Law invariant risk measures on $L^\infty(\mathbb{R}^d)$

Ivar Ekeland*

CEREMADE and Institut de Finance, Université Paris-Dauphine

Walter Schachermayer†

University of Vienna, Faculty of Mathematics

August 5, 2011

Abstract

Kusuoka (2001) has obtained explicit representation theorems for comonotone risk measures and, more generally, for law invariant risk measures. These theorems pertain, like most of the previous literature, to the case of scalar-valued risks.

Jouini-Meddeb-Touzi (2004) and Burgert-Rüschendorf (2006) extended the notion of risk measures to the vector-valued case. Recently Ekeland-Galichon-Henry (2009) obtained extensions of the above theorems of Kusuoka to this setting. Their results were confined to the regular case.

In general, Kusuoka's representation theorem for comonotone risk measures also involves a singular part. In the present work we give a full generalization of Kusuoka's theorems to the vector-valued case. The singular component turns out to have a richer structure than in the scalar case.

1 Introduction

Coherent risk measures have been intensively studied since their introduction in the seminal paper by Artzner, Delbaen, Eber and Heath[1]. We recall their

*The first author gratefully acknowledges the support of the National Science and Engineering Research Council of Canada

†The second author gratefully acknowledges financial support from the Austrian Science Fund (FWF) under grant P19456, from the European Research Council (ERC) under grant FA506041 and from the Vienna Science and Technology Fund (WWTF) under grant MA09-003. Furthermore this work was financially supported by the Christian Doppler Research Association (CDG).

definition in the slightly more general setting of convex risk measures [6], [7].

Definition 1.1. *A function $\varrho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ is a convex risk measure if*

- *(normalisation)* $\varrho(0) = 0$
- *(monotonicity)* $X \geq Y \Rightarrow \varrho(X) \leq \varrho(Y)$,
- *(cash invariance)* $\varrho(X + m) = -m + \varrho(X)$, *for* $m \in \mathbb{R}$,
- *(convexity)* $\varrho(\alpha X + (1-\alpha)Y) \leq \alpha\varrho(X) + (1-\alpha)\varrho(Y)$, *for* $0 \leq \alpha \leq 1$.

If, in addition, ϱ is positively homogeneous, i.e., $\varrho(\lambda X) = \lambda\varrho(X)$, for $\lambda \geq 0$, we say that ϱ is a coherent risk measure.

Throughout this paper $(\Omega, \mathcal{F}, \mathbb{P})$ denotes a standard probability space, i.e., free of atoms and such that $L^2(\Omega, \mathcal{F}, \mathbb{P})$ is separable. In fact, all our results hold true without this separability assumptions, but we don't want to elaborate on this level of generality which seems to be of little relevance in the applications.

A number of papers ([2],[14],[8],[5]) have pointed out that in certain situations it is desirable to pass to risk measures defined for $\mathbb{R}^d$-valued random variables $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) = L^\infty(\mathbb{R}^d)$ modeling *portfolio vectors.* Instead of $\mathbb{R}$-valued random variables $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, modeling portfolios expressed in terms of a unique numéraire, we now consider $\mathbb{R}^d$-valued bounded random variables. We refer to the above quoted papers for a discussion of the economic aspects. Here is a mathematical definition.

Definition 1.2. *Fix $d \in \mathbb{N}$. A function $\varrho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \to \mathbb{R}$ is a* convex risk measure in dimension d *if*

(i) *(normalisation)* $\varrho(0) = 0$,

(ii) *(monotonicity)* $X \geq Y \Rightarrow \varrho(X) \leq \varrho(Y)$,

(iii) *(cash invariance)* $\varrho(X + m\mathbf{e}) = -m + \varrho(X)$, *for* $m \in \mathbb{R}$,

(iv) *(convexity)* $\varrho(\alpha X + (1-\alpha)Y) \leq \alpha\varrho(X) + (1-\alpha)\varrho(Y)$, *for* $0 \leq \alpha \leq 1$.

We call ϱ a coherent risk measure in dimension d *if , in addition, we have*

(v) *(positive homogeneity)* $\varrho(\lambda X) = \lambda\varrho(X)$, *for* $\lambda \geq 0$.

The order in (ii) is the usual order in $\mathbb{R}^d$, namely $X \geq Y$ if $X_i \geq Y_i$, for $1 \leq i \leq d$. We denote by $\mathbb{R}^d_+$ the set of all vectors $x \in \mathbb{R}^d$ with $x \geq 0$, and we shall often write L^1_+ for $L^1(\mathbb{R}^d_+)$. We denote by $\mathbf{e} \in L^\infty(\mathbb{R}^d)$ the constant vector $\mathbf{e}= (1, 1, \ldots, 1)$.

The notion of cash invariance is not as obvious a generalization of the scalar case as it might seem at first glance. For example, Burgert and Rüschendorf [2] use a different (stronger) notion of cash invariance, namely

$$(iii') \qquad \varrho(X + me_i) = -m + \varrho(X), \quad \text{for } 1 \leq i \leq d, \qquad m \in \mathbb{R}. \tag{1}$$

Clearly (iii') implies (iii) (after renormalizing ϱ by the factor d). From an economic point of view there are pros and cons for adapting the point of view of (iii) or (iii') (compare [8] for an ample discussion of the economic aspects). In the present paper we do not want to elaborate on the economics but rather focus on the mathematical aspects. As (iii) is the more general concept, we have chosen (iii) as the definition of cash invariance in order to obtain results in maximal generality. We shall indicate below which specializations have to be made if one chooses definition (iii').

The following definition, due to Sh. Kusuoka, makes sense in the d-dimensional just as in the one-dimensional case.

Definition 1.3. *([11], Def. 3) A function $\varrho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \to \mathbb{R}$ is called* law invariant *if $law(X) = law(Y)$ implies that $\varrho(X) = \varrho(Y)$.*

In this paper we shall extend two well-known theorems from the one-dimensional to the d-dimensional case.

We start with Kusuoka's representation of comonotone risk measures in the one-dimensional case. Recall ([11], Def. 6) that two scalar random variables X, Y are *comonotone* if

$$(X(\omega) - X(\omega'))\,(Y(\omega) - Y(\omega')) \geq 0, \qquad \mathbb{P}(d\omega) \otimes \mathbb{P}(d\omega') - \text{a.s.}$$

and that map $\varrho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ is *comonotone* if

$$\varrho(X + Y) = \varrho(X) + \varrho(Y),$$

for any comonotone pair $X, Y \in L^\infty$.

An example of a comonotone coherent risk measure is, for $F \in L^1_+(\Omega, \mathcal{F}, \mathbb{P})$ normalized by $\mathbb{E}[F] = 1$, the function

$$\varrho_F(X) = \sup \left\{ \mathbb{E}[-\widetilde{X} F] : \widetilde{X} \sim X \right\} \tag{2}$$

where $\widetilde{X} \sim X$ means that $law(X) = law(\widetilde{X})$ (compare [11]). Note that

$$\begin{aligned} \varrho_F(X) &= sup\left\{\mathbb{E}[-\tilde{X}F] : \tilde{X} \sim X\right\} \\ &= sup\left\{\mathbb{E}[-X\tilde{F}] : \tilde{F} \sim F\right\} \\ &= sup\left\{\mathbb{E}[-\tilde{X}\tilde{F}] : \tilde{X} \sim X, \tilde{F} \sim F\right\}. \end{aligned}$$

We now rephrase Kusuoka's theorem in a form which will be suitable for the generalization to the d-dimensional case.

Theorem 1.4. *([11], Th. 7): For a law invariant convex risk measure $\varrho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ the following are equivalent.*

(i) ϱ is a comonotone risk measure.

(ii) There is $F \in L^1_+(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{E}[F] = 1$, and $0 \le s \le 1$ such that

$$\varrho(X) = s \ \operatorname{ess\,sup}(-X) + (1-s)\varrho_F(X). \tag{3}$$

(iii) ϱ is strongly coherent*, i.e. for $X, Y \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ we have*

$$\varrho(X) + \varrho(Y) = \sup_{X \sim \widetilde{X}} \varrho(\widetilde{X} + Y).$$

A thorough discussion of this remarkable theorem is postponed to Appendix A. There are several ways to extend the notion of comonotonicity from the one- to the d-dimensional case: see [5], [13],[15]. In this paper, we will concentrate on d-dimensional strong coherence, following the definition of Ekeland, Galichon and Henry [5] to extend the notion of comonotonicity from the one- to the d-dimensional case; compare the recent paper [13] which elucidates the issue. On the other hand, Ekeland, Galichon and Henry have extended the notion of strong coherence from the one- to the d-dimensional case.

Definition 1.5. *([5], Def. 2): A convex risk measure $\varrho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \to \mathbb{R}$ in dimension d is* strongly coherent *if*

$$\varrho(X) + \varrho(Y) = \sup\{\varrho(\widetilde{X} + Y) : X \sim \widetilde{X}\}, \qquad X, Y \in L^\infty(\mathbb{R}^d). \tag{4}$$

Observe that a strongly coherent risk measure ϱ is coherent. Indeed, for rational $\lambda > 0$ and $X \in L^\infty(\mathbb{R}^d)$, we quickly deduce from (4) and convexity that $\varrho(\lambda X) = \lambda \varrho(X)$; by continuity this property extends to real $\lambda > 0$. It is also obvious, by considering $Y = 0$, that strong coherence implies law invariance.

The risk measures of the form ϱ_F defined in (2) have been generalized to the d-dimensional case by Rüschendorf [14].

Definition 1.6. *([2, 14]): Let* $F \in L^1_+(\mathbb{R}^d)$ *be normalized by* $\mathbb{E}\left[|F|_{l^1_d}\right] = \mathbb{E}\left[\sum_{i=1}^d |F_i|\right] = 1$. *The* maximal correlation risk measure in the direction F *is defined as*

$$\varrho_F(X) = \sup\left\{\mathbb{E}[(-\widetilde{X}|F)] : \widetilde{X} \sim X\right\}. \tag{5}$$

where $(\cdot\,|\cdot)$ *denotes the inner product in* $\mathbb{R}^d$.

Again we note that ϱ_F only depends on the law of F.

We now can formulate the "regular version" of the extension of Kusuoka's theorem to the d-dimensional case (compare also [15, Theorem 2.2]). Recall that the Mackey topology on L^∞ is the topology of uniform convergence over all weakly compact subsets of L^1. For instance, the unit ball of $L^p, 1 < p \leq \infty$, is weakly compact in L^1, so the map $x \to \|x\|_{L^q}, 1 \leq q < \infty$, is continuous in the Mackey topology.

Theorem 1.7. *For a law invariant, convex risk measure* $\varrho : L^\infty(\mathbb{R}^d) \to \mathbb{R}$ *the following are equivalent.*

(i) ϱ *is strongly coherent and continuous with respect to the Mackey topology* $\tau(L^\infty(\mathbb{R}^d), L^1(\mathbb{R}^d))$.

(ii) There is $F \in L^1_+(\mathbb{R}^d)$ *normalized by* $\mathbb{E}\left[|F|_{l^1_d}\right] = \mathbb{E}\left[\sum_{i=1}^d |F_i|\right] = 1$ *such that*

$$\varrho(X) = \varrho_F(X).$$

Comparing this theorem to Kusuoka's Theorem 1.4 it corresponds to the case where in (3) the "singular mass" s equals zero or, equivalently, when ϱ is continuous from below (see Corr. 4.74 in [6]).

The above theorem was proved by Ekeland, Galichon and Henry [5] in the framework of $L^2(\mathbb{R}^d)$ which is in natural duality with itself (see also [15] for a simpler proof). We have reformulated this result for the space $L^\infty(\mathbb{R}^d)$ equipped with the Mackey topology induced by $L^1(\mathbb{R}^d)$ to obtain an if and only if result.

We still note that, if we define cash invariance as in (1), the above theorem remains valid, provided we change the normalization of F to $\mathbb{E}[|F_i|] = 1$, for $i = 1, \ldots, d$.

In [5] the question remained open how the generalization of Kusuoka's theorem to $\mathbb{R}^d$ reads in the general case. In other words, what takes the place of the singular part s when we extend Kusuoka's theorem to $\mathbb{R}^d$?

Denote by S^d the unit simplex in $\mathbb{R}^d$

$$S^d = \left\{ \xi \in \mathbb{R}^d_+, \qquad \sum_{i=1}^{d} \xi_i = 1 \right\}.$$

Definition 1.8. *(i) For every $\xi \in S^d$, we define the* worst case risk measure ϱ_ξ in the direction ξ *by*

$$\varrho_\xi(X) = \operatorname{ess\,sup}(-X|\xi) = \operatorname{ess\,sup}\left(-\sum_{i=1}^{d} \xi_i X_i \right).$$

(ii) More generally, for a probability measure μ on S^d, we define ϱ_μ as the "mixture"

$$\varrho_\mu(X) = \int_{S^d} \varrho_\xi(X)\ d\mu(\xi). \tag{6}$$

We shall verify below that ϱ_μ is a strongly coherent risk measure.

