the latter follows from the convexity of the cumulant-generating function $K_Y(t)$ in Y. This convexity result is highly important when the $\mathsf{EV@R}$ is incorporated into large-scale stochastic optimization problems, such as portfolio optimization and distributionally robust optimization (Ahmadi-Javid and Fallah-Tafti (2017), and Postek et al. (2016)).

$\mathsf{EV@R}$ is given as a nonlinear optimization problem in (1). We first characterize the conditions when the optimal value in (1) is attained.

Remark 2.10. As a function of t, the objective function in (1) tends to $+\infty$ as the parameter t tends to zero, and tends to $\operatorname{ess\,sup}(Y)$ as t tends to ∞. The infimum in (1) is thus either attained at some $t^* \in (0, \infty)$ or as t tends to ∞.

Proposition 2.11 (The optimal parameter in (1) to compute $\mathsf{EV@R}$)**.** *The following are equivalent for $Y \in E'$:*

(i) *The infimum in* (1) *is attained at some optimal parameter $t^* \in (0, \infty)$*

(ii) $\mathsf{V@R}_\alpha(Y) < \operatorname{ess\,sup}(Y)$

(iii) $P(Y = \operatorname{ess\,sup}(Y)) < 1 - \alpha$

(iv) $\mathsf{EV@R}_\alpha(Y) < \operatorname{ess\,sup}(Y)$.

Proof. The infimum in (1) is *not* attained if an increasing parameter t improves the objective, and this is the case if and only if

$$\frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY} \geq \operatorname{ess\,sup}(Y)$$

for all $t > 0$. This is equivalent to $\mathbb{E}\, e^{t(Y - \operatorname{ess\,sup}(Y))} \geq 1 - \alpha$ for all $t > 0$. As $Y - \operatorname{ess\,sup}(Y) \leq 0$, this holds if and only if $P(Y = \operatorname{ess\,sup}(Y)) \geq 1 - \alpha$. The equivalence of the other conditions are obvious. □

We start by comparing the values of $\mathsf{EV@R}$ for different confidence levels α, which is used to show the equivalence of the norms induced by $\mathsf{EV@R}$ at different confidence levels.

Proposition 2.12. *For $0 < \alpha \leq \alpha' < 1$ and $Y \in E'$ it holds that*

$$\mathsf{EV@R}_\alpha(Y) \leq \mathsf{EV@R}_{\alpha'}(Y). \tag{10}$$

Conversely, if $Y \geq 0$ (i.e., Y is in the nonnegative cone) it holds true that

$$\mathsf{EV@R}_{\alpha'}(Y) \leq \frac{\log(1-\alpha')}{\log(1-\alpha)} \cdot \mathsf{EV@R}_\alpha(Y). \tag{11}$$

Proof. The inequality (10) is evident by writing $\frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY} = \frac{1}{t} \log \frac{1}{1-\alpha} + \frac{1}{t} \log \mathbb{E}\, e^{tY}$ as $\log \frac{1}{1-\alpha} \leq \log \frac{1}{1-\alpha'}$.

To accept (11), observe that

$$\frac{1}{t} \log \frac{1}{1-\alpha'} \mathbb{E}\, e^{tY} = \frac{1}{t} \left(\log \frac{1}{1-\alpha} \right) \cdot \frac{\log(1-\alpha')}{\log(1-\alpha)} + \frac{1}{t} \log \mathbb{E}\, e^{tY}.$$

As $Y \geq 0$, it follows that $\log \mathbb{E}\, e^{tY} \geq 0$ whenever $t > 0$. Hence, as $\frac{\log(1-\alpha')}{\log(1-\alpha)} \geq 1$,

$$\begin{aligned} \frac{1}{t} \log \frac{1}{1-\alpha'} \mathbb{E}\, e^{tY} &\leq \frac{1}{t} \cdot \frac{\log(1-\alpha')}{\log(1-\alpha)} \log \frac{1}{1-\alpha} + \frac{1}{t} \cdot \frac{\log(1-\alpha')}{\log(1-\alpha)} \log \mathbb{E}\, e^{tY} \\ &= \frac{\log(1-\alpha')}{\log(1-\alpha)} \cdot \frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY}. \end{aligned}$$

By taking the infimum among all $t > 0$ in the latter inequality, it follows that

$$\mathsf{EV@R}_{\alpha'}(Y) \leq \frac{\log(1-\alpha')}{\log(1-\alpha)} \, \mathsf{EV@R}_{\alpha}(Y),$$

which is the assertion. □

Remark 2.13 (Exclusion of the special cases $\alpha = 0$ and $\alpha = 1$). The cases $\alpha = 0$ and $\alpha = 1$ are excluded in the previous proposition. Indeed, for $\alpha = 0$, $\mathsf{EV@R}_0(Y) = \mathbb{E}\, Y$; while for $\alpha = 1$, $\mathsf{EV@R}_1(Y) = \operatorname{ess\,sup} Y$. Hence, we do not further consider these special cases in what follows.

It is observed in Pichler (2013) and Kalmes and Pichler (2017) that coherent risk measures induce semi-norms on vector spaces contained in their domains. In the sequel, we define the EV@R norm for $Y \in E^{**}$ by

$$\|Y\| := \mathsf{EV@R}_{\alpha}(|Y|). \tag{12}$$

The sets E and E^{**} do not depend on the confidence level, and further, by Proposition 2.12 the norms (12) are equivalent for different confidence levels $0 < \alpha < 1$. For this reason, we do not explicitly indicate the confidence level α in the notation of the norm $\|\cdot\| = \mathsf{EV@R}_{\alpha}(|\cdot|)$ in the rest of the paper. Recall that for the special cases $\alpha = 0$ and $\alpha = 1$, which are excluded in the following, the EV@R norm coincides with the L^1 and L^∞ norms, respectively.

The next theorem establishes that E (E^{**}, resp.) equipped with the EV@R norm $\|\cdot\|$ is a Banach space. This immediately implies that E^* equipped with the dual EV@R norm $\|\cdot\|^*$ is a Banach space.

Theorem 2.14. *For $0 < \alpha < 1$, the pairs*

$$(E, \|\cdot\|) \text{ and } (E^{**}, \|\cdot\|)$$

are (different) Banach spaces for the norm $\|\cdot\| = \mathsf{EV@R}_{\alpha}(|\cdot|)$.

Proof. It is first shown that E and E^{**} are complete. We start by considering E^{**} first. To demonstrate completeness, let Y_n be a Cauchy sequence. For $\varepsilon > 0$, one can find $n > 0$ such that $\|Y_n - Y_m\| < \varepsilon$ whenever $m > n$, and thus $\|Y_m\| \leq \|Y_n\| + \|Y_m - Y_n\| < \|Y_n\| + \varepsilon$. Hence, $\lim_{n\to\infty} \|Y_n\|$ exists, and $\|Y_n\| < C$ for some constant $C < \infty$. The norms are, therefore, uniformly bounded.

Now recall that $\|Y_n\| = \mathsf{EV@R}_{\alpha}(|Y_n|)$, so we can choose $t_n > 0$ in (1) to have

$$\frac{1}{t_n} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{t_n |Y_n|} < C. \tag{13}$$

Next, note that $\mathbb{E}\, e^{t_n |Y_n|} \geq 1$; hence,

$$\frac{1}{t_n} \log \frac{1}{1-\alpha} \leq \frac{1}{t_n} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{t_n |Y_n|} < C.$$

It follows that

$$t_n > t^* := \frac{1}{C} \log \frac{1}{1-\alpha} > 0,$$

and $\mathbb{E}\, e^{t|Y_n|}$ is well-defined, by (13), for every $t < t^*$.

Now recall from Remark 2.7 that

$$\mathbb{E}\,|Y| \le \mathsf{EV@R}_\alpha(|Y|).$$

As Y_n is a Cauchy sequence for $\|\cdot\| = \mathsf{EV@R}_\alpha(|\cdot|)$, it follows that Y_n is a Cauchy sequence for L^1 as well. From completeness of L^1, we conclude that the Cauchy sequence Y_n has a limit $Y \in L^1$. It remains to be shown that $Y \in E^{**}$ and that $\mathsf{EV@R}(|Y - Y_n|) \to 0$.

Denote the cumulative distribution function of the random variable Y by $F_Y(x) := P(Y \le x)$, and its generalized inverse by

$$F_Y^{-1}(p) := \mathsf{V@R}_p(Y) = \inf\{x : \; F_Y(x) \ge p\}$$

(see (8)). It follows from convergence in L^1 that Y_n converges in distribution, that is, $F_{Y_n}(y) \to F_Y(y)$ for every point y at which $F_Y(\cdot)$ is continuous, and consequently $F^{-1}_{|Y_n|}(\cdot) \to F^{-1}_{|Y|}(\cdot)$ (Vaart (1998, Chapter 21)). Hence, by Fatou's lemma,

$$\begin{aligned} \mathbb{E}\, e^{t|Y|} &= \int_0^1 \exp\left(tF^{-1}_{|Y|}(p)\right)\mathrm{d}p = \int_0^1 \liminf_{n\to\infty} \exp\left(tF^{-1}_{|Y_n|}(p)\right)\mathrm{d}p \\ &\le \liminf_{n\to\infty} \int_0^1 \exp\left(tF^{-1}_{|Y_n|}(p)\right)\mathrm{d}p = \liminf_{n\to\infty} \mathbb{E}\, e^{t|Y_n|} \le (1-\alpha)e^{t^* C} \end{aligned} \tag{14}$$

whenever $t < t^*$. It follows that $\|Y\| \le C < \infty$, and thus $Y \in E^{**}$.

