



# A concept of copula robustness and its applications in quantitative risk management

Henryk Zähle<sup>1</sup>

Received: 20 July 2021 / Accepted: 30 March 2022 / Published online: 13 September 2022  
© The Author(s) 2022

## Abstract

In financial and actuarial applications, marginal risks and their dependence structure are often modelled separately. While it is sometimes reasonable to assume that the marginal distributions are ‘known’, it is usually quite involved to obtain information on the copula (dependence structure). Therefore copula models used in practice are quite often only rough guesses. For many purposes, it is thus relevant to know whether certain characteristics derived from  $d$ -variate risks are robust with respect to (at least small) deviations in the copula. In this article, a general concept of copula robustness is introduced and criteria for copula robustness are presented. These criteria are illustrated by means of several examples from quantitative risk management. The concept of aggregation robustness introduced by Embrechts et al. (Finance Stoch. 19:763–790, 2015) can be embedded in our framework of copula robustness.

**Keywords** Copula · Fréchet class ·  $L^p$ -weak topology · Risk measure · Portfolio optimisation

**Mathematics Subject Classification (2020)** 62H05 · 60B10 · 90C17 · 91G70

**JEL Classification** C02 · C60 · G11

## 1 Introduction

As pointed out by Embrechts et al. [17] and McNeil et al. [33, Sect. 6.2.1], in financial mathematics and actuarial science, marginal risks and their dependence structure are often modelled separately. While the marginal risks of a  $d$ -variate risk are identified with probability distributions  $\mu_1, \dots, \mu_d$  on the real line, the dependence structure

---

✉ H. Zähle  
[zaehle@math.uni-sb.de](mailto:zaehle@math.uni-sb.de)

<sup>1</sup> Department of Mathematics, Saarland University, Campus E2 4, D-66123 Saarbrücken, Germany

is most often modelled by a  $d$ -variate copula  $C$ . The distribution function  $F_\mu$  of the joint distribution  $\mu$  is then given by

$$F_\mu(x_1, \dots, x_d) = C(F_{\mu_1}(x_1), \dots, F_{\mu_d}(x_d)) \quad \text{for all } x_1, \dots, x_d \in \mathbb{R}, \quad (1.1)$$

where  $F_{\mu_i}$  is the distribution function of  $\mu_i$ .

In practical applications, quantitative risk managers and actuaries are interested in various aspects  $\mathcal{T}_d(\mu)$  of the joint distribution  $\mu$  of the individual risks. An important example is  $\mathcal{T}_d = \mathcal{R}_{A_d}$  with

$$\mathcal{R}_{A_d}(\mu) := \mathcal{R}(\mu \circ A_d^{-1}), \quad (1.2)$$

where  $\mathcal{R}$  is the risk functional corresponding to some distribution-invariant ‘downside’ risk measure and  $A_d : \mathbb{R}^d \rightarrow \mathbb{R}$  is a fixed Borel-measurable map regarded as an aggregation map in the spirit of McNeil et al. [33, Sect. 6.2.1]. Standard examples for the aggregation map are  $A_d(x_1, \dots, x_d) := \sum_{i=1}^d x_i$  and the three other maps presented in Example 4.4 below. Note that  $\mu \circ A_d^{-1}$  is the distribution of  $A_d(X_1, \dots, X_d)$  when  $(X_1, \dots, X_d)$  is a random vector distributed according to  $\mu$ . Therefore  $\mathcal{R}_{A_d}(\mu)$  can be seen as the downside risk of the aggregate position  $A_d(X_1, \dots, X_d)$ .