We now can formulate the general extension of Kusuoka's theorem to the vector-valued case which is the main result of this paper.

Theorem 1.9. *For a law invariant convex risk measure $\varrho : L^\infty(\mathbb{R}^d) \to \mathbb{R}$ in dimension d the following are equivalent.*

(i) ϱ is strongly coherent.

(ii) There is a number $0 \le s \le 1$, a probability measure μ on S^d, and a function $F \in L^1_+(\mathbb{R}^d)$, normalized by $\mathbb{E}\left[\sum_{i=1}^{d} |F_i|\right] = 1$, such that

$$\varrho(X) = s\varrho_\mu(X) + (1-s)\varrho_F(X). \tag{7}$$

Let us still discuss what happens to the above theorem if, following Rüschendorf [14], and Burgert and Rüschendorf [2], we define cash invariance by (1). We show in Remark 5.2 after the proof of Theorem 1.9 that the condition equivalent to strong coherence in the above theorem then reads as

(ii') For $i = 1, \dots, d$, there are numbers $0 \le s_i \le 1$ and functions $F_i \in L^1(\mathbb{R}_+)$ normalized by $\mathbb{E}[|F_i|] = 1$, such that,

$$\varrho(X) = \sum_{i=1}^{d} s_i \ \operatorname{ess\ sup}(-X_i) + (1-s_i)\varrho_{F_i}(X), \tag{8}$$

where $F = (F_1, \dots, F_d)$.

Note that condition (ii) is, from a mathematical point of view, more subtle than (ii'), as it involves the general μ-mixtures of the risk measures

ϱ_ξ in the direction ξ, while (ii') only involves the risk measure ϱ_{e_i} in the directions of the unit vectors e_i. This is one of the reasons why we adapted the more general notion of cash invariance in Definition 1.2.

We now pass to a second theme which again consists in a generalization of results of Kusuoka [11], Rüschendorf [14], [15] and Ekeland, Galichon, Henry [5]. Denoting by $\mathcal{P}$ the set of functions $F \in L^1(\mathbb{R}^d_+)$ normalized by $\mathbb{E}\left[\sum_{i=1}^{d} |F_i|\right] = 1$, and by $\mathcal{M}^1_+(S^d)$ the probability measures on S^d, we can state the following representation result for law invariant convex risk measures. The emphasis is on the fact that the theorem below involves a *max* rather than a *sup* (compare [14] for a version of this theorem where the *max* below is replaced by a *sup* as well as [15, Theorem 2.2], where a *max* rather than a *sup* type result is given).

Theorem 1.10. *Assume that $\varrho : L^\infty(\mathbb{R}^d) \to \mathbb{R}$ is a convex, law invariant risk measure in dimension d. Then there is a function $v : [0,1] \times \mathcal{P} \times \mathcal{M}^1_+(S^d) \to [0,\infty]$ such that*

$$\varrho(X) = \max_{(s,F,\mu)\in[0,1]\times\mathcal{P}\times\mathcal{M}^1_+(S^d)} \{s\varrho_\mu(X) + (1-s)\varrho_F(X) - v(s,F,\mu)\} \tag{9}$$

The law invariant risk measure ϱ is coherent if and only if v can be chosen to take only values in $\{0,\infty\}$.

The remainder of the paper is organized as follows. In section 2 we study law invariant, convex, closed subsets $C \subseteq \mathcal{P} \subseteq L^1(\mathbb{R}^d_+)$; they are the polar sets of law invariant coherent risk measures $\varrho = \varrho_C$ in dimension d. We identify a property, called strong coherence, of the set C which is equivalent to the strong coherence of ϱ_C. The main result of this section is Proposition 2.9: for a strongly coherent set $C \subseteq \mathcal{P}$, there is $0 \le s \le 1$ such that C uniquely decomposes as $C = (1-s)C^r + sC^s$. Here C^r is a *weakly compact* strongly coherent subset of $\mathcal{P}$, while $C^s \subseteq \mathcal{P}$ has the property that all extreme points of the σ^*-closure of C^s in $L^1(\mathbb{R}^d)^{**}$ are purely singular. This decomposition will turn out in Section 5 to correspond to the decomposition (7) in Theorem 1.9.

In Section 3 we consider the case when the set C satisfies $C = C^r$, i.e. the weakly compact case. We thus obtain a proof of Theorem 1.7.

In Section 4 we analyze the other extreme case when $C = C^s$. We then find a representation of the polar function $\varrho = \varrho_C$ as being of the form ϱ_μ (see (6)).

Finally, in Section 5, we put things together, obtaining proofs of Theorems 1.9 and 1.10.

2 Convex law invariant sets in $L^1(\mathbb{R}^d)$

In this section, we investigate closed, convex, law-invariant subsets C of $\mathcal{P} \subset L^1(\mathbb{R}^+_d)$, where $\mathcal{P}$ is the set of vector probability densities normalized by $\mathbb{E}[\sum_{i=1}^d F_i] = 1$. We shall define what it means for such a subset to be *strongly coherent* (Definition 2.3), and we show that C is strongly coherent if and only if the associated risk measure:

$$\varrho(X) = \sup_{F \in C} \langle -X, F \rangle$$

is strongly coherent. We then show that such a subset decomposes into a weighed sum:

$$C = (1-t)C^r + tC^s$$

where C^r is a weakly compact, convex, law-invariant subset of $\mathcal{P}$ and C^s is "totally singular" in a sense made precise in Proposition 2.9. Both C^r and C^s then are strongly coherent (Lemma 2.10).

We review some general functional analytic results. Fix a vector space E equipped with a locally convex topology τ. Denote by E^* its topological dual, and fix a convex, bounded subset $C \subseteq E$.

We start with a well-known result which seems to be of folklore type. Recall that a point $x \in C$ is *extremal* (or extreme) if it is not a convex combination of two different points in C.

Proposition 2.1. *Let x be an extremal point of a convex, τ-compact set $C \subseteq E$. The slices of the form*

$$S(f, \varepsilon) = \{y \in C : \langle y, f \rangle > \langle x, f \rangle - \varepsilon\}, \qquad f \in E^*, \quad \varepsilon > 0,$$

form a basis for the relative τ-neighborhoods of x in C.

Proof: As C is assumed to be τ-compact, the τ- and the weak, i.e. $\sigma(E, E^*)$-topology coincide on C. Let V be a weak neighborhood of x. There are $f_1, \ldots, f_n$ in E^* and $\varepsilon > 0$ such that

$$V \supseteq \{y \in E : \quad \langle y, f_i \rangle > \langle x, f_i \rangle - \varepsilon, \qquad i = 1, \ldots, n\}.$$

Denote by C_i the set

$$C_i = \{y \in C : \quad \langle y, f_i \rangle \le \langle x, f_i \rangle - \varepsilon\}, \quad i = 1, \ldots, n,$$

which are compact, convex subsets of C. The convex hull

$$\tilde{C} = \left\{ \sum_{i=1}^n \mu_i y_i \quad : \quad y_i \in C_i, \mu_i \ge 0, \sum_{i=1}^n \mu_i = 1 \right\}$$

is compact, convex too and does not contain the extremal point $x \in C$. Hence by Hahn-Banach we may separate x from $\tilde{C}$ by a functional $f \in E^*$ which yields the assertion.

■

Proposition 2.2. *(compare [5]): Let C be a convex, compact set in E. For a subset $K \subseteq C \times C$ the following are equivalent*

(i) $\overline{conv}(K) = C \times C$

(ii) $\overline{K} \supseteq \mathcal{E}(C) \times \mathcal{E}(C) = \mathcal{E}(C \times C)$

(iii) $\Phi_K(f, g) = \Phi_C(f) + \Phi_C(g), \qquad f, g \in E^*.$

The bar above denotes the closure with respect to the topology of $E \times E$, and $\mathcal{E}(C)$ denotes the extremal points of the set C. By Φ_C we denote the polar function of C, i.e.

$$\Phi_C(f) = \sup\{\langle x, f\rangle : x \in C\}, \qquad f \in E^*. \tag{10}$$

In other words Φ_C is the Legendre transform

$$\Phi_C(f) = \chi_C^*(f) = \sup_{x \in E}\{-\chi_C(x) + \langle x, f\rangle\},$$

of the indicator function

$$\chi_C(x) = \begin{cases} 0, & \text{for } x \in C, \\ \infty, & \text{otherwise.} \end{cases}$$

<u>*Proof:*</u> $(i) \Rightarrow (ii)$: Clearly we have $\mathcal{E}(C) \times \mathcal{E}(C) = \mathcal{E}(C \times C)$.

If (ii) were wrong, we could apply Proposition 2.1 to separate an extremal point $(x, y) \in \mathcal{E}(C \times C)$ from $\bar{K}$ by an element $(f, g) \in E^* \times E^*$ which yields a contradiction to (i).

$(ii) \Rightarrow (iii)$: Assuming (ii) we have

$$\begin{aligned} \Phi_K(f, g) &= \sup\{\langle (x, y), (f, g)\rangle : (x, y) \in K\} \\ &= \sup\{\langle x, f\rangle + \langle y, g\rangle : (x, y) \in \mathcal{E}(C) \times \mathcal{E}(C)\} \\ &= \Phi_C(f) + \Phi_C(g). \end{aligned}$$

$(iii) \Rightarrow (i)$: If (i) were false, we could find by Hahn Banach $f, g \in E^*$ such that

$$\sup\{\langle (x, y), (f, g)\rangle : (x, y) \in K\} < \sup\{\langle (x, y), (f, g)\rangle : (x, y) \in C \times C\}$$

which contradicts (iii).

$\blacksquare$

Note that for the equivalence of (i) and (iii) in Proposition 2.2 the compactness assumption is not needed. In other words, $(i) \Leftrightarrow (iii)$ holds true for closed, convex sets $C \subseteq E$.

We now consider a closed, convex subset $C \subseteq \mathcal{P}$, where $\mathcal{P}$ again denotes the set of vector probability densities

$$\mathcal{P} = \left\{ F \in L^1(\mathbb{R}^d_+) : \mathbb{E}\left[|F|_{l^1_d} \right] = \mathbb{E}\left[\sum_{i=1}^d |F_i| \right] = 1 \right\}, \tag{11}$$

which is a bounded subset of the Banach space $L^1(\mathbb{R}^d)$. In general, C will not be compact with respect to the $\sigma(L^1(\mathbb{R}^d), L^\infty(\mathbb{R}^d))$- topology. But passing to the closure $\bar{C}$ of C in the Banach space bi-dual $L^1(\mathbb{R}^d)^{**} = L^\infty(\mathbb{R}^d)^*$ with respect to the $\sigma^* := \sigma(L^\infty(\mathbb{R}^d)^*, L^\infty(\mathbb{R}^d))$-topology we always find a σ^*-compact, convex subset $\bar{C}$ of $L^\infty(\mathbb{R}^d)^*$.

Letting

$$\Phi_C(X) = \sup_{F \in C} \langle -X, F \rangle, \text{ for } X \in L^\infty(\mathbb{R}^d),$$

we find a coherent risk measure in dimension d (compare [1], [14]).

We denote by $\mathcal{T}$ the set of bijective, bi-measurable, measure preserving maps $\tau : \Omega \to \Omega$. The subsequent definition relates Proposition 2.2 with the concept of strong coherence.

Definition 2.3. *Let C be a closed, convex, law invariant subset of $\mathcal{P} \subseteq L^1(\mathbb{R}^d_+)$. Define $K \subseteq C \times C$ as*

$$K = \{(F, F \circ \tau) : F \in C, \tau \in \mathcal{T}\}.$$

We say that C is strongly coherent if

$$(SC) \qquad \overline{conv}(K) = C \times C \subset L^1_+ \times L^1_+. \tag{12}$$

It follows from Proposition 2.2 and the subsequent remark that a closed, convex, law invariant subset C of $\mathcal{P}$ is strongly coherent if and only if the risk measure $\varrho(X) = \Phi_C(-X)$ satisfies

$$\begin{aligned} \varrho(X) + \varrho(Y) &= \sup \{\langle -X, F \rangle + \langle -Y, F \circ \tau \rangle : F \in C, \tau \in \mathcal{T}\} \\ &= \sup \left\{ \langle -X - Y \circ \tau^{-1}, F \rangle : F \in C, \tau \in \mathcal{T} \right\} \\ &\leq \sup \left\{ \varrho(X + \tilde{Y}) : \tilde{Y} \sim Y \right\}. \end{aligned}$$

As the reverse inequality

$$\varrho(X) + \varrho(Y) \geq \sup\{\varrho(X + \tilde{Y}) : \tilde{Y} \sim Y\}$$

always holds true by the law invariance of C and the subsequent Proposition 2.4, we conclude that C is strongly coherent if and only if the corresponding risk measure $\varrho(X) = \Phi_C(-X)$ is strongly coherent.