It remains to be shown that $\mathsf{EV@R}_\alpha(|Y_n - Y|) \to 0$. For $\varepsilon > 0$ choose N large enough so that $\mathsf{EV@R}_\alpha(|Y_n - Y_m|) < \varepsilon$ for $m, n > N$. Therefore, there must be parameters t that satisfy

$$\frac{1}{t}\log\frac{1}{1-\alpha} \le \frac{1}{t}\log\frac{1}{1-\alpha}\,\mathbb{E}\, e^{t|Y_n - Y_m|} < \varepsilon,$$

and thus we can assume

$$t > t^* := \frac{1}{\varepsilon}\log\frac{1}{1-\alpha}. \tag{15}$$

By Jensen's inequality for $t > t^*$,

$$\begin{aligned} \frac{1}{t}\log\frac{1}{1-\alpha}\,\mathbb{E}\, e^{t|Y_n - Y_m|} &= \frac{1}{t}\log\frac{1}{1-\alpha}\,\mathbb{E}\left(e^{t^*|Y_n - Y_m|}\right)^{t/t^*} \ge \frac{1}{t}\log\frac{1}{1-\alpha}\left(\mathbb{E}\, e^{t^*|Y_n - Y_m|}\right)^{t/t^*} \\ &= \frac{1}{t}\log\frac{1}{1-\alpha} + \frac{1}{t^*}\log\mathbb{E}\, e^{t^*|Y_n - Y_m|} \ge \frac{1}{t^*}\log\mathbb{E}\, e^{t^*|Y_n - Y_m|}. \end{aligned}$$

Taking the infimum with respect to $t > t^*$ reveals that $\mathbb{E}\, e^{t^*|Y_n - Y_m|} \le e^{t^*\varepsilon}$ whenever $m, n > N$.

It follows from convergence in probability (a consequence of convergence in L^1) that there is a subsequence converging almost surely to Y. Without loss of generality we assume that Y_n is this sequence. Then,

$$\mathbb{E}\, e^{t^*|Y - Y_n|} = \mathbb{E} \lim_{m\to\infty} e^{t^*|Y_n - Y_m|} \le \liminf_{m\to\infty} \mathbb{E}\, e^{t^*|Y_n - Y_m|} \le e^{t^*\varepsilon}$$

by Fatou's lemma. Finally

$$\frac{1}{t^*}\log\frac{1}{1-\alpha}\,\mathbb{E}\,e^{t^*|Y-Y_n|}\le\frac{1}{t^*}\log\frac{1}{1-\alpha}e^{t^*\varepsilon}\le\frac{1}{t^*}\log\frac{1}{1-\alpha}+\varepsilon=2\varepsilon$$

due to (15). It follows that $\mathsf{EV@R}(|Y-Y_n|)\to 0$, and thus E^{**} is complete and a Banach space.

The proof that E is complete is along the same lines as for completeness of E^{**} (particularly (14)), except that it is not necessary to find a number $t^*>0$ for which all moments $\mathbb{E}\,e^{tY}$ are well defined; by definition, this is the case for $Y_n\in E$. This completes the proof. ◻

3 Comparison with normed spaces

This section relates the entropic spaces with the L^p spaces in the first subsection. To further specify the nature of the entropic spaces, the $\mathsf{EV@R}$ norm is related to specific Orlicz and Luxembourg norms in the subsequent subsection.

3.1 Comparison with L^p spaces

In this section, we proceed with a comparison of the $\mathsf{EV@R}$ norm with the L^p-norms (see Fig. 1).

Theorem 3.1. *For* $\|\cdot\|=\mathsf{EV@R}_\alpha(|\cdot|)$ *at every confidence level* $0<\alpha<1$*, it holds that*

$$\|Y\|_1\le\|Y\|,\ Y\in E^{**}\quad\text{ and }\quad\|Y\|\le\|Y\|_\infty,\ Y\in L^\infty. \tag{16}$$

Further, for every $1<p<\infty$ *there is a finite constant* $c\le\max\left\{1,\frac{p-1}{\log\frac{1}{1-\alpha}}\right\}$ *such that*

$$\|Y\|_p\le c\cdot\mathsf{EV@R}_\alpha(|Y|),\ Y\in E^{**}. \tag{17}$$

It holds particularly that $L^\infty\subset E\subset E^{**}\subset\bigcap_{p\ge 1}L^p$.

Proof. The first part follows from Remark 2.7. Without loss of generality we shall assume throughout this proof that $Y\ge 0$. To prove (17), first assume that $\alpha=1-e^{1-p}$. Then,

$$\frac{1}{t}\log\frac{1}{1-\alpha}\,\mathbb{E}\,e^{tY}=\frac{1}{t}\log\frac{e^{1-p}}{1-\alpha}+\frac{1}{t}\log\mathbb{E}\,e^{p-1}e^{tY}=\frac{1}{t}\left(\varphi\left(e^{p-1}\,\mathbb{E}\,e^{tY}\right)\right)^{1/p} \tag{18}$$

where $\varphi(x)=\big(\log(x)\big)^p$. Now note that

$$\varphi''(x)=\frac{p}{x^2}\big(\log x\big)^{p-2}(p-1-\log x).$$

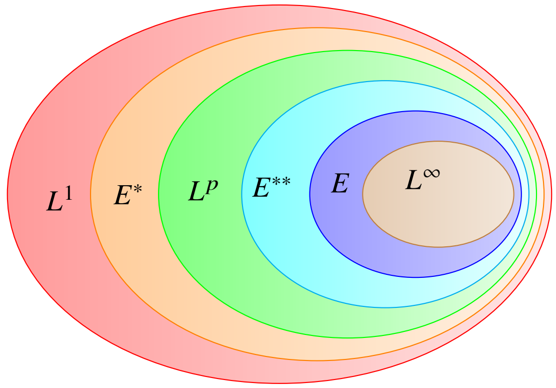

Figure 1: Cascading arrangement of the entropic spaces E, E^*, and E^{**}; and their relation to the L^p spaces, $1 < p < \infty$

Hence, the function $\varphi(\cdot)$ is concave (i.e., $\varphi'' \le 0$) provided that the argument x satisfies $x \ge e^{p-1}$; this is the case in (18), as $\mathbb{E}\, e^{tY} \ge 1$. We apply Jensen's inequality and obtain

$$\begin{aligned}\frac{1}{t}\log\frac{1}{1-\alpha}\,\mathbb{E}\, e^{tY} &= \frac{1}{t}\left(\varphi\left(\mathbb{E}\, e^{p-1+tY}\right)\right)^{1/p} \ge \frac{1}{t}\left(\mathbb{E}\,\varphi\left(e^{p-1+tY}\right)\right)^{1/p} \\ &= \frac{1}{t}\left(\mathbb{E}(p-1+tY)^p\right)^{1/p} \ge \frac{1}{t}\left(\mathbb{E}(t\,Y)^p\right)^{1/p} = \|Y\|_p,\end{aligned} \tag{19}$$

as $t\,Y \ge 0$. Taking the infimum in (19) among all $t > 0$ reveals that

$$\|Y\|_p \le \mathsf{EV@R}_{1-e^{1-p}}(|Y|).$$

The assertion follows finally from Proposition 2.12. ◻

Theorem 3.2. *L^∞ and simple functions are dense in E, but* not *dense in E^{**} for the norm $\|\cdot\| = \mathsf{EV@R}_\alpha(|\cdot|)$ at any $0 < \alpha < 1$.*

Remark 3.3. Note that simple functions and L^∞ are of course dense whenever $\alpha = 0$, as they are dense in L^1.

Proof. Let Y follow the exponential distribution, i.e., $P(e^{-Y} \le u) = u$ for $u > 0$. The moment generating function of this distribution is finite only for $t < 1$,

$$\mathbb{E}\, e^{tY} = \frac{1}{1-t},$$

which shows that $Y \in E^{**} \setminus E$. Further, note that for $Y_n := \min\{Y, n\}$

$$\mathbb{E}\, e^{t|Y-Y_n|} \ge 1,$$

and hence for $0 < \alpha < 1$ and $t < 1$

$$\frac{1}{t}\log\frac{1}{1-\alpha}\,\mathbb{E}\,e^{t|Y-Y_n|} \geq \frac{1}{t}\log\frac{1}{1-\alpha} \geq \log\frac{1}{1-\alpha} > 0.$$

Now if S_n is any step function with $S_n \leq n$, then $|Y - S_n| \geq |Y - Y_n|$; and the result follows from monotonicity.