More generally, one could consider  $\mathcal{T}_d = \mathcal{R}_{\mathfrak{A}_d}$  with

$$\mathcal{R}_{\mathfrak{A}_d}(\mu) := \inf\{\mathcal{R}(\mu \circ A_d^{-1}) : A_d \in \mathfrak{A}_d\} = \inf\{\mathcal{R}_{A_d}(\mu) : A_d \in \mathfrak{A}_d\}, \quad (1.3)$$

where  $\mathfrak{A}_d$  is a fixed set of Borel-measurable maps  $A_d : \mathbb{R}^d \rightarrow \mathbb{R}$ . If there exists an  $A_d^* \in \mathfrak{A}_d$  at which the infimum in (1.3) is attained, then  $\mathcal{R}_{\mathfrak{A}_d}(\mu)$  can be seen as the smallest possible risk of a position  $A_d(X_1, \dots, X_d)$  derived from the single risks  $X_1, \dots, X_d$  with joint distribution  $\mu$  through a function  $A_d \in \mathfrak{A}_d$ . It is worth noting that ‘risk’ here does not necessarily mean downside risk, but can also be for instance a mean–downside risk mixture which is the target value in many portfolio optimisation problems. For details, see Sect. 5.2, in particular Remark 5.5.

Of course, there are many other examples for  $\mathcal{T}_d$ . One of them is the optimal value in a multi-period portfolio optimisation problem that is addressed in Sect. 6.2. In this example, the role of  $\mu$  is played by the joint distribution of the relative price changes of the  $d$  risky assets that are available on the considered financial market.

When starting from separate models for the copula and the marginal distributions, it is reasonable to regard  $\mathcal{T}_d$  as a functional of the copula  $C$  and the marginal distributions  $\mu_1, \dots, \mu_d$  via

$$\mathfrak{T}_d(C, \mu_1, \dots, \mu_d) := \mathcal{T}_d\left(\mathfrak{p}_d\left(C(F_{\mu_1}, \dots, F_{\mu_d})\right)\right), \quad (1.4)$$

where  $\mathfrak{p}_d$  assigns to a  $d$ -variate distribution function its corresponding Borel probability measure on  $\mathbb{R}^d$ .

In [33, Sect. 6.2.1], McNeil et al. point out that practitioners are often required to work only with partial information. For instance, in some situations, it is possible to obtain (sufficient) information on  $\mu_1, \dots, \mu_d$ , but it is much more difficult to obtain information on the dependence structure. Carrying this to the extreme, McNeil et al. assume that  $\mu_1, \dots, \mu_d$  are fully known and  $C$  is fully unknown. In this case, one

cannot specify  $\mathfrak{T}_d(C, \mu_1, \dots, \mu_d)$ , because  $C$  is unknown. This leads to the ‘Fréchet problem’ of specifying the range of the map  $C \mapsto \mathfrak{T}_d(C, \mu_1, \dots, \mu_d)$ . In the special case where  $\mathfrak{T}_d$  takes values in  $\mathbb{R}$ , this is often related to finding (sharp) upper and lower bounds for this map. There is a vast literature dealing with this problem; see for instance the works of Rüschendorf [43], [44, Chap. 4], Embrechts and Puccetti [14], Embrechts et al. [15], Puccetti [39], Embrechts et al. [17] and the references cited therein.

In the present paper, a related but different problem is addressed. Still in the case where  $\mu_1, \dots, \mu_d$  are known (and fixed), assume that  $\hat{C}$  is a guess for the true copula  $C$ . It might be based on an expert opinion, a statistical estimation, or the like. Of course, as a guess,  $\hat{C}$  can differ from  $C$ . It is clear that a deviation of  $\hat{C}$  from  $C$  can imply a significant difference between  $\mathfrak{T}_d(\hat{C}, \mu_1, \dots, \mu_d)$  and  $\mathfrak{T}_d(C, \mu_1, \dots, \mu_d)$ . On the other hand, one might ask whether the difference remains small if the deviation of  $\hat{C}$  from  $C$  is small. This question was raised and answered by Embrechts et al. [17] in the context of (1.2) with  $A_d(x_1, \dots, x_d) := \sum_{i=1}^d x_i$ . Krätschmer et al. [28, Sect. 4.2.4] took up this concept and generalised the respective result of [17]. In fact, in the latter two references, continuity of the functional  $\mathcal{T}_d$  at the probability measure  $\mathfrak{p}_d(C(F_{\mu_1}, \dots, F_{\mu_d}))$  (with fixed marginal distributions  $\mu_1, \dots, \mu_d$  having finite  $p$ th moments) was not considered with respect to a metric on the set of copulas, but with respect to the (relative) weak topology on the set of  $d$ -variate distributions (with marginal distributions  $\mu_1, \dots, \mu_d$ ). However, it can be seen from Theorem 3.10 below that this is equivalent when the set of copulas is equipped with the supremum distance.