We have used the following proposition which is a straightforward extension of a result of Jouini et al. ([8], Lemma A.4) in the scalar case. Its proof carries over verbatim to the vectorial case.

Proposition 2.4. *Fixing a closed, convex subset $C \subseteq \mathcal{P} \subseteq L^1(\mathbb{R}^d)$ the following are equivalent:*

(i) C is law invariant, i.e., $F \in C$ and $\widetilde{F} \sim F$ implies that $\widetilde{F} \in C$.

(i') Φ_C is law invariant, i.e., $\Phi_C(X) = \Phi_C(\widetilde{X})$, for $X \sim \widetilde{X}$.

(ii) $C_\tau := \{F \circ \tau : F \in C\} = C$, for each $\tau \in \mathcal{T}$.

(ii') $\Phi_C = \Phi_C \circ \tau$, for each $\tau \in \mathcal{T}$.

(ii") $\bar{C}_\tau := \{\beta \circ \tau : \beta \in L^\infty(\mathbb{R}^d)^*, \beta \in \bar{C}\} = \bar{C}$, for each $\tau \in \mathcal{T}$.

The notation of *(ii")* deserves some explanation: $\bar{C}$ denotes the σ^*-closure of C in $L^1(\mathbb{R}^d)^{**}$ and the functional $\beta \circ \tau$ on $L^\infty(\mathbb{R}^d)$ is defined as $\langle X, \beta \circ \tau \rangle = \langle X \circ \tau^{-1}, \beta \rangle$.

The next result again is due to Jouini et al. in the scalar case ([8], Proposition 4.1). Denote by $\bar{\mathcal{P}}$ the σ^*-closure of $\mathcal{P}$ in $L^1(\mathbb{R}^d)^{**}$.

Proposition 2.5. *Let $\widetilde{C}$ denote a non-empty, σ^*-closed, convex subset of $\bar{\mathcal{P}}$ such that $\widetilde{C}_\tau := \{\beta \circ \tau : \beta \in \widetilde{C}\} = \widetilde{C}$, for each $\tau \in \mathcal{T}$.*
Then $C := \widetilde{C} \cap L^1(\mathbb{R}^d)$ is σ^*-dense in $\widetilde{C}$.

Proof: For $\beta \in L^\infty(\mathbb{R}^d)^*$ we may define the expectation $\mathbb{E}[\beta] \in \mathbb{R}^d$ (compare [9]): consider the subspace of $L^\infty(\mathbb{R}^d)$ consisting of the constant functions which we may identify with $\mathbb{R}^d$ in an obvious way. The restriction of β to this space defines a linear functional on $\mathbb{R}^d$, namely $\mathbb{E}[\beta]$. Of course, if $\beta \in L^1(\mathbb{R}^d) \subseteq L^\infty(\mathbb{R}^d)^*$, this definition coincides with the usual definition of the expectation of a random variable.

More generally, for a finite sub sigma-algebra $\mathcal{G}$ of $\mathcal{F}$ we may, by reasoning on the atoms of $\mathcal{G}$, well-define $\mathbb{E}[\beta|\mathcal{G}]$ which is a simple function in $L^1(\mathbb{R}^d)$ (compare [9]). Observe that, for $\beta \in \overline{\mathcal{P}}$, we have $\mathbb{E}[\beta|\mathcal{G}] \in \mathcal{P}$.

Next we show that, for $\beta \in \widetilde{C}$, we have that $\mathbb{E}[\beta]$, considered as a constant $\mathbb{R}^d$-valued function, is in $\widetilde{C}$ too. To do so it will suffice to show that, for $X_1, \dots, X_K \in L^\infty(\mathbb{R}^d)$ and $\varepsilon > 0$, there is $\gamma \in \widetilde{C}$ such that

$$|\langle X_k, \gamma - \mathbb{E}[\beta]\rangle| < \varepsilon, \qquad k = 1, \dots, K.$$

Similarly as in ([9], Proof of Lemma 4.2), we find, for $\eta > 0$, natural numbers $M \leq N$ and a partition $A_1, \dots, A_N$ of Ω into $\mathcal{F}-$measurable sets of probability N^{-1} such that
(i) $osc\{X_k|A_i\} < \eta, \qquad$ for $k = 1, \dots, K$ and for $i = M+1, \dots, N$
(ii) $M/N < \eta$.

Here $osc\{X_k|A_i\}$ denotes the essential oscillation of X_k on A_i, i.e., the smallest number $a \geq 0$ such that $|X_k(\omega) - X_k(\omega')|_{l_d^\infty} \leq a$, for $\mathbb{P} \otimes \mathbb{P}$ almost all $(\omega, \omega') \in A_i$.

Continuing as in ([9], Proof of Lemma 4.2) we find, for each permutation $\pi : \{1, \dots, N\} \to \{1, \dots, N\}$, a measure preserving transformation $\tau^\pi : \Omega \to \Omega$, mapping A_i onto $A_{\pi(i)}$. Defining

$$\gamma = \frac{1}{N!} \sum_\pi \beta \circ \tau^\pi$$

we infer from the law invariance and convexity of $\widetilde{C}$ that $\gamma \in \widetilde{C}$. For each $1 \leq k \leq K$ we then may estimate

$$\left\| \frac{1}{N!} \sum_\pi X_k \circ \tau^\pi - \mathbb{E}[X_k] \right\|_{L^\infty(\mathbb{R}^d)} \leq \eta + \frac{M}{N} osc\{X\}$$

so that

$$\langle X_k, \gamma - \mathbb{E}[\beta]\rangle = \left\langle \frac{1}{N!} \sum_\pi X_k \circ \tau^\pi - \mathbb{E}[X_k], \beta \right\rangle < \eta + 2\eta \|X\|_{L^\infty(\mathbb{R}^d)}.$$

The same argument, localized to the atoms of an arbitrary finite sub-sigma-algebra $\mathcal{G}$ of $\mathcal{F}$, yields

$$\mathbb{E}[\beta|\mathcal{G}] \in \widetilde{C} \cap L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) = \widetilde{C} \cap \mathcal{P}.$$

To show that $\widetilde{C} \cap L^1(\mathbb{R}^d)$ is, in fact, σ^*-dense in $\widetilde{C}$, it now suffices to note that, for $\beta \in \widetilde{C}$, the net $(\mathbb{E}[\beta|\mathcal{G}])_\mathcal{G}$, where $\mathcal{G}$ runs through the directed set of finite sub-sigma-algebras of $\mathcal{F}$, converges to β with respect to the $\sigma^* = \sigma(L^\infty(\mathbb{R}^d)^*, L^\infty(\mathbb{R}^d))$ topology.

$\blacksquare$

Exactly as in ([9], Proof of Theorem 2.2) we quickly deduce from Proposition 2.4 and Proposition 2.5 the automatic Fatou property of a law invariant convex risk measure on $L^\infty(\mathbb{R}^d)$. We summarize this fact in the next theorem.

Theorem 2.6. *Let $\varrho : L^\infty(\mathbb{R}^d) \to \mathbb{R}$ be a* law invariant*, convex risk measure. Then ϱ is lower semi-continuous with respect to the $\sigma(L^\infty(\mathbb{R}^d), L^1(\mathbb{R}^d))$ topology. In other words, ϱ has the Fatou property or, using the terminology of [6], is continuous from above.*

Hence, defining the conjugate function ϱ_* of ϱ with respect to the dual pair $\langle L^1(\mathbb{R}^d), L^\infty(\mathbb{R}^d)\rangle$ by

$$\varrho_*(F) = \sup\left\{\langle -X, F\rangle - \varrho(X) : X \in L^\infty(\mathbb{R}^d)\right\}, \qquad F \in L^1(\mathbb{R}^d),$$

we obtain the duality formula

$$\varrho(X) = \sup\left\{\langle -X, F\rangle - \varrho_*(F) : F \in L^1(\mathbb{R}^d)\right\}, \qquad X \in L^\infty(\mathbb{R}^d).$$

The function ϱ_* takes finite values only on $\mathcal{P}$, and the convex risk measure ϱ is coherent if and only if ϱ_* only takes the values 0 and ∞.

$\blacksquare$

In the rest of this section we again consider a convex, closed set $C \subseteq \mathcal{P} \subseteq L^1(\mathbb{R}^d)$ which we now assume to be strongly coherent (Definition 2.3). We denote by $\bar{C}$ the σ^*-closure of C in $\bar{\mathcal{P}} \subseteq L^1(\mathbb{R}^d)^{**} = L^\infty(\mathbb{R}^d)^*$. Each $\beta \in \bar{\mathcal{P}}$ admits a Hahn decomposition $\beta = \beta^r + \beta^s$, where β^r is the regular part, which we identify with a function $F \in L^1_+(\mathbb{R}^d)$, and where β^s is purely singular, i.e., for $\varepsilon > 0$ there is $A \in \mathcal{F}, \mathbb{P}[A] < \varepsilon$ such that $\beta^s = \beta^s \mathbb{1}_A$. Here we define

$$\langle X, \beta^s \mathbb{1}_A\rangle := \langle X\mathbb{1}_A, \beta^s\rangle, \qquad \text{for } X \in L^\infty(\mathbb{R}^d).$$

We also define the *total variation measure* $|\beta| \in L^\infty_+(\mathbb{R})^*$ by

$$|\beta|[A] = \langle \mathbf{e}\mathbb{1}_A, \beta\rangle.$$

Lemma 2.7. *Let C be a strongly coherent subset of $\mathcal{P}$. Define the function $\sigma : \ ]0,1] \to]0,1]$ as*

$$\begin{aligned}\sigma(\delta) &= sup\left\{\mathbb{E}[\mathbb{1}_A |F|_{l^1_d}] : F \in C, \mathbb{P}[A] \le \delta\right\} \\ &= sup\left\{|\beta|[A] : \beta \in \bar{C}, \mathbb{P}[A] \le \delta\right\},\end{aligned} \tag{13}$$

and let

$$\sigma(C) := \lim_{\delta \to 0} \sigma(\delta). \tag{14}$$

For $\beta = \beta^r + \beta^s \in \bar{C}$ *we then have* $\|\beta^s\|_{L^1(\mathbb{R}^d, |\cdot|_{l^1_d})^{**}} \le \sigma(C)$*. The set*

$$\tilde{C} = \left\{ \beta = \beta^r + \beta^s \in \bar{C} : \|\beta^s\|_{L^1(\mathbb{R}^d, |\cdot|_{l^1_d})^{**}} = \sigma(C) \right\} \tag{15}$$

is σ^**-dense in* $\bar{C}$*, convex, and contains the extreme points* $\mathcal{E}(\bar{C})$ *of* $\bar{C}$*.*

Proof: Recall that C is strongly coherent if and only if the corresponding risk measure $\varrho(X) = \Phi_C(-X)$ is strongly coherent. The equality of the first and second line in (13) then follows from Theorem 2.6. Clearly $\delta \to \sigma(\delta)$ is an increasing function from $]0,1]$ to $]0,1]$ so that $\sigma(C) \in [0,1]$ is well-defined.

Let $\hat{\beta} \in \mathcal{E}(\bar{C})$ be an extreme point of $\bar{C}$ and $V(\hat{\beta})$ a relative σ^*-neighborhood of $\hat{\beta}$ in $\bar{C}$.

Defining

$$\sigma_{V(\hat{\beta})}(\delta) = sup \left\{ |\beta|[A] : \beta \in V(\hat{\beta}), \mathbb{P}[A] \le \delta \right\}, \qquad 0 < \delta \le 1,$$

we trivially obtain that

$$\sigma_{V(\hat{\beta})}(\delta) \le \sigma(\delta). \tag{16}$$

We claim that equality holds true in (16). Indeed, suppose that there is some $\delta \in]0,1]$ and $\alpha > 0$ such that

$$\sigma_{V(\hat{\beta})}(\delta) \le \sigma(\delta) - \alpha.$$

For every extreme point $\check{\beta}$ of $\bar{C}$ and every $A \in \mathcal{F}$ with $0 < \mathbb{P}[A] \le \delta$ we have

$$\check{\beta}[A] \le \sigma(\delta) - \alpha. \tag{17}$$

Indeed, by the strong coherence of C and Proposition 2.4 we can find a net $(\beta_\alpha)_{\alpha \in I}$ which σ^*-converges to $\hat{\beta}$, as well as a net $(\tau_\alpha)_{\alpha \in I}$ in $\mathcal{T}$ such that $(\beta_\alpha \circ \tau_\alpha)_{\alpha \in I}$ does σ^*-converge to $\check{\beta}$. For α big enough, we get $\beta_\alpha \in V(\hat{\beta})$ so that $\beta_\alpha[A] \le \sigma(\delta) - \alpha$, for every $A \in \mathcal{F}$ with $\mathbb{P}[A] \le \delta$. This property carries over to $\beta_\alpha \circ \tau_\alpha$ and therefore also to $\check{\beta}$, thus showing (17).