To accept that simple functions (and consequently L^∞) are dense in E, let $Y \in E$ be fixed. By Hölder's inequality,

$$\begin{aligned}\mathbb{E}\,e^{tY\mathbb{1}_{\{Y\geq n\}}} =& P(Y \leq n) + \mathbb{E}\,\mathbb{1}_{\{Y\geq n\}}e^{tY} \\ \leq& 1 + \left(\mathbb{E}\,\mathbb{1}^2_{\{Y\geq n\}}\right)^{1/2}\left(\mathbb{E}\,e^{2tY}\right)^{1/2} \\ =& 1 + P(Y \geq n)^{1/2}\left(\mathbb{E}\,e^{2tY}\right)^{1/2}.\end{aligned}$$

Recall the following extended version of Markov's inequality,

$$P(Y \geq n) \leq \frac{\mathbb{E}\,\varphi(Y)}{\varphi(n)}, \tag{20}$$

where $\varphi(\cdot)$ is an increasing function and $\varphi(n) > 0$. We choose $\varphi(y) := \exp(2ty)$ in (20), and it follows that

$$\mathbb{E}\,e^{tY\mathbb{1}_{\{Y\geq n\}}} \leq 1 + \left(\frac{\mathbb{E}\,e^{2tY}}{e^{2tn}}\right)^{1/2}\cdot\left(\mathbb{E}\,e^{2tY}\right)^{1/2} = 1 + \frac{\mathbb{E}\,e^{2tY}}{e^{tn}}.$$

Consequently,

$$\mathsf{EV@R}_\alpha\left(Y\mathbb{1}_{\{Y\geq n\}}\right) \leq \frac{1}{t}\log\frac{1}{1-\alpha}\left(1 + \frac{\mathbb{E}\,e^{2tY}}{e^{tn}}\right) \leq \frac{1}{t}\log\frac{2}{1-\alpha}, \tag{21}$$

provided that n is sufficiently large. As $Y \in E$, the moment generating function exists for all $t > 0$. Hence, the latter equation is valid for every $t > 0$, and it follows that

$$0 \leq \mathsf{EV@R}_\alpha(Y - Y_n) \leq \mathsf{EV@R}_\alpha\left(Y\mathbb{1}_{\{Y\geq n\}}\right) \xrightarrow{n\to\infty} 0$$

by monotonicity, where $Y_n = \min\{Y, n\}$.

Finally, note that Y_n is bounded, and thus it can be approximated sufficiently close by simple functions (step functions). This completes the proof. □

Remark 3.4. We want to point out that all moments of Y exist whenever $Y \in E$, and thus

$$Y \in E \implies Y \in L^p$$

for all $1 \leq p < \infty$. However, E is not closed in L^p, and a converse relation to Theorem 3.1 does not hold true on E. More specifically, there is no finite constant $c < \infty$ for which $\mathsf{EV@R}_\alpha(|Y|) \leq c\,\|Y\|_p$ holds true for every $Y \in E$ (except $p = \infty$).

We provide an explicit counterexample: consider the random variables Y_n with $P(Y_n = n) = {}^1\!/\!_{n^p}$ and $P(Y_n = 0) = 1 - {}^1\!/\!_{n^p}$. Then, $\|Y_n\|_p = 1$ and

$$\frac{1}{t}\log\frac{1}{1-\alpha}\,\mathbb{E}\,e^{tY_n} = \frac{1}{t}\log\frac{1}{1-\alpha}\left(1-\frac{1}{n^p}+\frac{1}{n^p}e^{tn}\right). \tag{22}$$

Now note that

$$(22) \geq \frac{1}{t}\log\frac{1}{1-\alpha} \tag{23}$$

$$(22) \geq \frac{1}{t}\log\frac{e^{tn}}{(1-\alpha)n^p} = n + \frac{1}{t}\log\frac{1}{1-\alpha} - \frac{p\log n}{t}. \tag{24}$$

The function in (23) is decreasing in t, while the function in (24) is increasing for $n \geq \left(\frac{1}{1-\alpha}\right)^{1/p}$. We choose $t = \frac{p\log n}{n}$ in (23) and (24), and it follows that $\mathsf{EV@R}_\alpha(Y_n) \geq \frac{n}{p\log n}\log\frac{1}{1-\alpha}$ from which the assertion is immediate.

3.2 Relation to Orlicz spaces

In this section, we demonstrate that the entropic spaces are equivalent to some Orlicz hearts and Orlicz spaces. Let Φ and Ψ be coercive convex functions with $\Phi(0) = \Psi(0) = 0$, which are conjugate functions to each other. These functions are called a *pair of complementary Young functions* in the context of Orlicz spaces, and they satisfy

$$y\,z \leq \Phi(y) + \Psi(z).$$

Definition 3.5. For a pair Φ and Ψ of complementary Young functions, consider the vector spaces

$$L^\Phi := \{Y \in L^0 : \mathbb{E}\,\Phi(t|Y|) < \infty \text{ for some } t > 0\}$$

$$L^\Phi_* := \{Y \in L^0 : \mathbb{E}\,\Phi(t|Y|) < \infty \text{ for all } t > 0\},$$

and define the norms

$$\|Y\|_\Phi := \sup_{\mathbb{E}\,\Psi(|Z|)\leq 1} \mathbb{E}\,YZ \tag{25}$$

$$\|Y\|_{(\Phi)} := \inf\left\{k > 0 : \mathbb{E}\,\Phi\left(\frac{|Y|}{k}\right) \leq 1\right\}. \tag{26}$$

The norms $\|\cdot\|_\Phi$ and $\|\cdot\|_{(\Phi)}$ are called the *Orlicz norm* and *Luxemburg norm*, respectively. The spaces L^Φ and L^Φ_* are called the *Orlicz space* and *Orlicz heart*, respectively. In an analogous way, we define L^Ψ and L^Ψ_*, as well as the norms $\|\cdot\|_\Psi$ and $\|\cdot\|_{(\Psi)}$.

Remark 3.6. The Orlicz norm $\|\cdot\|_\Phi$ and the Luxemburg norm $\|\cdot\|_{(\Phi)}$ are equivalent,

$$\|Y\|_{(\Phi)} \leq \|Y\|_\Phi \leq 2\,\|Y\|_{(\Phi)},$$

over L^Φ (Pick et al. (2012, Theorem 4.8.5)). Hence, we present our results only for the Orlicz norm in the following.

In the rest of this paper, we consider the pair

$$\Phi(y) := \begin{cases} y & \text{if } y \leq 1, \\ e^{y-1} & \text{if } y \geq 1 \end{cases} \quad \text{and} \quad \Psi(z) := \begin{cases} 0 & \text{if } z \leq 1, \\ z\log z & \text{if } z \geq 1 \end{cases} \tag{27}$$

of complementary Young functions. It is clear that $\Phi(y)$ and $e^y - 1$ are equivalent, in the sense that they generate the same Orlicz hearts and spaces.

We have the following relations among the entropic spaces, and the Orlicz hearts and spaces corresponding with Young function $\Phi(\cdot)$ given in (27).

Theorem 3.7. *It holds that $E = L^\Phi_*$, $E^* = L^\Psi_* = L^\Psi$, and $E^{**} = L^\Phi$. Indeed, for $0 < \alpha < 1$, the norms*

$$\|\cdot\| = \mathsf{EV@R}_\alpha(|\cdot|) \quad \textit{and} \quad \|\cdot\|_\Phi$$

are equivalent on E^{**}, where Φ is the Young function (27). Particularly, it holds that

$$\|Y\| \leq c\cdot\|Y\|_\Phi \qquad \textit{for all } Y \in E^{**}$$

for some $c \leq \max\left\{e,\, \log\frac{1}{1-\alpha}\right\}$.

Proof. By employing the inequality $\log y \leq y - 1$ and the elementary inequality

$$\log\frac{1}{1-\alpha} - 1 + e^y \leq c\cdot(1 + \Phi(y)),$$

which is valid for $c = \max\left\{e,\, \log\frac{1}{1-\alpha}\right\}$, it follows that

$$\log\frac{1}{1-\alpha}\,\mathbb{E}\,e^{t|Y|} = \log\frac{1}{1-\alpha} + \log\mathbb{E}\,e^{t|Y|} \leq \log\frac{1}{1-\alpha} - 1 + \mathbb{E}\,e^{t|Y|} \leq c\big(1 + \mathbb{E}\,\Phi(t|Y|)\big),$$

and hence

$$\mathsf{EV@R}_\alpha(|Y|) \leq c\cdot\inf_{t>0}\frac{1}{t}\big(1 + \mathbb{E}\,\Phi(t|Y|)\big). \tag{28}$$

By Krasnosel'skii and Rutickii (1961, Theorem 10.5) (see also Pick et al. (2012, Remark 4.8.9 (i))), the Orlicz norm has the equivalent expression

$$\|Y\|_\Phi = \inf_{t>0}\frac{1}{t}\big(1 + \mathbb{E}\,\Phi(t|Y|)\big).$$

Therefore, the assertion

$$\mathsf{EV@R}_\alpha(Y) \le c \cdot \|Y\|_\Phi \quad \text{for all } Y \in L^\Phi \tag{29}$$

follows from (28). Note that we have particularly proved that $Y \in L^\Phi \implies Y \in E^{**}$, i.e., $L^\Phi \subset E^{**}$.