Despite this equivalence, it might be a little more accessible for some readers to measure the difference between two dependence structures directly through the difference between the corresponding copulas, in particular if one starts from separate models for the copula and the marginal distributions. If one follows this approach, one ought to take into account that a  $d$ -variate distribution  $\mu$  with fixed marginal distributions  $\mu_1, \dots, \mu_d$  depends on the copula  $C$  only through the values that  $C$  takes on  $\text{ran} F_{\mu_1} \times \dots \times \text{ran} F_{\mu_d} (\subseteq [0, 1]^d)$ , where  $\text{ran} F_{\mu_i}$  is the range of  $F_{\mu_i}$ . This is apparent from (1.1) and suggests to measure the distance between copulas (in the considered framework) only on  $\text{ran} F_{\mu_1} \times \dots \times \text{ran} F_{\mu_d}$ .

We propose to say that the functional  $\mathcal{T}_d$  underlying  $\mathfrak{T}_d$  (recall Eq. (1.4)) is *copula robust* if for any ‘admissible’ univariate distributions  $\mu_1, \dots, \mu_d$ , the map  $C \mapsto \mathfrak{T}_d(C, \mu_1, \dots, \mu_d)$  is continuous with respect to pointwise (or uniform) convergence on  $\text{ran} F_{\mu_1} \times \dots \times \text{ran} F_{\mu_d}$ , where it is assumed that  $\mathcal{T}_d$  (and thus  $\mathfrak{T}_d$ ) takes values in a topological space. By ‘admissible’ we mean that one can find at least one copula  $C$  such that the probability measure  $\mathfrak{p}_d(C(F_{\mu_1}, \dots, F_{\mu_d}))$  is contained in the domain of  $\mathcal{T}_d$ . The precise definition of copula robustness is given in Sect. 3. The required notation and terminology as well as some auxiliary results are given before in Sect. 2. It is worth mentioning that Theorem 2.3 provides a generalisation of Deheuvels’ [10] copula convergence theorem and that Corollary 2.9 provides a characterisation of weak convergence in Fréchet classes of  $d$ -variate distributions.

In the second part of the paper, we discuss three examples for copula robust functionals  $\mathcal{T}_d$ . First, in Sect. 4, we address the quantification of the ‘downside risk’ of aggregate financial positions. It will be seen that the functional in (1.2) is copula

robust under mild assumptions (Sect. 4.2). The relation of copula robustness to the concept of aggregation robustness of Embrechts et al. [17] (Sect. 4.3) as well as copula robustness of inf-convolution functionals (Sect. 4.4) are also discussed in detail. Second, in Sect. 5, we address stochastic programming problems. It can be inferred from results of Claus et al. [9] that the optimal value of a general stochastic programming problem depends copula robustly on the distribution of the underlying  $d$ -variate input random variable  $Z$ . This covers in particular classical one-period portfolio optimisation problems (where the role of  $Z$  is played by the vector of the relative price changes of  $d$  risky assets) and therefore backs in a way a hypothesis of Saida and Prigent [45]. In [45, Sect. 1], they conclude from their numerical investigations that ‘investors must more take care of the specification of the marginal distribution than of the copula function’. Third, in Sect. 6, we address multi-period portfolio optimisation problems and derive results that are similar to those in the one-period case. The main tool in this context is Theorem 6.2 which is a variant of a result of Müller [35] about the continuous dependence of the value function on the transition function in a Markov decision model. Theorem 6.2 is of independent interest and contributes to the general theory of Markov decision processes.