To show that $\sigma_{V(\beta)}(\delta) \ge \sigma(\delta)$, note that there is some (not necessarily extremal) $\bar{\beta} \in \bar{C}$ and an element $\bar{A} \in \mathcal{F}, 0 < \mathbb{P}[\bar{A}] \le \delta$ such that $V(\bar{\beta}) := \{\beta \in \bar{C} : |\beta|[A] > \sigma(\delta) - \frac{\alpha}{2}\}$ is a σ^*-neighborhood of $\bar{\beta}$. Since $C \subset \mathcal{P}, \bar{C}$ is compact in $(L^\infty)^*$. By Krein-Milman there are extreme points $\beta_1, \dots, \beta_n$ and convex weights $\mu_1, \dots, \mu_n$ such that

$$\sum_{i=1}^n \mu_i \beta_i \in V(\bar{\beta})$$

so that

$$\sum_{i=1}^{n} \mu_i |\beta_i|[A] > \sigma(\delta) - \frac{\alpha}{2}$$

in contradiction to (17). Hence $\sigma_{V(\hat{\beta})}(\delta) = \sigma(\delta)$, for each relative σ^*-neighborhood $V(\hat{\beta})$ of an extreme point $\hat{\beta} \in \mathcal{E}(\bar{C})$, thus showing (16)

It follows that, for every extreme point $\hat{\beta} = \hat{\beta}^r + \hat{\beta}^s \in \mathcal{E}(\bar{C})$, we have $\|\hat{\beta}^s\|_{L^1(\mathbb{R}^d,|\cdot|_{l^1_d})^{**}} = \sigma(c)$. Indeed, we may find, for $\varepsilon > 0$, a decreasing sequence $(A_n)_{n=1}^\infty$ in $\mathcal{F}$ with $\lim_{n\to\infty} \mathbb{P}[A_n] = 0$, and a decreasing sequence $V_n(\hat{\beta})$ of relative σ^*-neighborhoods of $\hat{\beta}$, such that $|\beta|[A_n] > \sigma(C) - (1 - 2^{-n})\varepsilon$, for each $\beta \in V_n(\hat{\beta})$, which readily shows that $\|\hat{\beta}^s\|_{L^1(\mathbb{R}^d,|\cdot|_{l^1_d})^{**}} \geq \sigma(C)$.

The fact that the set $\tilde{C}$ defined in (15) is σ^*-dense in $\bar{C}$ now follows from Krein-Milman: the convex combinations of the extreme points of $\bar{C}$ are σ^*-dense in $\bar{C}$.

■

We now shall decompose C into a weighted sum of a "regular" set $C^r \subseteq \mathcal{P}$ and a "purely singular" set $C^s \subseteq \mathcal{P}$. Supposing $0 < \sigma(C) < 1$ (in the cases $\sigma(C) = 0$ and $\sigma(C) = 1$ the decomposition will be trivial) and using the notation (15), define

$$C^r = \left\{ \beta^r / (1 - \sigma(C)) : \text{ there is } \beta \in \tilde{C} \text{ with } \beta = \beta^r + \beta^s \right\}, \tag{18}$$

$$\bar{C}^s = \overline{conv} \left\{ \beta^s / \sigma(C) : \text{ there is } \beta \in \tilde{C} \text{ with } \beta = \beta^r + \beta^s \right\}, \tag{19}$$

where $\overline{conv}$ denotes the σ^*-closed convex hull. Finally, let $C^s = \bar{C}^s \cap L^1(\mathbb{R}^d)$.

Lemma 2.8. *Under the above hypotheses C^r is a weakly compact, convex, law invariant subset of $\mathcal{P}$.*

Proof: Convexity and law invariance being rather obvious, let us show that C^r is uniformly integrable. This follows from the definition of $\sigma(\cdot)$. For $\beta \in \tilde{C}$ and $A \in \mathcal{F}$ we have by (13) and (14)

$$\mathbb{P}[A] < \delta \Rightarrow |\beta^r|[A] \leq \sigma(\delta) - \sigma(C).$$

As regards the closedness of C^r, let $(\beta_n)_{n=1}^\infty = (\beta_n^r + \beta_n^s)_{n=1}^\infty$ be a sequence in $\tilde{C}$ such that $(\beta_n^r)_{n=1}^\infty$ converges to β_0^r in the norm of $L^1(\mathbb{R}^d)$.

Any σ^*-cluster-point β_0 of $(\beta_n)_{n=1}^\infty$ will then be an element of $\tilde{C}$ that has a Hahn-decomposition $\beta_0 = \beta_0^s + \beta_0^r$ for some purely singular β_0^s, so that $\beta_0^r \in C^r$.

■

Proposition 2.9. *Under the above hypotheses we have*

$$C = (1 - \sigma(C))\ C^r + \sigma(C)\ C^s. \tag{20}$$

Proof: Since $C^r \subset L^1(\mathbb{R}^d)$, it is enough to prove that $\bar{C} = (1-\sigma(C))C^r + \sigma(C)\bar{C}^s$. For the set $\tilde{C}$ defined in (15) we have $\tilde{C} \subseteq (1-\sigma(C))\ C^r + \sigma(C)\ \bar{C}^s$. As the right hand side is a convex, σ^*-compact subset of $L^1(\mathbb{R}^d)^{**}$ we also have

$$\bar{C} = \overline{conv}(\tilde{C}) \subseteq (1 - \sigma(C))\ C^r + \sigma(C)\bar{C}^s. \tag{21}$$

Conversely, fix extremal elements $\hat{\beta} = \hat{\beta}^r + \hat{\beta}^s$ and $\check{\beta} = \check{\beta}^r + \check{\beta}^s$ in $\bar{C}$. We shall show that $\hat{\beta}^r + \check{\beta}^s$ is in $\bar{C}$ too. This will prove the reverse inclusion in (21). Indeed, the elements $\hat{\beta}^r$ (resp. $\check{\beta}^s$) originating from *extremal* elements $\hat{\beta}$ and $\check{\beta}$ of $\bar{C}$ in the above way, form a set whose convex hull is dense in $(1-\sigma(C))C^r$ (resp. $s\bar{C}^s$) with respect to the norm (resp. σ^*) topology.

It follows from the assumption of strong coherence of C and Proposition 2.4. that there is a net $(\hat{\beta}_\alpha)_{\alpha\in I} = (\hat{\beta}^r_\alpha + \hat{\beta}^s_\alpha)_{\alpha\in I}$ in $\tilde{C}$ which σ^*-converges to $\hat{\beta}$, as well as a net $(\tau_\alpha)_{\alpha\in I}$ in $\mathcal{T}$ such that $(\hat{\beta}_\alpha \circ \tau_\alpha)_{\alpha\in I}$ σ^*-converges to $\check{\beta}$.

Fix a decreasing sequence $(A_n)_{n=1}^\infty \in \mathcal{F}$ with $\lim_{n\to\infty} \mathbb{P}[A_n] = 0$ such that $\hat{\beta}^s = \lim_{n\to\infty} \hat{\beta}\mathbb{1}_{A_n}$, the limit now holding true in the norm topology of $L^1(\mathbb{R}^d)^{**}$. Then $(\hat{\beta}_\alpha\mathbb{1}_{A_n})_{\alpha\in I}$ σ^*-converges to $\hat{\beta}\mathbb{1}_{A_n}$, for each $n \in \mathbb{N}$. Observe that $(\hat{\beta}_\alpha\mathbb{1}_{A_n} \circ \tau_\alpha)_{\alpha\in I}$ σ^*-converges (after possibly passing to a σ^*-converging subnet) to some $\check{\beta}_n := \bar{\beta}^r_n + \check{\beta}^s$, where $\bar{\beta}^r_n$ is in $L^1(\mathbb{R}^d_+)$ satisfying $\lim_{n\to\infty} \|\bar{\beta}^r_n\|_{L^1(\mathbb{R}^d)} = 0$.

Define $\bar{\beta}_{n,\alpha}$ and $\bar{\beta}_n$ by:

$$\bar{\beta}_{n,\alpha} = \hat{\beta}_\alpha\mathbb{1}_{\Omega\setminus A_n} + \hat{\beta}_\alpha\mathbb{1}_{A_n} \circ \tau_\alpha,$$

$$\begin{aligned} \bar{\beta}_n &= \lim_{\alpha\in I} \bar{\beta}_{n,\alpha} \\ &= \hat{\beta}^r\mathbb{1}_{\Omega\setminus A_n} + \check{\beta}^s + \bar{\beta}^r_n. \end{aligned}$$

As

$$\lim_{n\to\infty} \|\hat{\beta}^r_n - \hat{\beta}^r_n\mathbb{1}_{A_n}\| + \lim_{n\to\infty} \|\bar{\beta}^r_n\| = 0$$

we readily obtain an element

$$\bar{\beta} = \lim_{n\to\infty} \bar{\beta}_n$$

in $\tilde{C}$ with Hahn decomposition $\bar{\beta} = \hat{\beta}^r + \check{\beta}^s$.

■

Lemma 2.10. *Under the above hypotheses the sets C^r as well as C^s are strongly coherent.*

The extremal points of $\bar{C}^s$ are purely singular and therefore the purely singular elements of $\bar{C}^s$ are σ^*-dense in $\bar{C}^s$.

Proof: Assume w.l.g. that $0 < \sigma(C) < 1$. Denoting by ϱ_C, ϱ_{C^r}, and ϱ_{C^s} the coherent risk measures induced by the respective sets, we infer from the preceding lemma that

$$\varrho_C = (1 - \sigma(C))\varrho_{C^r} + \sigma(C)\ \varrho_{C^s}.$$

As we assumed that the set C is strongly coherent, we have that ϱ_C is strongly coherent. This implies that ϱ_{C^r} and ϱ_{C^s} are both strongly coherent too which in turn implies the strong coherence of C^r and C^s.

The final assertion follows from Lemma 2.7.

■

Proposition 2.9 allows to separate the analysis of strongly coherent sets $C \subseteq \mathcal{P}$ into two extreme cases: either $C = C^r$ is weakly compact or $C = C^s$ is purely singular as in the previous lemma. This will be done in the next two sections.

3 The regular case

In this section we prove Theorem 1.7, where the strongly coherent risk measure $\varrho : L^\infty(\mathbb{R}^d) \to \mathbb{R}$ is Mackey continuous with respect to the dual pair $\langle L^\infty(\mathbb{R}^d), L^1(\mathbb{R}^d)\rangle$.

Proof of Theorem 1.7: $(ii) \Rightarrow (i)$: obvious.

$(i) \Rightarrow (ii)$: Given a strongly coherent, convex risk measure $\varrho : L^\infty(\mathbb{R}^d) \to \mathbb{R}$ we know by the remark after Definition 1.5 that ϱ is coherent. By Theorem 2.6 we know that there is a closed convex subset $C \subseteq \mathcal{P}$ such that

$$\varrho(X) = \sup\left\{\langle -X, F\rangle : F \in C\right\}.$$

By assumption ϱ is Mackey continuous with respect to $\langle L^\infty(\mathbb{R}^d), L^1(\mathbb{R}^d)\rangle$ which implies that C is weakly compact in $L^1(\mathbb{R}^d)$.

A classical theorem of R. Phelps [12] implies that a weakly compact subset of a Banach space is the closed convex hull of its strongly exposed points. Recall that $\widehat{F}$ is strongly exposed if there is some $\widehat{F} \in C$ and $X \in L^\infty(\mathbb{R}^d)$ such that, for $(F_n)_{n=1}^\infty \in C$ verifying

$$\lim_{n\to\infty} \langle X, F_n\rangle = \sup_{F\in C} \langle X, F\rangle$$

we have that

$$\lim_{n\to\infty} \|\widehat{F} - F_n\|_{L^1(\mathbb{R}^d)} = 0.$$

We want to show that $\varrho = \varrho_{\widehat{F}}$, i.e. for $Y \in L^\infty(\mathbb{R}^d)$

$$\varrho_{\hat{F}}(Y) := \sup\left\{\mathbb{E}[(-Y|\widetilde{F})] : \widetilde{F} \sim \widehat{F}\right\} = \varrho(Y).$$

As ϱ is strongly coherent we deduce from Proposition 2.2 and Definition 2.3 that

$$\varrho(X) + \varrho(Y) = \sup\left\{\langle -X, F\rangle + \langle -Y, F\circ\tau\rangle : F \in C, \tau \in \mathcal{T}\right\}.$$

If $(F_n, \tau_n)_{n=1}^\infty$ is a maximizing sequence in the above equation, we must have

$$\varrho(X) = \langle -X, \widehat{F}\rangle = \lim_{n\to\infty}\langle -X, F_n\rangle$$

so that $(F_n)_{n=1}^\infty$ norm-converges to $\widehat{F}$. Hence

$$\varrho(Y) = \lim_{n\to\infty} \sup\left\{\langle -Y, \widetilde{F}_n\rangle : \widetilde{F}_n \sim F_n\right\}$$

which implies

$$\varrho(Y) = \sup\left\{\langle -Y, \widetilde{F}\rangle : \widetilde{F} \sim \widehat{F}\right\}.$$

The proof of Theorem 1.7 now is complete.