To prove the converse inequality, let $Y \in E^{**}$. Then, by Definition 2.1, there is a number $t^* > 0$ such that $\mathbb{E}\, e^{tY} < \infty$ is finite whenever $t \le t^*$. Hence, the Luxemburg norm $\|Y\|_{(\Phi)}$, given in (26), which is equivalent to the Orlicz norm on L^Φ, is also finite because of $\Phi(x) \le \frac{1}{e} e^x$. We can, thus, conclude that $\|Y\|_\Phi < \infty$, i.e., $Y \in L^\Phi$.

Now consider the identity map

$$i : \left(L^\Phi, \|\cdot\|_\Phi\right) \to \left(E^{**}, \mathsf{EV@R}_\alpha(|\cdot|)\right),$$

which is bounded ($\|i\| \le c$ by (29)). By the above reasoning, i is bijective; and because $(E^{**}, \mathsf{EV@R}_\alpha(|\cdot|)$ is a Banach space by Theorem 2.14, it follows from the bounded inverse theorem (open mapping theorem, Rudin (1973, Corollary 2.12)) that the inverse i^{-1} is continuous as well, i.e., there is a constant $c' < \infty$ such that

$$\|Y\|_\Phi \le c' \cdot \mathsf{EV@R}_\alpha(|Y|) \quad \text{for all } Y \in E^{**}.$$

Finally note that the function $\Psi(\cdot)$ in Eq. (27) satisfies the Δ_2-condition, i.e., $\Psi(2x) \le k\, \Psi(x)$ for every $k > 2$ whenever x is large enough. Then, it follows from Pick et al. (2012, Proposition 4.12.3) that $L_*^\Psi = L^\Psi$, from which the remaining statement is immediate. This completes the proof. □

4 Duality

This section first establishes the duality relations of the entropic spaces. In the second part, we provide an explicit expression for the EV@R dual norm. To complete the dual description, we explicitly give the Hahn–Banach functionals, where the Hahn–Banach functional for a norm is an optimal functional for the optimization problem corresponding to the dual norm. The last two subsections provide the dual and Kusuoka representations of the EV@R, as well as explicit expressions for optimal densities in both representations. A new version of the classical Donsker-Varadhan formula, and its dual formula, is also presented.

4.1 Duality of entropic spaces

Theorem 3.7 above makes a series of results developed for Orlicz spaces available for the Banach spaces $(E, \|\cdot\|)$, its dual $(E, \|\cdot\|)^*$, and $(E^{**}, \|\cdot\|)$. We derive the following relations for these Banach spaces. They reveal that E^* and E^{**} are the norm dual and bidual spaces of E.

Theorem 4.1. *For $\|\cdot\| = \mathsf{EV@R}_\alpha(|\cdot|)$ at every $0 < \alpha < 1$ and the pair of complementary Young functions (27), it holds that*

(i) $(E, \|\cdot\|) \cong \left(L_*^\Phi, \|\cdot\|_\Phi\right)$

(ii) $(E^*, \|\cdot\|^*) \cong \left(L_*^\Psi, \|\cdot\|_\Psi\right) \cong \left(L^\Psi, \|\cdot\|_\Psi\right)$

(iii) $(E^{**}, \|\cdot\|) \cong \left(L^\Phi, \|\cdot\|_\Phi\right)$.

Further, the duality relations

(iv) $(E, \|\cdot\|)^* \cong (E^*, \|\cdot\|^*) \cong \left(L^\Psi, \|\cdot\|_\Psi\right)$

(v) $(E, \|\cdot\|)^{**} \cong (L^\Psi, \|\cdot\|_\Psi)^* \cong \left(L^\Phi, \|\cdot\|_\Phi\right) \cong (E^{**}, \|\cdot\|)$

hold true (here, $\cong$ denotes the continuous homomorphism and the superscript * the dual space).

Proof. By employing Pick et al. (2012, Theorem 4.13.6), we deduce that

$$\left(L_*^\Phi, \|\cdot\|_\Phi\right)^* \cong \left(L^\Psi, \|\cdot\|_\Psi\right). \tag{30}$$

It is evident that $E = L_*^\Psi$ by the definition of the spaces, and the equivalence of the norms has already been established in Theorem 3.7 in a broader context. Thus, (i) and (iv) follow from (30). Moreover, (ii) and (iii) follow from Theorem 3.7 by noting (30).

As for (v), we recall that $\Psi(\cdot)$ in Eq. (27) satisfies the Δ_2-condition. Then, it follows from Pick et al. (2012, Proposition 4.12.3) that $L_*^\Psi = L^\Psi$, and further from Pick et al. (2012, Theorem 4.13.6), that the dual space of L^Ψ is L^Φ, i.e.,

$$\left(L^\Psi, \|\cdot\|_\Psi\right)^* \cong \left(L^\Phi, \|\cdot\|_\Phi\right) \cong \left(E^{**}, \mathsf{EV@R}_\alpha(|\cdot|)\right).$$

Hence, we get (v) by employing (iii) and (i). This completes the proof. □

Remark 4.2. One should note that Φ does *not* satisfy the Δ_2-condition, and hence the space $(E, \|\cdot\|)$ is *not* reflexive. This follows as well from $E^{**} \supsetneqq E$. Moreover, by (iii) in Theorem 4.1, the norm dual of E^{**} can be determined using the results available in the literature of Orlicz spaces (see e.g., Section 3 of Rao (1968) and Section 2.2 of Biagini and Frittelli (2008)).

Remark 4.3. The duality relations in Theorem 4.1 are specified in the above-mentioned references as well. One should note that the following natural mapping (the inner product on $L^\Phi \times L^\Psi$):

$$\langle Y, Z\rangle \mapsto \mathbb{E}\, Y Z$$

is the bilinear form considered to study duality results in this section, such as dual norms. Indeed, for each Z in the dual space L^Ψ, the functional $\langle \cdot, Z\rangle$ is the associated continuous linear functional on the primal space L^Φ.

4.2 Explicit expression of dual norm

The dual norms corresponding to the Orlicz and Luxemburg norms associated with Young functions in (27) cannot probably be derived explicitly. To see this, for example, the dual norm $\|\cdot\|_\Phi^*$ is defined by

$$\|Z\|_\Phi^* = \sup_{\|Y\|_\Phi \le 1} \mathbb{E}\, YZ,$$

where $\|\cdot\|_\Phi$ is given by the rather involved expression (25); recall that the latter is a supremum over a *continuum* of random variables.

However, for the $\mathsf{EV@R}$, which is equivalent to the above mentioned norms by Theorem 3.7, the following theorem makes a simple explicit expression for the dual norm

$$\|Z\|^* := \sup_{\mathsf{EV@R}_\alpha(|Y|) \le 1} \mathbb{E}\, YZ$$

available, which is a supremum over a single parameter. Indeed, the norms $\|\cdot\|^*$, $\|\cdot\|_\Phi^*$, and $\|\cdot\|_\Psi$ ($\|\cdot\|_{(\Phi)}^*$, and $\|\cdot\|_{(\Psi)}$) are notably equivalent by Theorem 4.1.

Theorem 4.4 ($\mathsf{EV@R}$ dual norm)**.** *For $0 < \alpha < 1$ and $Z \in E^*$ (i.e., $Z \in L^\Psi$), we have the explicit expression*

$$\|Z\|^* = \sup_{c>0} \frac{\mathbb{E}\, |Z| \log\left(\frac{|Z|}{c} \vee 1\right)}{\log \frac{1}{1-\alpha}\, \mathbb{E}\left(\frac{|Z|}{c} \vee 1\right)} \tag{31}$$

for the dual norm $\|\cdot\|^*$ on E^*, where $x \vee y = \max\{x, y\}$.

Proof. Without loss of generality, we may assume that $Z \ge 0$. Then, it holds that

$$\begin{aligned}\|Z\|^* &= \sup_{Y\neq 0} \frac{\mathbb{E}\, YZ}{\mathsf{EV@R}(|Y|)} = \sup_{Y\neq 0} \frac{\mathbb{E}\, YZ}{\inf_{t>0} \frac{1}{t} \log \mathbb{E}\, \frac{1}{1-\alpha} e^{t|Y|}} \\ &= \sup_{Y\neq 0} \sup_{t>0} \frac{\mathbb{E}\, \frac{Y}{t} Z}{\frac{1}{t} \log \mathbb{E}\, \frac{1}{1-\alpha} e^{|Y|}} = \sup_{Y\neq 0} \frac{\mathbb{E}\, YZ}{\log \frac{1}{1-\alpha}\, \mathbb{E}\, e^{|Y|}}.\end{aligned} \tag{32}$$

Now define $Y_c := \log\left(\frac{Z}{c} \vee 1\right)$ and observe that $Y_c \ge 0$ for every $c > 0$. Hence,

$$\|Z\|^* \ge \sup_{c>0} \frac{\mathbb{E}\, Z \log\left(\frac{Z}{c} \vee 1\right)}{\log \frac{1}{1-\alpha}\, \mathbb{E}\left(\frac{Z}{c} \vee 1\right)}, \tag{33}$$

which is the first inequality required to prove (31).