■

We summarize our findings in the subsequent proposition which is a more abstract reformulation of Theorem 1.7.

Proposition 3.1. *A Mackey continuous risk measure $\varrho : L^\infty(\mathbb{R}^d) \to \mathbb{R}$ is strongly coherent if and only if there is $\widehat{F} \in \mathcal{P}$ such that*

$$\varrho = \varrho_{\widehat{F}}.$$

Defining $C = \overline{conv}\{\widehat{F}\circ\tau : \tau \in \mathcal{T}\}$ the point $\widehat{F}$ is strongly exposed in C by some $-X \in L^\infty(\mathbb{R}^d)$ and we have

$$\varrho(X) = \max\left\{\langle -X, F\rangle : F \in C\right\} = \langle -X, \widetilde{F}\rangle.$$

If $F \in C$ is another strongly exposed point in C we have

$$law(\widehat{F}) = law(F)$$

and

$$\overline{conv}(F\circ\tau : \tau \in \mathcal{T}) = C.$$

4 The purely singular case

In this section we analyze the "purely singular" case where we assume that C satisfies the "pure singularity" condition i.e. $\sigma(C) = 1$ in Lemma 2.7.

We know that the extremal points of $\bar{C}$ are purely singular. An ordinary point of $\bar{C}$ need not be singular, however, since $\bar{C}$ is σ^*-compact, we have, by the Krein-Milman theorem:

$$\bar{C} = \overline{conv}\{\beta \in \bar{C} : \beta \text{ is singular}\} \tag{22}$$

where $\overline{conv}$ denotes the σ^*-closed convex hull. We then have that every $\beta \in C$ is some sort of "integral" convex combination of purely singular measures (Choquet's theorem). We associate with any purely singular $\beta \in \bar{\mathcal{P}}$ a probability $\mu(\beta)$ on S^d (Definition 4.5). We then show that, if ϱ is the (strongly coherent) risk measure associated with $\bar{C}$

$$\varrho(X) = \sup\{\langle -X, \beta\rangle : \beta \in \bar{C}\},$$

we have $\varrho(X) = \varrho_\mu(X)$, where the right-hand side is defined by formula (6), and $\mu = \mu(\beta)$ for some purely singular $\beta \in \bar{C}$ (Proposition 4.7).

The main result of this section is the following analogue to Theorem 1.7.

Proposition 4.1. *For a law invariant, convex risk measure $\varrho : L^\infty(\mathbb{R}^d) \to \mathbb{R}$ the following are equivalent.*

(i) ϱ is strongly coherent and the polar set of ϱ

$$\bar{C} = \left\{\beta \in \bar{\mathcal{P}} : \varrho(X) \geq \langle -X, \beta\rangle, X \in L^\infty(\mathbb{R}^d)\right\} \tag{23}$$

satisfies the pure singularity condition (22).

(ii) There is a probability measure μ on S^d such that

$$\varrho(X) = \varrho_\mu(X) := \int_{S^d} \operatorname{ess\,sup}(-X, \xi)\ d\mu(\xi). \tag{24}$$

The proof will be postponed to the end of this section.

For $\beta \in L^\infty_+(\mathbb{R}^d)^*$ we define the *total variation measure* $|\beta| \in L^\infty_+(\mathbb{R})^*$ by

$$|\beta|[A] = \langle \mathbf{e}\mathbb{1}_A, \beta\rangle.$$

If $\beta \in \bar{\mathcal{P}}$ we clearly have that $|\beta|(\Omega) = 1$, hence $|\beta|$ is a normalized, positive, finitely additive measure on $(\Omega, \mathcal{F})$, vanishing on the null sets.

To a purely singular $\beta \in \bar{\mathcal{P}}$ we want to associate a Borel probability measure $\mu = \mu(\beta)$ on S^d as in Definition 1.8.

We first assume that $\beta = \beta^s$ is of the following "simple" form, corresponding to simple functions in the case of $L^1(\mathbb{R}^d)$:

$$\beta = \sum_{j=1}^{M} \xi_j \left(|\beta| \mathbb{1}_{G_j}\right) \tag{25}$$

where $(\xi_j)_{j=1}^M \in S^d$ and $\mathcal{G} = (G_1, \ldots, G_M)$ is a partition of Ω into $\mathcal{F}$-measurable sets of strictly positive measure.

Definition 4.2. *For a purely singular $\beta \in \bar{\mathcal{P}}$ of the simple form* (25) *we define the element $\mu = \mu(\beta) \in \mathcal{M}^1_+(S^d)$ as*

$$\mu = \sum_{j=1}^{M} |\beta|[G_j]\ \delta_{\xi_j}. \tag{26}$$

To extend this notion to general purely singular elements $\beta \in \bar{\mathcal{P}}$ we have to approximate β in the norm of $L^\infty(\mathbb{R}^d)^*$ by simple elements. For $\mathcal{G} = (G_1, \ldots, G_M)$ as above, where in the sequel we identify a partition $\mathcal{G}$ with the sigma-algebra generated by $\mathcal{G}$, we define the conditional expectation with respect to $|\beta|$, given $\mathcal{G}$, as

$$\beta_{\mathcal{G}} := \mathbb{E}_{|\beta|}[\beta|\mathcal{G}] := \sum_{j=1}^{M} \xi_j(|\beta| \mathbb{1}_{G_j}) \tag{27}$$

where the elements $\xi_j \in S^d$ are defined as

$$(\xi_j)_i = \frac{\langle e_i \mathbb{1}_{G_j}, \beta\rangle}{\langle \mathbf{e} \mathbb{1}_{G_j}, \beta\rangle}, \qquad i = 1, \ldots, d,$$

with the convention $\frac{0}{0} = 0$ (only those j where $|\beta|[G_j] > 0$ matter in (27) above).

For a purely singular β, the simple $\beta_{\mathcal{G}}$ is purely singular too. It is rather obvious that $\beta_{\mathcal{G}}$ converges to β along the filter of finite partitions $\mathcal{G}$ in the σ^* topology of $L^\infty(\mathbb{R}^d)^*$. In fact, we even get norm-convergence as shown by the next result.

Lemma 4.3. *Let $\beta \in \bar{\mathcal{P}} \subseteq L^\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)^*$ be a purely singular element.*

For $\varepsilon > 0$, there is a finite partition $\mathcal{G} = (G_1, \ldots, G_M)$ such that, for every refinement $\mathcal{H} = (H_1, \ldots, H_K)$ of $\mathcal{G}$ we have

$$\|\beta - \beta_{\mathcal{H}}\|_{L^\infty(\mathbb{R}^d)^*} < \varepsilon.$$

Proof: Let $\mathcal{G} = (G_1, \dots, G_M)$ be any partition of Ω into $\mathcal{F}$-measurable sets and let

$$\beta_{\mathcal{G}} = \sum_{j=1}^{M} \xi_j(|\beta| \mathbb{1}_{G_j})$$

be as in (27) above. As $\xi_j \in S^d$ we have

$$d^{-\frac{1}{2}} \leq |\xi_j|_{l^2_d} \leq 1.$$

We define the function $V(\beta_{\mathcal{G}})$ as

$$V(\beta_{\mathcal{G}}) = \sum_{j=1}^{M} |\xi_j|^2_{l^2_d}\ p_j,$$

where $p_j = |\beta|(G_j)$. Note that $V(\beta_g) \leq 1$.

Let $\mathcal{H} = (H_{1,1}, \dots, H_{1,K_1}, H_{2,1}, \dots, H_{2,K_2}, \dots, H_{M,1}, \dots, H_{M,K_M})$ be a refinement of $\mathcal{G}$ into sets of strictly positive measure such that

$$G_j = \bigcup_{k=1}^{K_j} H_{j,k}, \qquad j = 1, \dots, M.$$

We get

$$V(\beta_{\mathcal{H}}) = \sum_{j=1}^{M} \sum_{k=1}^{K_j} |\xi_{j,k}|^2_{l^2_d}\ p_{j,k}$$

where $\xi_{j,k} = \mathbb{E}_{|\beta|}[\beta | H_{j,k}]$ and $p_{j,k} = |\beta|[H_{j,k}]$. As $\sum_{k=1}^{K_j} p_{j,k}\xi_{j,k} = p_j \xi_j$ we obtain the "Pythagorean" relation

$$\begin{aligned} V(\beta_{\mathcal{H}}) &= \sum_{j=1}^{M} \sum_{k=1}^{K_j} |\xi_{j,k}|^2_{l^2_d}\ p_{j,k} \\ &= \sum_{j=1}^{M} |\xi_j|^2_{l^2_d}\ p_j + \sum_{j=1}^{M} \sum_{k=1}^{K_j} |\xi_{j,k} - \xi_j|^2_{l^2_d}\ p_{j,k} \\ &= V(\beta_{\mathcal{G}}) + \sum_{j=1}^{M} \sum_{k=1}^{K_j} |\xi_{j,k} - \xi_j|^2_{l^2_d}\ p_{j,k}. \end{aligned}$$

For $\delta > 0$ and $\mathcal{G}$ such that $V(\beta_{\mathcal{G}}) > \sup_{\mathcal{H}} V(\beta_{\mathcal{H}}) - \delta$ we conclude that the last term is smaller than δ. Noting that the diameter of S^d is $\sqrt{2}$ so that $|\xi_{j,k} - \xi_j|^2_{l^2_d} \leq 2$, we obtain that, for $\varepsilon > 0$, there is $\delta > 0$ such that

$$\sum_{j=1}^{M} \sum_{k=1}^{K_j} |\xi_{j,k} - \xi_j|^2_{l^2_d}\ p_{j,k} < \delta \Rightarrow \sum_{j=1}^{M} \sum_{k=1}^{K_j} |\xi_{j,k} - \xi_j|_{l^2_d}\ p_{j,k} = \|\beta_{\mathcal{G}} - \beta_{\mathcal{H}}\|_{L^\infty(\mathbb{R}^d)^*} < \varepsilon$$

Hence, for this choice of $\mathcal{G}$,

$$\|\beta - \beta_{\mathcal{G}}\|_{L^\infty(\mathbb{R}^d)^*} \leq \sup_{\mathcal{H}\supseteq\mathcal{G}} \|\beta_{\mathcal{H}} - \beta_G\|_{L^\infty(\mathbb{R}^d)^*} \leq \varepsilon,$$

and the same inequality holds true for every $\mathcal{H} \supseteq \mathcal{G}$.

■

Lemma 4.3 allows us to extend the notion of the measure $\mu(\beta)$ associated to a simple, purely singular $\beta \in \bar{\mathcal{P}}$ to a general purely singular element $\beta \in \bar{\mathcal{P}}$ via the following Lipschitz continuity result with respect to the Wasserstein-distance on $\mathcal{M}^1_+(S^d)$. Recall that for two probability measures μ, ν on the compact space S^d, equipped with the metric $d(\xi, \eta) = |\xi - \eta|_{l^2_d}$, the Wasserstein-distance of μ and ν is defined as

$$d(\mu, \nu) = \inf\{\|f - g\|_{L^2(\mathbb{R}^d)} : \ law(f) = \mu, \ law(g) = \nu\}.$$

Lemma 4.4. *Let β, γ be purely singular elements in $\bar{\mathcal{P}}$ of the simple form* (25)*, i.e.*

$$\beta = \sum_{j=1}^{M} \xi_j(|\beta|\mathbb{1}_{G_j}), \ \gamma = \sum_{k=1}^{K} \eta_k(|\gamma|\mathbb{1}_{H_k})$$

and μ, ν the associated measures by (26)*. Then*

$$d(\mu, \nu) \leq 5\|\beta - \gamma\|_{L^1(\mathbb{R}^d, |\cdot|_{l^1_d})^{**}}. \tag{28}$$

Proof: Passing to a common refinement of the partitions $\mathcal{G} = (G_1, \ldots, G_M)$ and $\mathcal{H} = (H_1, \ldots, H_K)$ corresponding to β and γ respectively, we may assume that $\mathcal{G} = \mathcal{H}$. We still denote this partition by $\mathcal{G} = (G_1, \ldots, G_M)$. We assume w.l.g. that $\mathbb{P}[G_j] > 0$, for each j.

Hence $\beta = \sum_{j=1}^M \xi_j|\beta|\mathbb{1}_{G_j}$ and $\gamma = \sum_{j=1}^M \eta_j|\gamma|\mathbb{1}_{G_j}$ where $(\xi_j)_{j=1}^M$ and $(\eta_j)_{j=1}^M$ are elements of S^d and $|\beta|$ and $|\gamma|$ are purely singular, normalized elements of $L^\infty_+(\mathbb{R})^*$.