To obtain the converse inequality observe that, by (32), $\lambda \ge \|Z\|^*$ is equivalent to

$$\mathbb{E}\, YZ - \lambda \cdot \log \frac{1}{1-\alpha}\, \mathbb{E}\, e^Y \le 0 \quad \text{for every} \quad Y \ge 0.$$

To maximize this latter expression with respect to *nonnegative* $Y \geq 0$, consider the Lagrangian

$$L(Y,\mu) := \mathbb{E}\, YZ - \lambda \cdot \log \frac{1}{1-\alpha}\, \mathbb{E}\, e^Y - \mathbb{E}\, Y\mu,$$

where μ is the Lagrangian multiplier associated with the constraint $Y \geq 0$. The Lagrangian L is differentiable, and its directional derivative with respect to Y in direction $H \in L^\Phi$ is

$$\frac{\partial}{\partial Y} L(Y,\mu)H = \mathbb{E}\, ZH - \lambda \cdot \frac{\frac{1}{1-\alpha}\,\mathbb{E}\, e^Y H}{\frac{1}{1-\alpha}\,\mathbb{E}\, e^Y} - \mathbb{E}\, \mu H = \mathbb{E}(Z-\mu)\cdot H - \frac{\lambda}{\mathbb{E}\, e^Y}\, \mathbb{E}\, e^Y \cdot H. \tag{34}$$

The derivative vanishes in every direction H, i.e., $\frac{\partial L}{\partial Y}(Y,\mu)H = 0$, and it follows from (34) that

$$Z - \mu = c \cdot e^Y \text{ a.s.}, \tag{35}$$

where $c := \frac{\lambda}{\mathbb{E}\, e^Y} > 0$ is a positive constant. By complementary slackness for the optimal $Y \geq 0$ and the associated multiplier $\mu \geq 0$ one may conclude that

$$Y > 0 \iff \mu = 0 \iff Z = ce^Y > c,$$

and hence, by (35),

$$Y = \begin{cases} \log \frac{Z}{c} & \text{if } Y \geq 0 \\ 0 & \text{if } Y = 0 \end{cases} = \log\left(\frac{Z}{c} \vee 1\right).$$

It is thus sufficient to consider Y_c in (33) to prove the remaining assertion. □

To evaluate the dual norm numerically, we continue by examining the objective function in (31).

Proposition 4.5. *The objective function*

$$\varphi_\alpha(c) := \frac{\mathbb{E}\,|Z| \log\left(\frac{|Z|}{c} \vee 1\right)}{\log \frac{1}{1-\alpha}\, \mathbb{E}\left(\frac{|Z|}{c} \vee 1\right)}$$

in the expression of the dual norm (31) *extends continuously to* $c = 0$*, and it holds that*

$$\lim_{c \downarrow 0} \varphi_\alpha(c) = \mathbb{E}\,|Z|. \tag{36}$$

Further, the supremum is attained at some $c \geq 0$*. If* $Z \neq 0$ *is bounded, then the optimal* c *satisfies* $0 \leq c < \|Z\|_\infty$.

Remark 4.6. The function $\varphi_\alpha(\cdot)$ can be continuously extended to $[0,\infty)$. $\varphi_\alpha(\cdot)$ is, however, not necessarily monotonic, convex, nor concave in general.

Proof. First, note that the objective is continuous and $\varphi_\alpha(\cdot) \geq 0$. One may consider the numerator and denominator separately for $c \to \infty$ to get

$$\varphi_\alpha(c) = \frac{\mathbb{E}\,|Z| \log\left(\frac{|Z|}{c} \vee 1\right)}{\log \frac{1}{1-\alpha}\, \mathbb{E}\left(\frac{|Z|}{c} \vee 1\right)} \xrightarrow{c\to\infty} \frac{0}{\log \frac{1}{1-\alpha}} = 0.$$

For the case $c \to 0$ we find that

$$\varphi_\alpha(c) = \frac{\mathbb{E}\,|Z| \log\left(\frac{|Z|}{c} \vee 1\right)}{\log \frac{1}{1-\alpha}\, \mathbb{E}\left(\frac{|Z|}{c} \vee 1\right)} = \frac{\log \frac{1}{c} \cdot \mathbb{E}\,|Z| \cdot (1 + o(1))}{\log \frac{1}{c} + o(1)} \xrightarrow{c\to 0} \mathbb{E}\,|Z|.$$

Finally, for $c \geq \|Z\|_\infty$ the numerator of φ_α is $\mathbb{E}\,|Z| \log\left(\frac{|Z|}{c} \vee 1\right) = 0$, and thus $\varphi_\alpha(c) = 0$, from which the remaining claim of the proposition follows by the continuity of $\varphi_\alpha(\cdot)$. □

4.3 Hahn–Banach functionals

In this section, we describe the Hahn–Banach functionals corresponding to $Y \in E^{**}$ and $Z \in E^*$ explicitly. That is, we identify the random variable $Z \in E^*$ which maximizes the objective function in the expression

$$\mathsf{EV@R}_\alpha(|Y|) = \sup_{Z\neq 0} \frac{\mathbb{E}\,YZ}{\|Z\|^*}, \tag{37}$$

and the random vairable $Y \in E^{**}$ maximizing the objective function in the follwing supremum

$$\|Z\|^* = \sup_{Y\neq 0} \frac{\mathbb{E}\,ZY}{\mathsf{EV@R}_\alpha(|Y|)}. \tag{38}$$

Proposition 4.7. *Let $Y \in E^{**}$ and suppose there is an optimal $t^* \in (0,\infty)$ for (1) when Y is replaced by $|Y|$, then*

$$Z := \operatorname{sign}(Y) \cdot e^{t^*|Y|}$$

maximizes (37), that is, for $\|\cdot\| = \mathsf{EV@R}_\alpha(|\cdot|)$ at any $0 < \alpha < 1$,

$$\|Y\| = \frac{\mathbb{E}\,YZ}{\|Z\|^*}. \tag{39}$$

Proof. Without loss of generality, we may assume $Y \geq 0$. Consider the function $\varphi(t) := \frac{1}{t} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY}$. Given $t' := \sup\{t \in (0,\infty) : m_Y(t) < +\infty\}$, the function $\varphi(\cdot)$ is continuously differentiable on $(0, t')$ with derivative

$$\varphi'(t) := -\frac{1}{t^2} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{tY} + \frac{1}{t} \cdot \frac{\frac{1}{1-\alpha} \mathbb{E}\, Y e^{tY}}{\frac{1}{1-\alpha} \mathbb{E}\, e^{tY}}.$$

The point $t^* \in (0, t')$ attaining the infimum satisfies $\varphi'(t^*) = 0$, i.e.,

$$\frac{1}{t^*} \log \frac{1}{1-\alpha} \mathbb{E}\, e^{t^* Y} = \frac{\mathbb{E}\, Y e^{t^* Y}}{\mathbb{E}\, e^{t^* Y}}, \tag{40}$$

so we have

$$\mathbb{E}\, YZ = \mathsf{EV@R}_\alpha(Y) \cdot \mathbb{E}\, Z,$$

and thus it is enough to demonstrate that $\|Z\|^* = \mathbb{E}\, Z$ for $Z = \exp(t^* Y)$. By (36), it is sufficient to prove

$$\frac{\mathbb{E}\, Z \log\left(\frac{Z}{c} \vee 1\right)}{\log \frac{1}{1-\alpha} \mathbb{E}\left(\frac{Z}{c} \vee 1\right)} \le \mathbb{E}\, Z$$

for all $c > 0$, or the stronger statement

$$\mathbb{E}\, Z \log(Z \vee c) \le \mathbb{E}\, Z \cdot \log \frac{1}{1-\alpha} \mathbb{E}(Z \vee c) \tag{41}$$

for all $c \ge 0$. Note that, by (40), $c = 0$ satisfies the latter equation. To accept (41), it is enough to demonstrate that the left hand side of (41) grows slower than its right hand side with respect to increasing c. This is certainly correct if the derivatives of (41) with respect to c satisfy the inequality

$$\mathbb{E}\, \frac{Z}{c} \mathbb{1}_{\{Z \le c\}} \le \mathbb{E}\, Z \cdot \frac{\mathbb{E}\, \mathbb{1}_{\{Z \le c\}}}{\mathbb{E}\, Z \vee c}.$$

The latter relation derives from

$$\frac{\mathbb{E}\, Z}{\mathbb{E}\, Z \vee c} = \frac{\mathbb{E}\, Z \mathbb{1}_{\{Z \le c\}} + \mathbb{E}\, Z \mathbb{1}_{\{Z > c\}}}{\mathbb{E}\, c \mathbb{1}_{\{Z \le c\}} + \mathbb{E}\, Z \mathbb{1}_{\{Z > c\}}} \ge \frac{\mathbb{E}\, Z \mathbb{1}_{\{Z \le c\}}}{\mathbb{E}\, c \mathbb{1}_{\{Z \le c\}}},$$

which is a consequence of $\mathbb{E}\, c \mathbb{1}_{\{Z \le c\}} \ge \mathbb{E}\, Z \mathbb{1}_{\{Z \le c\}}$. This completes the proof. □

The following proposition addresses and answers the converse question, i.e., given Z, what is the random variable Y to obtain equality in (38)?