Suppose that $\|\beta - \gamma\|_{L^1(\mathbb{R}^d, |\cdot|_{l^1_d})^{**}} < \varepsilon$, for $\varepsilon > 0$. Consider the restrictions $\beta|_{\mathcal{G}}$ and $\gamma|_{\mathcal{G}}$ of β and γ to the finite sigma-algebra $\mathcal{G}$. We denote the corresponding Radon-Nikodym derivatives with respect to $\mathbb{P}$ by F and G:

$$F := \frac{d\beta|_{\mathcal{G}}}{d\mathbb{P}} = \sum_{j=1}^{M} b_j\xi_j\mathbb{1}_{G_j}, \quad G := \frac{d\gamma|_{\mathcal{G}}}{d\mathbb{P}} = \sum_{j=1}^{M} c_j\eta_j\mathbb{1}_{G_j},$$

where

$$b_j = \frac{|\beta|[G_j]}{\mathbb{P}[G_j]}, \quad c_j = \frac{|\gamma|[G_j]}{\mathbb{P}[G_j]}.$$

Clearly

$$\|F-G\|_{L^1(\mathbb{R}^d,|\cdot|_{l^1_d})} \leq \|\beta-\gamma\|_{L^1(\mathbb{R}^d,|\cdot|_{l^1_d})^{**}} < \varepsilon.$$

For $a_j = \max(b_j, c_j)$ we find that $1 \leq a := \sum_{j=1}^M a_j \leq 1 + \|F-G\|_{L^1(\mathbb{R}^d,|\cdot|_{l^1_d})} < 1+\varepsilon$. The density

$$\frac{d\lambda}{d\mathbb{P}} := \sum_{j=1}^M \frac{a_j}{a} \mathbb{1}_{G_j}$$

defines a probability measure λ on $(\Omega, \mathcal{F})$. The positive weights $(p_j)_{j=1}^M$

$$p_j = \min(\lambda[G_j], b_j, c_j)$$

satisfy $1 - 2\varepsilon \leq p := \sum_{j=1}^M p_j \leq 1$.

We define the measures $\tilde{\mu}$ and $\tilde{\nu}$ of total mass $\tilde{\mu}(S^d) = \tilde{\nu}(S^d) = p$ as

$$\tilde{\mu} = \sum_{j=1}^M p_j \delta_{\xi_j} \quad \text{and} \quad \tilde{\nu} = \sum_{j=1}^M p_j \delta_{\eta_j}.$$

There is an obvious transport of the measure $\tilde{\mu} \in \mathcal{M}_+(S^d)$ to $\tilde{\nu} \in \mathcal{M}_+(S^d)$ which maps each piece $p_j\delta_{\xi_j}$ to $p_j\delta_{\eta_j}$. For the corresponding transport cost we find

$$\sum_{j=1}^M p_j |\xi_j - \eta_j|_{l^1_d} \leq \|F-G\|_{L^1(\mathbb{R}^d,|\cdot|_{l^1_d})} < \varepsilon.$$

Noting that $(S^d, |\cdot|_1)$ has diameter 2 and choosing an arbitrary transport that maps the remaining mass $\mu - \tilde{\mu}$ to $\nu - \tilde{\nu}$, we obtain from $(\mu - \tilde{\mu})(\Omega) < 2\varepsilon$ and $(\nu - \tilde{\nu})(\Omega) < 2\varepsilon$ the desired estimate (28)

$$d(\mu, \nu) < 2\,.\,2\varepsilon + \varepsilon < 5\varepsilon.$$

■

The two previous lemmas justify the following concept.

Definition 4.5. *For a purely singular $\beta \in \bar{\mathcal{P}}$ we define the Borel probability measure $\mu := \mu(\beta) \in \mathcal{M}_+^1(S^d)$ as*

$$\mu(\beta) = \lim_{\mathcal{G}} \mu(\beta_{\mathcal{G}}),$$

where $\mathcal{G}$ runs through the directed set of finite partitions $(G_1, \ldots, G_M)$ of Ω into sets G_j of strictly positive $\mathbb{P}$-measure, and the convergence takes place with respect to the Wasserstein distance on $\mathcal{M}_+^1(S^d)$.

The map $\mu(\cdot) : \beta \to \mu(\beta)$ is law invariant in the following sense: for a measure preserving transformation $\tau \in \mathcal{T}$ and β as above we have $\mu(\beta \circ \tau) = \mu(\beta)$. Indeed, it suffices to observe that τ maps the finite partitions $\mathcal{G}$ of Ω bijectively onto themselves.

Lemma 4.6. *Let $\beta \in \bar{\mathcal{P}}$ be purely singular. For $X \in L^\infty(\mathbb{R}^d)$ we have*

$$\sup_{\tau \in \mathcal{T}} \langle -X, \beta \circ \tau \rangle = \int_{S^d} \text{ess } \sup(-X|\xi) \; d\mu(\beta)(\xi). \tag{29}$$

For $\varepsilon > 0$ denote by A_ε the set

$$A_\varepsilon = \left\{ (-X|\xi) < \text{ess } \sup(-X|\xi) - \varepsilon, \; \textit{for each } \xi \in S^d \right\}. \tag{30}$$

Then, for $\varepsilon > 0$ and a maximizing sequence $(\tau_n)_{n=1}^\infty$ in (29) we have

$$\lim_{n \to \infty} |\beta \circ \tau_n|[A_\varepsilon] = 0. \tag{31}$$

<u>Proof:</u> Let $\beta \in \bar{\mathcal{P}}$ be purely singular and of the simple form (25)

$$\beta = \sum_{j=1}^{M} \xi_j \left(|\beta| \mathbb{1}_{G_j} \right) \tag{32}$$

so that

$$\mu(\beta) = \sum |\beta|[G_j] \; \delta_{\xi_j}.$$

Fix $X \in L^\infty(\mathbb{R}^d)$ and $\tau \in \mathcal{T}$. Noting that $\mu(\beta \circ \tau) = \mu(\beta)$ we find

$$\begin{aligned} \int_{S^d} \text{ess}\sup(-X, \xi) d\mu(\beta)(\xi) &= \int_{S^d} \text{ess}\sup(-X \circ \tau^{-1}, \xi) d\mu(\beta)(\xi) \\ &= \sum_{j=1}^{M} \text{ess}\sup(-X \circ \tau^{-1}, \xi_j) |\beta|[G_j] \\ &\geq \sum_{j=1}^{M} \langle -X, \beta \circ \tau \mathbb{1}_{G_j} \rangle = \langle -X, \beta \circ \tau \rangle. \end{aligned} \tag{33}$$

Applying Lemma 4.3 and 4.4 the inequality carries over to general purely singular $\beta \in \bar{\mathcal{P}}$.

To prove the reverse inequality in (29) assume again that β is of the simple form (32).

For $\varepsilon > 0$ find elements $(x_j)_{j=1}^M$ in $\mathbb{R}^d$, and disjoint sets $(B_j)_{j=1}^M$ in $\mathcal{F}$ with $\mathbb{P}[B_j] > 0$ such that

$$\text{ess sup}(-X|\xi_j) = (-x_j|\xi_j) \text{ and } \quad B_j \subseteq \{|X - x_j|_{l_d^2} < \varepsilon\}.$$

As β is purely singular we may find $A \in \mathcal{F}$, with $0 < \mathbb{P}[A] \leq \min_{1\leq j\leq M} \mathbb{P}[B_j]$ such that $\beta = \beta \mathbb{1}_A$. Let $\tau \in \mathcal{T}$ be any measure preserving transformation of Ω such that τ^{-1} maps $G_j \cap A$ into B_j. We then have

$$\begin{aligned}
\sum_{j=1}^M |\beta|[G_j] \operatorname{ess\,sup}(-X, \xi_j) &= \sum_{j=1}^M |\beta|[G_j](-x_j, \xi_j) \\
&\geq \sum_{j=1}^M \langle -X \circ \tau^{-1}, |\beta| \mathbb{1}_{A\cap G_j} \rangle - \varepsilon \\
&= \langle -X, \beta \circ \tau \rangle - \varepsilon
\end{aligned}$$

which yields

$$\int_{S^d} \operatorname{ess\,sup}(-X|\xi)\ d\mu(\beta)(\xi) \leq \sup_{\tau\in\mathcal{T}} \langle -X, \beta \circ \tau \rangle.$$

By continuity and Lemma 4.3 this relation again passes from elements of the form (32) to general $\beta \in \bar{\mathcal{P}}$ which readily shows (29).

Finally let us prove (31) where again we first assume that β is of the simple form (32). Fix $\varepsilon > 0$ and a maximizing sequence $(\tau_n)_{n=1}^\infty$ in (29) and suppose that there is $\alpha > 0$ s.t.

$$|\beta \circ \tau_n|[A_\varepsilon] \geq \alpha, \qquad n \in \mathbb{N}. \tag{34}$$

We may suppose that A_ε is an element of the sigma-algebra generated by the partition $(G_j)_{j=1}^M$ and we may split $\{1, \dots, M\}$ into $I \cup J$ such that $j \in J$ iff $G_j \subseteq A_\varepsilon$.

We then have as in (33) above, for $n \in \mathbb{N}$,

$$\begin{aligned}
&\sum_{j=1}^{M} \operatorname{ess\,sup}(-X \circ \tau_n^{-1} | \xi_j) |\beta|(G_j) \\
\geq &\sum_{j \in I} \operatorname{ess\,sup}(-X \circ \tau_n^{-1} | \xi_j) |\beta|(G_j) \\
+ &\sum_{j \in J} \left(\operatorname{ess\,sup}(-X \circ \tau_n^{-1} | \xi_j) - \varepsilon \right) |\beta|(G_j) + \varepsilon\alpha \\
\geq &\sum_{j=1}^{M} \langle -X, \beta \circ \tau_n \mathbb{1}_{G_j} \rangle + \varepsilon\alpha \\
= &\langle -X, \beta \circ \tau_n \rangle + \varepsilon\alpha,
\end{aligned}$$

which contradicts (29). So (34) is not possible with $\alpha > 0$.

■

Fix again $\bar{C}$ to be a σ^*-closed, convex, law invariant subset of $\bar{\mathcal{P}}$ satisfying the "pure singularity" condition (22). Denote by $K \subseteq \bar{C} \times \bar{C}$ the set

$$K = \left\{ (\beta, \beta \circ \tau) : \beta \in \bar{C}, \beta = \beta^s, \tau \in \mathcal{T} \right\}.$$

We suppose in the sequel that $\bar{C}$ satisfies the following strong coherence property analogous to (12)

$$(SC^s) \qquad \bar{K} \supseteq \mathcal{E}(\bar{C}) \times \mathcal{E}(\bar{C}), \tag{35}$$

where $\mathcal{E}(\bar{C})$ denotes the extreme points of $\bar{C}$ and $\bar{K}$ the σ^*-closure of K(compare Proposition 2.1).

Recall from the previous section that a decisive tool in the proof of Theorem 1.7 was the existence of $X_0 \in L^\infty(\mathbb{R}^d)$ which *strongly exposes* the weakly compact subset $C \subseteq L^1(\mathbb{R}^d)$. In the present context a somewhat analogous role is taken by elements $X_0 \in L^\infty(\mathbb{R}^d)$ described by the subsequent lemma.

Proposition 4.7. *Let $\bar{C}$ be a σ^*-closed, convex, law invariant subset of $\bar{\mathcal{P}}$ satisfying the strong coherence property* (35) *and the pure singularity property* (22). *Define $\varrho(X) = \sup\{\langle -X, \beta \rangle : \beta \in \bar{C}\}$.*

Then there is $\mu = \mu(\beta) \in \mathcal{M}_+^1(S^d)$ such that, for $X \in L^\infty(\mathbb{R}^d)$,

$$\varrho(X) = \varrho_\mu(X) = \int_{S^d} \text{ess }\sup(-X|\xi) d\mu(\xi). \tag{36}$$

Fix $X_0 \in L^\infty(\mathbb{R}^d)$ such that the support of the law of X_0 is the unit ball of $(\mathbb{R}^d, |\cdot|_2)$ and $\lim_{r\to 1} \mathbb{P}[|X_0|_{l_d^2} < r] = 1$. Then for each $\beta \in \bar{C}$ such that $\varrho(X_0) = \langle -X_0, \beta\rangle$ we have that β is purely singular and

$$\varrho(X_0) = \|\beta\|_{L^\infty(\mathbb{R}^d,|\cdot|_d^2)^*} = \int_{S^d} |\xi|_{l_d^2}\, d\mu(\xi) \tag{37}$$

where $\mu = \mu(\beta)$. For each sequence $(\beta_n)_{n=1}^\infty$ of purely singular elements in $\bar{C}$, such that

$$\varrho(X_0) = \lim_{n\to\infty} \langle -X_0, \beta_n\rangle$$

the sequence $(\mu(\beta_n))_{n=1}^\infty$ converges to μ in the Wasserstein-distance of $\mathcal{M}_+^1(S^d)$.