Proposition 4.8. *Let $c^* > 0$ be optimal in (31) for $Z \in E^*$. Then,*

$$Y := \operatorname{sign}(Z) \cdot \log\left(\frac{|Z|}{c^*} \vee 1\right) \tag{42}$$

satisfies the equality

$$\mathbb{E}\, YZ = \|Y\| \cdot \|Z\|^*.$$

Proof. Without loss of generality, we may again assume that $Z \ge 0$. By (31) and the definition of Y in (42) we have that

$$\|Z\|^* = \frac{\mathbb{E}\, YZ}{\log \frac{1}{1-\alpha} \mathbb{E}\, e^Y},$$

and hence

$$\mathbb{E}\, YZ \le \|Z\|^* \cdot \mathsf{EV@R}_\alpha(|Y|) \le \|Z\|^* \cdot \log \frac{1}{1-\alpha} \mathbb{E}\, e^Y = \mathbb{E}\, YZ,$$

from which we conclude the assertion. □

4.4 Dual representation and extended Donsker–Varadhan formula

We derive the dual representation of $\mathsf{EV@R}$ for $Y \in E^{**}$ from the characterizations already established for the norms in the previous section. As stated in the introduction, the dual representation given in the following theorem was already shown for $Y \in L^\infty$ and $Y \in E$ in the literature.

Theorem 4.9 (Dual representation of $\mathsf{EV@R}$)**.** *For every $0 < \alpha < 1$ and $Y \in E^{**}$, the $\mathsf{EV@R}$ has the representation*

$$\mathsf{EV@R}_\alpha(Y) = \sup\left\{\mathbb{E}\, YZ : \mathbb{E}\, Z = 1,\ Z \geq 0 \textit{ and } \mathbb{E}\, Z \log Z \leq \log \frac{1}{1-\alpha}\right\}. \tag{43}$$

Proof. The proof of Proposition 4.7 can be applied to similarly show that the equation

$$\mathsf{EV@R}_\alpha(Y) = \frac{\mathbb{E}\, YZ}{\|Z\|^*}. \tag{44}$$

holds for some density $Z \in E^*$. This implies the following representation:

$$\mathsf{EV@R}_\alpha(Y) = \sup\left\{\mathbb{E}\, YZ : \mathbb{E}\, Z = 1,\ Z \geq 0 \text{ and } \|Z\|^* \leq 1\right\}.$$

From the representation (31) of the norm and (36) it follows that

$$\frac{\mathbb{E}\, Z \log\left(\frac{Z}{c} \vee 1\right)}{\log \frac{1}{1-\alpha}\, \mathbb{E}\left(\frac{Z}{c} \vee 1\right)} \leq \mathbb{E}\, Z.$$

By the same reasoning as in the previous proof, we have

$$\mathbb{E}\, Z \log\left(Z \vee c\right) \leq \mathbb{E}\, Z \cdot \log \frac{1}{1-\alpha}\, \mathbb{E}\left(Z \vee c\right)$$

for every $c \geq 0$. By setting $c = 0$ and assuming $\mathbb{E}\, Z = 1$ (which follows from the translation invariance property (P1)), we obtain

$$\mathbb{E}\, Z \log Z \leq \log \frac{1}{1-\alpha}.$$

This concludes the proof.

□

Remark 4.10. A random variable Z with $\mathbb{E}\, Z = 1$ and $Z \geq 0$ represents a density and

$$Q(B) := \mathbb{E}\, \mathbb{1}_B Z$$

is a measure, which is absolutely continuous with respect to the probability measure P. Therefore, the expression $\mathbb{E}\, Z \log Z$ represents the relative entropy of Q with resepct to P (also known as Kullback–Leibler divergence of Q from P, denoted by $D(Q \parallel P)$) (cf. (45)).

Remark 4.11 (A financial interpretation). A nice financial interpretation of $\mathsf{EV@R}(Y)$ can be given based on the dual representation presented in Theorem 4.9. Namely, $\mathsf{EV@R}(Y)$ is the largest expected value of Y over all probability measures Q which are absolutely continuous with respect to P and have a Kullback-Leibler divergence $D(Q \parallel P)$ less than some positive value (informally, all measures Q within a ball around P defined based on the relative entropy), that is,

$$\mathsf{EV@R}_\alpha(Y) = \sup_{Q \ll P} \left\{ \mathbb{E}_Q(Y) : D(Q \parallel P) = \mathbb{E}_P \left(\tfrac{dQ}{dP} \log \tfrac{dQ}{dP} \right) \leq -\log(1-\alpha) \right\}. \tag{45}$$

This gives the interpretation of $\mathsf{EV@R}(Y)$ as a model risk quantification when it is compared to $\mathbb{E}_P(Y)$ in the presence of *model ambiguity*, which has been widely studied over the past several decades in economics, finance, and statistics (see the seminal papers Ellsberg (1961) and Gilboa and Schmeidler (1989); and also Ahmadi-Javid (2012c), Breuer and Csiszár (2016), Watson and Holmes (2016), and references therein).

We further explicitly describe the maximising $Z \in E^*$ for the dual representation (43) of $\mathsf{EV@R}$.

Proposition 4.12 (The maximizing density of the dual representation of $\mathsf{EV@R}$)**.** *Suppose that* $Y \in E^{**}$. *If* $\mathsf{V@R}_\alpha(Y) < \operatorname{ess\,sup}(Y)$ *(or equivalently,* $P(Y = \operatorname{ess\,sup}(Y)) < 1-\alpha$*), the supremum in* (43) *is attained for*

$$Z^* := \frac{e^{t^* Y}}{\mathbb{E}\, e^{t^* Y}},$$

where $t^* > 0$ *is the optimal parameter for* (1)*. This density satisfies*

$$\mathbb{E}\, Z^* \log Z^* = \log \frac{1}{1-\alpha}.$$

If $\mathsf{V@R}_\alpha(Y) = \operatorname{ess\,sup}(Y)$*, or equivalently,* $P(Y = \operatorname{ess\,sup}(Y)) \geq 1-\alpha$*, the supremum is attained for*

$$Z^* = \begin{cases} P\big(Y = \operatorname{ess\,sup}(Y)\big)^{-1} & \text{if } Y = \operatorname{ess\,sup}(Y), \\ 0 & \text{else.} \end{cases}$$

Proof. From Proposition 2.11, $\mathsf{V@R}_\alpha(Y) < \operatorname{ess\,sup}(Y)$ if and only if the optimal parameter t^* for (1) is a finite value in $(0, \infty)$. Then, to obtain the equality, consider (40) in the proof of Proposition 4.7, which can be rewritten as

$$\frac{1}{t^*} \log \frac{1}{1-\alpha}\, \mathbb{E}\, e^{t^* Y} = \frac{\mathbb{E}\, Y e^{t^* Y}}{\mathbb{E}\, e^{t^* Y}} = \mathbb{E}\, Y Z^*, \tag{46}$$

and thus $\mathsf{EV@R}_\alpha(Y) = \mathbb{E}\, Y Z^*$.

It is further evident that $Z^* \geq 0$, $\mathbb{E}\, Z^* = 1$, and it holds that

$$\mathbb{E}\, Z^* \log Z^* = \mathbb{E}\, \frac{e^{t^* Y}}{\mathbb{E}\, e^{t^* Y}} \left(t^* Y - \log \mathbb{E}\, e^{t^* Y} \right).$$

By (46) it follows that

$$\mathbb{E}\, Z^* \log Z^* = \log \frac{1}{1-\alpha}\, \mathbb{E}\, e^{t^* Y} - \log \mathbb{E}\, e^{t^* Y} = \log \frac{1}{1-\alpha}.$$

If $\mathsf{V@R}_\alpha(Y) = \operatorname{ess\,sup}(Y)$, the remaining assertion follows from $\mathsf{EV@R}_\alpha(Y) = \operatorname{ess\,sup}(Y)$. ◻

Corollary 4.13 (Donsker–Varadhan variational formula)**.** *The* Donsker–Varadhan variational formula

$$\mathbb{E}\, Z \log Z = \sup \left\{ \mathbb{E}\, Y Z - \log \mathbb{E}\, e^Y :\ Y \in E \right\} \tag{47}$$

holds true; further, the latter expression is finite only for densities $Z \in E^*$. Conversely, for any $Y \in E$

$$\log \mathbb{E}\, e^Y = \sup \left\{ \mathbb{E}\, Y Z - \mathbb{E}\, Z \log Z :\ Z \in E^*,\ \mathbb{E}\, Z = 1,\ Z \geq 0 \right\}. \tag{48}$$

Remark 4.14. The equation (47) is an alternative expression for the Kullback-Leibler divergence (see Remark 4.10) based on convex conjugate duality. In the classical Donsker–Varadhan variational formula (see Lemma 1.4.3 and Appendix C.2 of Dupuis and Ellis (1997)), the supremum is taken over $Y \in L^\infty$. The formula (47) shows that the optimal value remains intact after extending the optimization domain to $Y \in E$. The dual formula (48) is also extended for a broader range of random variables $Y \in E$, while the old version is given for $Y \in L^\infty$ (see Proposition 1.4.2 of Dupuis and Ellis (1997)). It should be remarked that the latter variational formula plays a key role in weak convergence approach to large deviations.

Proof. To verify the Donsker–Varadhan variational formula (47), we define the convex and closed set

$$\zeta_\alpha := \left\{ Z \in E^* :\ \mathbb{E}\, Z = 1, Z \geq 0 \text{ and } \mathbb{E}\, Z \log Z \leq \log \frac{1}{1-\alpha} \right\} \tag{49}$$

and its support function

$$\mathbb{I}_{\zeta_\alpha}(Z) := \begin{cases} 0 & \text{if } Z \in \zeta_\alpha, \\ +\infty & \text{else.} \end{cases}$$

By the Fenchel–Moreau duality theorem (Shapiro et al. (2014, Theorem 7.5)) and the dual representation of $\mathsf{EV@R}$, we have that

$$\begin{aligned} \mathbb{I}_{\zeta_\alpha}(Z) &= \sup_{Y \in E} \mathbb{E}\, Y Z - \mathsf{EV@R}_\alpha(Y) \\ &= \sup_{t>0,\, Y \in E} \mathbb{E}\, Y Z - \frac{1}{t} \log \frac{1}{1-\alpha}\, \mathbb{E}\, e^{tY} \\ &= \sup_{Y \in E,\, t>0} \frac{1}{t} \left(\mathbb{E}\, Y Z - \log \mathbb{E}\, e^Y - \log \frac{1}{1-\alpha} \right). \end{aligned}$$

It follows that $Z \in \zeta_\alpha$ if and only if

$$\mathbb{E}\, Y Z - \log \mathbb{E}\, e^Y \leq \log \frac{1}{1-\alpha} \qquad \text{for all } Y \in E. \tag{50}$$

The Donsker–Varadhan variational formula follows now by considering all values $0 < \alpha < 1$ in (49) and (50). The converse similarly follows by convex conjugate duality. ◻

4.5 Kusuoka representation

The $\mathsf{EV@R}$ is a law invariant (version independent) coherent risk measure for which a Kusuoka representation can be obtained (Kusuoka (2001) and Pflug and Römisch (2007)). We derive the Kusuoka representation of $\mathsf{EV@R}$ from its dual representation (Theorem 4.9; see also Delbaen (2018)).

Proposition 4.15 (Kusuoka representation)**.** *The Kusuoka representation of the* $\mathsf{EV@R}$ *for* $Y \in E^{**}$ *is*

$$\mathsf{EV@R}_\alpha(Y) = \sup_\mu \int_0^1 \mathsf{AV@R}_p(Y)\mu(\mathrm{d}p), \tag{51}$$

where the supremum is among all probability measures μ *on* $[0,1)$ *for which the associated distortion function*

$$\sigma_\mu(p) := \int_0^p \frac{1}{1-u}\mu(\mathrm{d}u)$$

satisfies

$$\int_0^1 \sigma_\mu(p)\log\left(\sigma_\mu(p)\right)\mathrm{d}p \le \log\frac{1}{1-\alpha}. \tag{52}$$

The supremum in (51) is attained for the measure μ_{σ^*} *associated with the distortion function*

$$\sigma^*(p) := F_{Z^*}^{-1}(p) = \mathsf{V@R}_p(Z^*), \tag{53}$$

where Z^* *is the optimal random variable addressed in Proposition 4.12.*

Remark 4.16 (Alternative formulation)*.* The Kusuoka representation (51) can be stated alternatively by

$$\mathsf{EV@R}_\alpha(Y) = \sup_\sigma \int_0^1 \sigma(u)\cdot\mathsf{V@R}_u(Y)\mathrm{d}u, \tag{54}$$

where the supremum is taken over the set of distortion functions σ (i.e., $\sigma(\cdot)$ is a nondecreasing density on $[0,1)$) satisfying (52) (with σ in lieu of σ_μ).

Remark 4.17. The formulation (54) of $\mathsf{EV@R}$ is the rearranged formulation of the dual representation (43) in Theorem 4.9. Indeed, Y and the optimal Z^* are comonotone random variables, and the functions $\sigma(u) = \mathsf{V@R}_u(Z^*)$ and $\mathsf{V@R}_u(Y)$ represent their nondecreasing rearrangements in (54) (Pflug and Römisch (2007)). The constraint (52) finally explicitly involves the Kullback–Leibler distance for the nondecreasing density Z^*.

Proof of Proposition 4.15. The representation is immediate from the dual formula (43) and the general description given in Pflug and Pichler (2014, pp. 100) or Pichler and Shapiro (2015).

To see the optimal measure in (53), when $t^* \in (0,\infty)$, observe that

$$\mathbb{E}\, Z^* \log Z^* = \log\frac{1}{1-\alpha}$$

by the definition of Z^* and Proposition 4.12, and further that

$$\mathbb{E}\, Z^* \log Z^* = \int_0^1 \sigma^*(u) \log \sigma^*(u) \mathrm{d}u$$

by the definition of σ^* in (53). Define next the measure

$$\mu^*(A) := \sigma^*(0) \cdot \delta_0(A) + \int_A (1-u) \mathrm{d}\sigma^*(u) \quad (A \subset [0,1) \text{ is measurable})$$

by Riemann–Stieltjes integration on the unit interval.

For this measure μ^* it holds that

$$\int_0^p \frac{1}{1-u} \mu^*(\mathrm{d}u) = \sigma^*(0) + \int_0^p \frac{1}{1-u}(1-u)\mathrm{d}\sigma^*(u) = \sigma^*(p),$$

and σ^* is therefore feasible for (51). Observe finally that

$$\begin{aligned}\int_0^1 \mathsf{AV@R}_\alpha(Y)\mu^*(\mathrm{d}\alpha) &= \sigma^*(0)\, \mathsf{AV@R}_0(Y) + \int_0^1 \frac{1}{1-\alpha} \int_\alpha^1 F_Y^{-1}(u)\mathrm{d}u\, (1-\alpha)\mathrm{d}\sigma^*(\alpha) \\ &= \int_0^1 \sigma^*(u) F_Y^{-1}(u)\mathrm{d}u = \mathbb{E}\, Z^* Y = \mathsf{EV@R}_\alpha(Y)\end{aligned}$$

by Riemann–Stieltjes integration by parts. Hence, the assertion follows. □

5 Summary and outlook

This paper considers the norm and dual norm induced by Entropic Value-at-Risk (EV@R), which are proven to be equivalent at different confidence levels $0 < \alpha < 1$. The paper also studies the related vector spaces E, E^*, and E^{**}, which are called the primal, dual, and bidual entropic spaces, respectively. The primal and bidual entropic spaces are model spaces over which the EV@R norm $\|\cdot\| := \mathsf{EV@R}(|\cdot|)$ is also well-defined. The spaces E^* and E^{**} are the norm dual and bidual spaces of the primal entropic space E; they are all not reflexive. The pairs $(E, \|\cdot\|)$, $(E^*, \|\cdot\|^*)$, and $(E^{**}, \|\cdot\|)$ are Banach spaces.

Moreover, it is shown that E is an Orlicz heart with E^{**} being its related Orlicz space, on which the associated Orlicz norm (Luxemburg norm) is proven to be equivalent to the EV@R norm. Similarly, it is shown that E^* is an Orlicz heart coinciding with its corresponding Orlicz space, over which both the dual EV@R norm and its related Orlicz norm (Luxemburg norm) are equivalent. These observations alert us to the fact that the topological results developed for risk measures over the Orlicz hearts and spaces can be directly applied to the entropic spaces.

Both primal and bidual entropic spaces are subsets of the intersection of all the L^p spaces, including many unbounded random variables, while both are significantly larger than L^∞. Though L^∞ is dense in E, the space E^{**} includes unbounded random variables which are not the limit of any sequence of bounded random variables. Therefore, considering that the natural topology on $\bigcap_{p\geq 1} L^p$ is not normable, one can use the larger spaces E and E^{**} as a model space instead of L^∞ when a flexible model space is needed over which all the L^p-based risk measures are well-defined.