Proof: We start with the final assertion. Let X_0 be as above and denote by $\hat{\beta}$ an extreme point of $\bar{C}$ on which $-X_0$ attains its maximum. By (22) we have that $\hat{\beta}$ is purely singular.

Fix an increasing sequence $(\mathcal{G}_n)_{n=1}^\infty$ of finite partitions of $\mathcal{F}$ such that $(\beta_{\mathcal{G}_n})_{n=1}^\infty$ converges to $\hat{\beta}$ in norm. We also assume that the set $\{|X_0|_{l_d^2} \le 1 - \frac{1}{n}\}$ is in the sigma-algebra $\mathcal{G}_n$. Drop n in the notation for the moment and write

$$\beta_{\mathcal{G}} = \sum_{j=1}^M \xi_j(|\beta|\mathbb{1}_{\mathcal{G}_j}).$$

As in the previous lemma, but using now that the support of X_0 is the unit ball of $(\mathbb{R}^d, |\cdot|_{l_d^2})$ we find

$$\begin{aligned}\langle -X_0, \beta_{\mathcal{G}}\rangle &= \sum_{j=1}^M |\xi_j|_{l_d^2} |\beta|[G_j] \\ &= \|\beta_{\mathcal{G}}\|_{L^\infty(\mathbb{R}^d,|\cdot|_{l_d^2})^*} \\ &= \int_{S^d} |\xi|_{l_d^2}\, d\mu(\beta_{\mathcal{G}})(\xi).\end{aligned} \tag{38}$$

By writing again $\mathcal{G} = \mathcal{G}_n$ and sending n to infinity we have shown (37).

Define the σ^*-neighborhoods V_n of $\hat{\beta}$

$$\begin{aligned}V_n &= \left\{\beta \in \bar{C} : \|\pi_n(\hat{\beta} - \beta)\|_{L^1(\mathcal{G}_n,\mathbb{R}^d)} < n^{-1}\right\}, \\ &= \left\{\beta \in \bar{C} : \|\pi_n(\hat{\beta}_{\mathcal{G}} - \beta)\|_{L^1(\mathcal{G}_n,\mathbb{R}^d)} < n^{-1}\right\},\end{aligned} \tag{39}$$

where $\pi_n : L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)^{**} \to L^1(\Omega, \mathcal{G}_n, \mathbb{P}; \mathbb{R}^d)$ denotes the restriction of $\beta \in L^1(\mathbb{R}^d)^{**}$ to the finite sigma-algebra $\mathcal{G}_n$ and where we identify this restriction with the Radon-Nikodym derivative $\frac{d\beta|_{\mathcal{G}_n}}{d\mathbb{P}}$ to obtain an element of the finite-dimensional space $L^1(\mathcal{G}_n;\mathbb{R}^d)$.

By the same "Pythagorean" reasoning as in Lemma 4.3 we conclude that, for every sequence of purely singular elements $\beta_n \in V_n$, we have that $\mu(\beta_n)$ converges to $\mu(\hat{\beta})$ in the Wasserstein-distance.

Let $\hat{\beta}'$ be another extreme point of C on which X_0 attains its maximum. Again we find a sequence of σ^*-neighborhoods V_n' defined in a similar way such that, for every sequence $(\beta_n')_{n=1}^\infty$ of purely singular elements in V_n', we have that $(\mu(\beta_n'))_{n=1}^\infty$ Wasserstein-converges to $\mu(\hat{\beta}')$.

We know from hypothesis (35) that, for each $n \in \mathbb{N}$, there is $\tau_n \in \mathcal{T}$ such that, for

$$V_n \circ \tau_n = \{\beta \circ \tau_n : \beta \in V_n\}$$

we have

$$V_n \circ \tau_n \cap V_n' \neq \emptyset.$$

The above set is relatively σ^*-open in $\bar{C}$ so that there is a simple, purely singular element $\beta_n \in V_n \circ \tau_n \cap V_n'$. We must have that $\mu(\beta_n)$ is close in the Wasserstein distance to $\mu(\hat{\beta})$ as well as to $\mu(\hat{\beta}')$ which implies, by passing to the limit $n \to \infty$, that

$$\mu(\hat{\beta}) = \mu(\hat{\beta}').$$

Hence for every extreme point $\beta \in C$ on which X_0 attains its maximum, we have

$$\mu(\hat{\beta}) = \mu(\beta),$$

and there is a sequence of σ^*-neighborhoods $V_n(\beta)$ such that, for each sequence of simple, purely singular elements $\beta_n \in V_n(\beta)$ we have

$$\lim_{n\to\infty} \mu(\beta_n) = \mu(\hat{\beta}), \tag{40}$$

with respect to the Wasserstein-distance.

Now let $\bar{\beta}$ be an arbitrary, not necessarily extremal, point of $\bar{C}$, where $-X_0$ attains its maximum.

Applying again Pythagoras we find σ^*-neighborhoods V_n of $\bar{\beta}$ of the form (39) such that, for every sequence $(\beta_n)_{n=1}^\infty \in V_n$, we have that $(\mu(\beta_n))_{n=1}^\infty$ Wasserstein-converges to $\mu(\bar{\beta})$. We have to show that $\bar{\beta}$ is purely singular and $\mu(\bar{\beta}) = \mu(\hat{\beta})$. For each $n \in \mathbb{N}$, there is a finite number $\hat{\beta}_1, \ldots, \hat{\beta}_m$ of extreme points on which X_0 attains its maximum, and convex weights $\mu_1, \ldots, \mu_m$ such that $\sum_{j=1}^m \mu_j \hat{\beta}_j \in V_n$. In addition, we may find relative σ^*-neighborhoods $\hat{V}_j$ of $\hat{\beta}_j$ in C such that

$$\sum_{j=1}^m \mu_j \hat{V}_j \subseteq V_n.$$

For each $j = 1, \dots, m$, we may find a purely singular $\beta_j \in \hat{V}_j$ such that the Wasserstein distance of $\mu(\beta_j)$ to $\mu(\hat{\beta}_j)$ is smaller than $\frac{1}{n}$. Hence V_n contains a purely singular element in $\sum_{j=1}^m \mu_j \hat{V}_j$ with Wasserstein-distance to $\mu(\hat{\beta})$ less than $\frac{1}{n}$ which yields that $\bar{\beta}$ is purely singular and $\mu(\hat{\beta}) = \mu(\bar{\beta})$.

Summing up, for every $\bar{\beta}$ in the face set

$$\bar{C}_{X_0} := \left\{ \beta \in \bar{C} : \langle -X_0, \beta \rangle = \sup_{\gamma} \langle -X_0, \gamma \rangle \right\}$$

we have that $\bar{\beta}$ is purely singular, $\mu(\bar{\beta}) = \mu(\hat{\beta})$ and that, for $k \in \mathbb{N}$, there is a σ^*-neighborhood $V_k(\bar{\beta})$ such that W-dist $(\mu(\beta_k), \mu(\hat{\beta})) < k^{-1}$, for each purely singular $\beta \in V_k(\bar{\beta})$. By compactness, there is a σ^*-neighborhood U_k of the σ^*-compact face $\bar{C}_{X_0}$ such that each purely singular $\beta \in U_k$ satisfies W-dist $(\mu(\beta), \mu(\hat{\beta})) < k^{-1}$.

If $(\beta_n)_{n=1}^\infty$ is a sequence as in the assertion of Proposition 4.7. i.e.

$$\lim_{n \to \infty} \langle -X_0, \beta_n \rangle = \sup_{\beta \in C} \langle -X, \beta \rangle, \tag{41}$$

then, for fixed $k \geq 0$, we have $\beta_n \in U_k$ for n large enough so that

$$\lim_{n \to \infty} W\text{-dist } (\mu(\beta_n), \mu(\hat{\beta})) = 0.$$

Letting $\mu = \mu(\hat{\beta})$ this proves the second part of Proposition 4.7.

For the first part it follows from the fact that $\bar{C}$ is strongly coherent (see formula (35)) that, for $X \in L^\infty(\mathbb{R}^d)$ we have

$$\begin{aligned} \varrho(X_0) + \varrho(X) &= \sup \left\{ \langle X_0, \beta \rangle + \langle X, \beta \circ \tau \rangle : \beta \in C, \tau \in \mathcal{T} \right\} = \\ &= \varrho(X_0) + \lim_{n \to \infty} \langle -X, \beta_n \circ \tau_n \rangle \end{aligned}$$

for some sequence $(\beta_n)_{n=1}^\infty$ of purely singular elements of C satisfying (41). It follows by the same argument as in the proof of Theorem 1.7 that (36) holds true.

■

Proof of Proposition 4.1.: $(i) \Rightarrow (ii)$: This implication is the first assertion of Proposition 4.7.

$(ii) \Rightarrow (i)$: If $\varrho(\cdot) = \varrho_\mu(\cdot)$ is of the form (24) then clearly ϱ is strongly coherent. Hence we only have to check that $\bar{C}$ defined in (23) satisfies the pure singularity condition (22).

Note that the elements $X \in L^\infty(\mathbb{R}^d)$ such that

$$\lim_{\varepsilon \to 0} \mathbb{P}[A_\varepsilon] = 1, \tag{42}$$

where $A_\varepsilon = \{(-X|\xi) < \operatorname{ess\,sup}(-X|\xi) - \varepsilon$, for each $\xi \in S^d\}$, are norm dense in $L^\infty(\mathbb{R}^d)$. If $\bar{C}$ would fail the singularity condition (22), we could find $X \in L^\infty(\mathbb{R}^d)$ satisfying (42) such that

$$\sup\left\{\langle -X, \beta\rangle : \beta \in \bar{C}\right\} > \sup\left\{\langle -X, \beta\rangle : \beta \in \bar{C}, \beta = \beta^s\right\}. \tag{43}$$

By compactness the sup on the left hand side is a max and attaind at some $\bar{\beta} \in \bar{C}$. As in the proof of Proposition 4.7 we deduce from (42) that $\bar{\beta}$ is purely singular, a contradiction to (43) finishing the proof of Proposition 4.1.

■

5 Proofs of the theorems

We have assembled all the ingredients to show our main result.

Proof of Theorem 1.9.: The implication $(ii) \Rightarrow (i)$ being obvious let us show $(i) \Rightarrow (ii)$.

We have seen (Proposition 2.9) that, for a given strongly coherent risk measure ϱ the polar set $C \subseteq \mathcal{P} \subseteq L^1(\mathbb{R}^d)$ decomposes as

$$C = (1-\sigma)C^r + \sigma\, C^s,$$

for some $\sigma \in [0,1]$ where C^r is weakly compact in $L^1(\mathbb{R}^d_+)$, while the extreme points of $\bar{C}^s$ are purely singular.

Hence

$$\varrho = (1-\sigma)\varrho_{C^r} + \sigma\ \varrho_{C^s}$$

and the result now follows from Theorem 1.7. and Proposition 4.7.

■

Before tackling the proof of Theorem 1.10 let us sum up our findings. In the regular setting we found in Theorem 1.7 that the general form of a strongly coherent $\tau(L^\infty(\mathbb{R}^d), L^1(\mathbb{R}^d))$-continuous risk measure ϱ is $\varrho = \varrho_F$ for some $F \in L^1_+(\mathbb{R}^d)$ normalized by $\mathbb{E}[\sum_{i=1}^d |F_i|] = 1$. In fact, ϱ only depends on the law of F, and the risk measures $\varrho = \varrho_F$ as above are in one to one correspondence with the weakly compact convex subsets C of $\mathcal{P}$ such that $C \circ \tau = C$ holds true, for $\tau \in \mathcal{T}$, and such that condition (SC) defined in (12) is satisfied by C. In this case each strongly exposed point of C has the same law as F (Proposition 3.1). Conversely starting with $F \in \mathcal{P}$ as above and defining C to be the closed, convex hull of $\{F \circ \tau, \tau \in \mathcal{T}\}$ we find the compact, convex set C corresponding to ϱ_F and F is a strongly exposed point of C.

In the purely singular setting (22) we found in Proposition 4.1 that in this case the general form of a strongly coherent risk measure is of the form $\varrho = \varrho_\mu$ as in (24). These risk measures are in one to one correspondence with the law invariant (i.e. $\bar{C} = \bar{C} \circ \tau$, for $\tau \in \mathcal{T}$) convex, compact subsets $\bar{C}$ of $\bar{\mathcal{P}} \subseteq L^1(\mathbb{R}^d)^{**}$ satisfying the pure singularity condition (22) and the strong coherence property (SC^s) defined in (35). The extreme points β of $\bar{C}$ are not (necessarily) strongly exposed with respect to the norm of the Banach space $L^1(\mathbb{R}^d)^{**}$, but there is a kind of strong exposition in terms of the Wasserstein distance of the measure μ on S^d (see Propostion 4.7). Conversely, starting with a purely singular element $\beta \in \bar{\mathcal{P}}$ and defining $\bar{C}$ as the σ^*-closed, convex hull of $\{\beta \circ \tau : \tau \in \mathcal{T}\}$, we find the σ^*-compact, convex subset $\bar{C}$ corresponding to $\mu(\beta)$. We could alternatively start with $\mu \in \mathcal{M}^1_+(S^d)$ and associate to μ the strongly coherent risk measure ϱ_μ.

For the general case we isolate the following corollary to the above results which will be used in the proof of Theorem 1.10 below.

Proposition 5.1. *Let $\beta \in \bar{\mathcal{P}}$ with Hahn decomposition $\beta = (1-\sigma)F + \sigma\beta^s$, where $0 \leq \sigma \leq 1, F \in \mathcal{P}$ and β^s is a purely singular element of $\bar{\mathcal{P}}$.*

$$\begin{aligned} C_F &= \overline{conv}\{F \circ \tau : \tau \in \mathcal{T}\} \\ C_{\beta^s} &= \overline{conv}\{\beta^s \circ \tau : \tau \in \mathcal{T}\} \\ C_\beta &= \overline{conv}\{\beta \circ \tau : \tau \in \mathcal{T}\}, \end{aligned}$$

where the first closure is taken w.r. to the norm of $L^1(\mathbb{R}^d)$, and the two subsequent ones taken w.r. to the σ^*-topology of $L^1(\mathbb{R}^d)^{**}$. Then

$$C_\beta = \sigma C_F + (1-\sigma)\ C_{\beta^s} \tag{44}$$

Hence, defining $\varrho : L^\infty(\mathbb{R}^d) \to \mathbb{R}$ by

$$\varrho(X) = \sup\{\langle -X, \beta \circ \tau\rangle : \tau \in \mathcal{T}\}$$

we get

$$\varrho = (1-\sigma)\varrho_F + \sigma\varrho_{\mu(\beta^s)}. \tag{45}$$

Proof: The set C_β obviously satisfies the strong coherence property (12), hence (44) follows from Proposition 2.9. Assertion (45) now follows from Lemma 2.10, Theorem 1.7, and Proposition 4.7.

■

Let us now pass to the setting of Theorem 1.10 where we consider a convex, law invariant risk measure ϱ in dimension d. Denote by

$$\varrho^* : L^\infty(\mathbb{R}^d)^* \to [0,\infty]$$

the Legendre transform

$$\varrho^*(\beta) = \sup\{\langle -X, \beta\rangle - \varrho(X) : X \in L^\infty(\mathbb{R}^d)\}.$$

By the norm-continuity of ϱ (in fact, ϱ is Lipschitz on $L^\infty(\mathbb{R}^d)$) we obtain the reverse formula

$$\varrho(X) = \sup\{\langle -X, \beta\rangle - \varrho^*(\beta) : \beta \in L^\infty(\mathbb{R}^d)^*\}.$$

In fact, the above sup is a max. Indeed, ϱ^* is a σ^*-l.s.c. function on $L^\infty(\mathbb{R}^d)^*$ taking finite values only on $\bar{\mathcal{P}}$. Hence, for fixed $X \in L^\infty(\mathbb{R}^d)$ the function

$$\beta \to \langle -X, \beta\rangle - \varrho^*(\beta) \tag{46}$$

is σ^*-u.s.c. and bounded from above on the σ^*-compact subset $\bar{\mathcal{P}}$; it therefore attains its maximum.

<u>Proof of Theorem 1.10.:</u> Using the above notation, fix $\beta \in \bar{\mathcal{P}}$ and Hahn-decompose β as $\beta = (1-\sigma)F + \sigma\ \beta^s$, where $0 \le \sigma \le 1, F \in \bar{P}$, and β^s a purely singular element of $\bar{\mathcal{P}}$.

We may associate to β the triple $(\sigma, F, \mu(\beta^s)) \in [0,1] \times \mathcal{P} \times \mathcal{M}_+^1(S^d)$ and define $\nu(\sigma, F, \mu) := \varrho^*(\beta)$. It follows from the law invariance of ϱ and the above discussion that ν is well-defined, i.e. if $\beta' \in \bar{\mathcal{P}}$ leads to the same triple (s, F, μ), then $\varrho(\beta) = \varrho(\beta')$. In fact ν depends on F only via the law of F, but it seems notationally easier to write ν as a function of (σ, F, μ) rather then as a function of $(\sigma, \mathrm{law}(F), \mu)$.

In any case, this well-defines a function ν on $[0,1] \times \mathcal{P} \times \mathcal{M}_+^1(S^d)$. In the extreme cases we need a little care: for $\sigma = 0$ we define $\nu(0, F, \mu) = \varrho^*(F)$, for all $\mu \in \mathcal{M}_+^1(S^d)$, and, for $\sigma = 1$, we define $\nu(1, F, mu) = \varrho^*(\beta)$ for all $F \in \mathcal{P}$, where β is chosen such that $\mu(\beta) = \mu$.

For every $(\sigma, F, \mu) \in [0,1] \times \mathcal{P} \times \mathcal{M}_+^1(S^d)$ we have

$$\varrho(X) \ge (1-\sigma)\varrho_F(X) + \sigma\ \varrho_\mu(X) - \nu(\sigma, F, \mu).$$

Indeed, let $\beta = (1-\sigma)F + \sigma\ \beta^s$ be such that $\mu = \mu(\beta^s)$. Then we get from the law-invariance of X

$$\begin{aligned}\varrho(X) &\ge \sup\{\langle -X, \beta\circ\tau\rangle - \varrho^*(\beta) : \tau \in \mathcal{T}\} \\ &= (1-s)_{\varrho_F}(X) + \sigma\ \varrho_\mu(X) - \nu(\sigma, F, \mu),\end{aligned}$$

where the last line follows from Proposition 5.1. This proves the inequality "$\geq$" in (9).

For the reverse inequality fix $X \in L^\infty(\mathbb{R}^d)$. As observed in (46), there is $\bar{\beta} = (1-\bar{\sigma})\bar{F} + \bar{\sigma}\ \bar{\beta}^s$ such that

$$\begin{aligned}
\varrho(X) &= \langle -X, \bar{\beta}\rangle - \varrho^*(\bar{\beta}) \\
&= \sup\{\langle -X, \bar{\beta}\circ\tau\rangle - \varrho^*(\bar{\beta}) : \tau \in \mathcal{T}\} \\
&= (1-\bar{\sigma}\varrho_F(X) + \bar{\sigma}\varrho_{\mu(\bar{\beta})}(X) - \nu(\bar{\sigma}, \bar{F}, \mu(\bar{\beta})),
\end{aligned}$$

which readily shows (9).

The final assertion of Theorem 1.10 is standard and straight-forward to prove.

■

Remark 5.2. If the risk measure ϱ in Theorem 1.9 satisfies, following Burgert and Rüschendorf [2], the cash invariance property (iii') defined in (1) rather than (iii), then it is straighthforward to check that ϱ is strongly coherent iff each of its coordinates $(\varrho_i)_{i=1}^d$ is strongly coherent. A direct application of Kusuoka's Theorem 1.4 now yields the characterisation (ii') in (8).

A Appendix:

We now give a more detailed discussion of Theorem 1.4 which we restate for the convenience of the reader.

Theorem A.1. *([11], Th. 7): For a law invariant convex risk measure $\varrho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ the following are equivalent.*

(i) ϱ is a comonotone risk measure.

(ii) There is $F \in L^1_+(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{E}[F] = 1$, and $0 \leq s \leq 1$ such that

$$\varrho(X) = s\ \operatorname{ess\,sup}(-X) + (1-s)\varrho_F(X).$$

(iii) ϱ is strongly coherent, *i.e. for $X, Y \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ we have*

$$\varrho(X) + \varrho(Y) = \sup_{X \sim \widetilde{X}} \varrho(\widetilde{X} + Y).$$

Firstly, we note that Kusuoka also imposed the Fatou property of ϱ in the formulation of (i). This additional assumption has been shown in [8] to automatically follow from the law invariance and can simply be dropped.

Secondly, Kusuoka formulated (ii) as

$$\varrho(X) = \int_0^1 \varrho_\alpha(X) m(d\alpha), \tag{47}$$

where $\varrho_\alpha(\cdot)$ denotes expected shortfall at level $\alpha \in [0,1]$, i.e.,

$$\varrho_\alpha(X) = \sup\{\mathbb{E}[-X|A] : A \in \mathcal{F}, \mathbb{P}[A] = \alpha],$$

for $0 \leq \alpha \leq 1$ and, for $\alpha = 0$, we let

$$\varrho_0(X) = \operatorname{ess\,sup}(-X).$$

The measure m in (47) runs through all probability measures on $[0,1]$. The link to the present representation (ii) is that the *"singular mass"* s in (ii) corresponds to $m(\{0\})$ in Kusuoka's representation.

As regards the regular part, note that, if m is a Dirac measure $m = \delta_\alpha$, for some $\alpha \in]0,1]$ in (47) then $\varrho = \varrho_\alpha = \varrho_{F_\alpha}$, where F_α takes the value α^{-1} on some set of measure α and zero otherwise.

Noting that only the law of F is relevant in definition (2) and assuming w.l.g. that $(\Omega, \mathcal{F}, \mathbb{P})$ equals $[0,1]$ equipped with Lebesgue measure on the Borel sets, we may represent F_α as

$$F_\alpha(t) = \begin{cases} \alpha^{-1}, & \text{for } 0 \leq t \leq \alpha \\ 0, & \text{otherwise.} \end{cases}$$

This representation has the feature that the functions $(F_\alpha)_{\alpha \in]0,1]}$ are (pairwise) comonotone.

If m in (47) is a probability measure on $]0,1]$, the corresponding F in (3) is given by

$$F = \int_0^1 F_\alpha m(d\alpha),$$

i.e.

$$F(t) = \int_t^1 \alpha^{-1} m(d\alpha), \qquad \text{for } t \in]0,1].$$

It is straightforward to check that we then have

$$\varrho(X) = \varrho_F(X), \qquad X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}).$$

For a thorough study of the correspondence of m and F and the relation to Choquet integrals we refer to [3] and [6].

Finally let us discuss item (iii) of *strong coherence* in Theorem 1.4: it is an easy exercise to verify that (iii) is equivalent to (i) in the one-dimensional case. The notion of *strong coherence* was introduced in [5] precisely for the purpose of extending the notion of comonotone risk measure to the vector valued case.

References

[1] P. Artzner, F. Delbaen, J.-M. Eber and D. Heath, "Coherent measures of risk", Mathematical Finance, 9, 203-228 (1999)

[2] C. Burgert and L. Rüschendorf, "Consistent risk measures for portfolio vectors", Insurance: Mathematics and Economics, 38, 289-297 (2006)

[3] R-A. Dana and M. Scarsini, "Optimal risk sharing with background risk", Journal of Economic Theory, 133, 152-176 (2007)

[4] F. Delbaen and W. Schachermayer, "The mathematics of arbitrage", Springer

[5] I. Ekeland, A. Galichon and M. Henry, "Comonotonic measures of multivariate risk", Preprint, 2009, to appear in Math. Fin.

[6] H. Föllmer, A. Schied, "Stochastic Finance - An Introduction in Discrete Time", de Gruyter Studies in Mathematics 27, Walter de Gruyter - Berlin New York, (2004)

[7] M. Fritelli and E. Rosazza Gianin, "Law invariant convex risk measures", Advances in Mathematical Economics, 7, 33-46 (2005)

[8] E. Jouini, M. Meddeb and N.Touzi, "Vector-valued coherent risk measures", Finance and Stochastics 4, 531-522 (2004)

[9] E. Jouini, W. Schachermayer, N.Touzi, "Law invariant risk measures have the Fatou property", Advances in Mathematical Economics, 9, 49-71 (2006)

[10] S. Kiesel, L. Rüschendorf, "On optimal allocation of risk vectors." Preprint (2010)

[11] S. Kusuoka, "On law invariant coherent risk measures", Advances in Mathematical Economics 3, 83-95 (2001)

[12] R. R. Phelps, "Dentability and extreme points in Banach spaces", Journal of Functional Analysis 16, 78-90 (1974)

[13] G. Pucetti, M. Scarsini, "Multivariate comonotonicity", Journal of Multivariate Analysis 101, 291-304 (2010)

[14] L. Rüschendorf, "Law invariant convex risk measures for portfolio vectors", Statistics and Decisions 24, 97-108 (2006)

[15] L. Rüschendorf, "Worst case portfolio vectors and diversification effects", Preprint (2010), Online First in Finance and Stochastics (23. November 2010)