The formulas of the EV@R dual norm and corresponding Hahn–Banach functionals are explicitly obtained This highlights the computational and analytical advantages of using EV@R norm and its dual norm in practice when dealing with risk measures over the entropic spaces. One should note that the dual norm and Hahn–Banach functionals are not explicitly known for the Orlicz (or Luxemburg) norm that is equivalent to the EV@R norm. Using the duality results, the dual and the Kusuoka representations of EV@R, as well as their corresponding maximizing densities, are derived explicitly. In addition, the duality analysis results in an extended version of Donsker–Varadhan variational formulas and its dual which involve unbounded random variables.

The analysis presented here shows that one can develop powerful analytical tools in probability theory based on risk measures. This promising avenue from modern risk measure theory to probability theory can be more investigated in future studies. For example, following the idea used in this paper, one can develop new norms based on the rich class of φ-entropic risk measures given by

$$\mathrm{ER}_{\varphi,\beta}(Y) = \sup\{\mathbb{E}\,YZ : \mathbb{E}\,Z = 1,\ Z \geq 0 \text{ and } \mathbb{E}\,\varphi(Z) \leq \beta\} \tag{55}$$

for an increasing convex function φ with $\varphi(1) = 0$ and $\beta \geq 0$. These information-theoretic risk measures are coherent, which can be efficiently computed using the primal representation

$$\mathrm{ER}_{\varphi,\beta}(Y) = \inf_{z>0,\,\mu\in\mathbb{R}} \left\{ z\mu + z\,\mathbb{E}\,\varphi^*\left(\frac{Y}{z} - \mu + \beta\right)\right\},$$

where φ^* is the conjugate of φ (Theorem 5.1 in Ahmadi-Javid (2012a)). The norm induced by $\mathrm{ER}_{\varphi,\beta}$ is conjectured to be equivalent to the Orlicz norm (Luxemburg norm) on the corresponding Orlicz space L^{φ^*} under mild conditions.

References

Ahmadi-Javid, A. (2011). "An information-theoretic approach to constructing coherent risk measures". In: *2011 IEEE International Symposium on Information Theory Proceedings*, pp. 2125–2127.

— (2012a). "Entropic value-at-risk: A new coherent risk measure". In: *Journal of Optimization Theory and Applications* 155.3, pp. 1105–1123.

— (2012b). "Addendum to: Entropic value-at-risk: A new coherent risk measure". In: *Journal of Optimization Theory and Applications* 155.3, pp. 1124–1128.

— (2012c). "Application of information-type divergences to constructing multiple-priors and variational preferences". In: *2012 IEEE International Symposium on Information Theory Proceedings*, pp. 538–540.

Ahmadi-Javid, A. and M. Fallah-Tafti (2017). “Portfolio optimization with entropic value-at-risk”. Optimization Online.

Arens, R. (1946). “The space L^{ω} and convex topological rings”. In: *Bulletin of the American Mathematical Society* 52.10, pp. 931–935.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). “Coherent measures of risk”. In: *Mathematical Finance* 9.3, pp. 203–228.

Bell, W. C. (1977). “On the normability of the intersection of L^p spaces”. In: *Proceedings of the American Mathematical Society* 66.2, pp. 299–304.

Bellini, F., B. Klar, A. Müller, and E. Rosazza Gianin (2014). “Generalized quantiles as risk measures”. In: *Insurance: Mathematics and Economics* 54, pp. 41–48.

Bellini, F. and E. Rosazza Gianin (2012). “Haezendonck–Goovaerts risk measures and Orlicz quantiles”. In: *Insurance: Mathematics and Economics* 51.1, pp. 107–114.

Bennett, C. and R. Sharpley (1988). *Interpolation of Operators*. Academic Press.

Biagini, S. and M. Frittelli (2008). “A unified framework for utility maximization problems: An Orlicz space approach”. In: *Annals of Applied Probability* 18.3, pp. 929–966.

Breuer, T. and I. Csiszár (2013). “Information geometry in mathematical finance: Model risk, worst and almost worst scenarios”. In: *2013 IEEE International Symposium on Information Theory Proceedings*, pp. 404–408.

Breuer, T. and I. Csiszár (2016). “Measuring distribution model risk”. In: *Mathematical Finance* 26.2, pp. 395–411.

Castillo, R. E. and H. Rafeiro (2016). *An Introductory Course in Lebesgue Spaces*. CMS books in Mathematics. Springer.

Cheridito, P. and T. Li (2008). “Dual characterization of properties of risk measures on Orlicz hearts”. In: *Mathematics and Financial Economics* 2.1, pp. 29–55.

— (2009). “Risk measures on Orlicz hearts”. In: *Mathematical Finance* 19.2, pp. 189–214.

Delbaen, F. (2018). “Remark on the paper "Entropic value-at-risk: A new coherent risk measure" by Amir Ahmadi-Javid”. In: *Risk and Stochastics*. Ed. by P. Barrieu. World Scientific, Available online at arXiv preprint, arXiv:1504.00640.

Delbaen, F. and K. Owari (2016). “On convex functions on the duals of Δ_2-Orlicz spaces”. In: *arXiv preprint, arXiv:1611.06218*.

Dupuis, P. and R. S. Ellis (1997). *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley.

Ellsberg, D. (1961). “Risk, ambiguity, and the Savage axioms”. In: *The Quarterly Journal of Economics* 75.4, pp. 643–669.

Farkas, W., P. Koch-Medina, and C. Munari (2015). “Measuring risk with multiple eligible assets”. In: *Mathematics and Financial Economics* 9.1, pp. 3–27.

Gilboa, I. and D. Schmeidler (1989). “Maxmin expected utility with non-unique prior”. In: *Journal of Mathematical Economics* 18.2, pp. 141–153.

Kaas, R., M. Goovaerts, J. Dhaene, and M. Denuit (2008). *Modern Actuarial Risk Theory*. 2nd. Springer.

Kaina, M. and L. Rüschendorf (2009). “On convex risk measures on L^p-spaces”. In: *Mathematical Methods of Operations Research* 69.3, pp. 475–495.

Kalmes, T. and A. Pichler (2017). “On Banach spaces of vector-valued random variables and their duals motivated by risk measures”. In: *Banach Journal of Mathematical Analysis, to appear.*

Kiesel, S. and L. Rüschendorf (2014). “Optimal risk allocation for convex risk functionals in general risk domains”. In: *Statisticas and Risk Modeling* 31.3–4, pp. 335–365.

Krasnosel’skii, M. A. and Y. B. Rutickii (1961). *Convex Functions and Orlicz Spaces*. Noordhoff Groningen.

Kupper, M. and G. Svindland (2011). “Dual reopresentation of monotone convex functions on L^0”. In: *Proceedings of the American Mathematical Society* 139.11, pp. 4073–4086.

Kusuoka, S. (2001). “On law invariant coherent risk measures”. In: *Advances in Mathematical Economics*. Ed. by S. Kusuok and T. Maruyama. Vol. 3. Springer, pp. 83–95.

Pflug, G. Ch. and A. Pichler (2014). *Multistage Stochastic Optimization*. Springer Series in Operations Research and Financial Engineering. Springer.

Pflug, G. Ch. and W. Römisch (2007). *Modeling, Measuring and Managing Risk*. World Scientific.

Pichler, A. (2013). “The natural Banach space for version independent risk measures”. In: *Insurance: Mathematics and Economics* 53.2, pp. 405–415.

— (2014). “Insurance pricing under ambiguity”. In: *European Actuarial Journal* 4.2, pp. 335–364.

— (2017). “A quantitative comparison of risk measures”. In: *Annals of Operations Research* 254.1-2, pp. 251–275.

Pichler, A. and A. Shapiro (2015). “Minimal representations of insurance prices”. In: *Insurance: Mathematics and Economics* 62, pp. 184–193.

Pick, L., A. Kufner, O. John, and S. Fučík (2012). *Function Spaces, 1*. De Gruyter Series in Nonlinear Analysis and Applications. De Gruyter.

Postek, K., D. den Hertog, and B. Melenberg (2016). “Computationally tractable counterparts of distributionally robust constraints on risk measures”. In: *SIAM Review* 58.4, pp. 603–650.

Rao, M. (1968). “Linear functionals on Orlicz spaces: General theory”. In: *Pacific Journal of Mathematics* 25.3, pp. 553–585.

Rudin, W. (1973). *Functional Analysis*. McGraw-Hill.

Shapiro, A., D. Dentcheva, and A. Ruszczyński (2014). *Lectures on Stochastic Programming*. 2nd edn. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics.

Shapiro, A. (2017). "Distributionally robust stochastic programming". In: *SIAM Journal on Optimization, to appear.*

Svindland, G. (2009). "Subgradients of law-invariant convex risk measures on L^1". In: *Statistics & Decisions* 27.2, pp. 169–199.

Vaart, A. W. van der (1998). *Asymptotic Statistics*. Cambridge University Press.

Watson, J. and C. Holmes (2016). "Approximate models and robust decisions". In: *Statistical Science* 31.4, pp. 465–489.