Throughout this paper,  $\|\cdot\|$  denotes any norm on  $\mathbb{R}^d$  and  $\langle \cdot, \cdot \rangle$  is the Euclidean scalar product defined by  $\langle x, y \rangle := \sum_{i=1}^d x_i y_i$  for any elements  $x = (x_1, \dots, x_d)$  and  $y = (y_1, \dots, y_d)$  of  $\mathbb{R}^d$ . Moreover, we set  $\mathbb{R}_+ := [0, \infty)$  and  $\mathbb{R}_{++} := (0, \infty)$ . The proofs of all results can be found in Appendix A.

## 2 Preliminary notation, terminology and results

### 2.1 Fréchet classes and copulas

For any  $d \in \mathbb{N}$ , let us use  $\mathcal{M}_d$  to denote the set of all Borel probability measures on  $\mathbb{R}^d$ . For any fixed  $\mu_1, \dots, \mu_d \in \mathcal{M}_1$ , denote by  $\mathcal{M}_d(\mu_1, \dots, \mu_d)$  the set of all  $\mu \in \mathcal{M}_d$  having marginals  $\mu_1, \dots, \mu_d$ , i.e., satisfying  $\mu \circ \pi_i^{-1} = \mu_i$  for any  $i = 1, \dots, d$ , where  $\pi_i : \mathbb{R}^d \rightarrow \mathbb{R}$  is the projection on the  $i$ th coordinate. The set  $\mathcal{M}_d(\mu_1, \dots, \mu_d)$  is known as *Fréchet class* associated with the univariate Borel probability measures  $\mu_1, \dots, \mu_d$ . The distribution function of a Borel probability measure  $\mu$  will be denoted by  $F_\mu$ .

By definition a  $d$ -variate copula is the distribution function  $C : [0, 1]^d \rightarrow [0, 1]$  of a Borel probability measure on  $[0, 1]^d$  whose marginal distributions are all given by the uniform distribution on  $[0, 1]$ . The latter condition ensures that each  $d$ -variate copula  $C$  is Lipschitz-continuous. Theorem 2.10.7 in Nelsen’s textbook [36] indeed shows that every  $d$ -variate copula  $C$  satisfies  $|C(u) - C(v)| \leq \|u - v\|_1$ , where  $\|x\|_1 := \sum_{i=1}^d |x_i|$  for any  $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ .

Let us denote by  $\mathbf{C}_d$  the set of all  $d$ -variate copulas. With any  $C \in \mathbf{C}_d$  and  $\mu_1, \dots, \mu_d \in \mathcal{M}_1$ , we associate an element  $\mu$  of the Fréchet class  $\mathcal{M}_d(\mu_1, \dots, \mu_d)$  through (1.1). It is indeed easily seen that the right-hand side of (1.1) defines a probability distribution function on  $\mathbb{R}^d$  and that the corresponding Borel probability measure  $\mu$  has  $\mu_1, \dots, \mu_d$  as its marginal distributions. Sklar’s theorem ([49]; see also [36, Theorem 2.10.9]) shows that the distribution function of any

element  $\mu$  of  $\mathcal{M}_d(\mu_1, \dots, \mu_d)$  admits the representation (1.1). That is, for any  $\mu \in \mathcal{M}_d(\mu_1, \dots, \mu_d)$ , one can find a copula  $C \in \mathbf{C}_d$  such that (1.1) holds. On the set  $\text{ran} F_{\mu_1} \times \dots \times \text{ran} F_{\mu_d}$ , the copula  $C$  is uniquely determined and given by

$$C(u_1, \dots, u_d) = F_{\mu}(F_{\mu_1}^{\leftarrow}(u_1), \dots, F_{\mu_d}^{\leftarrow}(u_d)), \quad (2.1)$$

where  $F_{\mu_i}^{\leftarrow}(u_i) := \inf\{x \in \mathbb{R} : F_{\mu_i}(x) \geq u_i\}$ . In particular, if  $F_{\mu_1}, \dots, F_{\mu_d}$  are all continuous, then the copula  $C$  is unique and given by (2.1) on the whole unit cube  $[0, 1]^d$ . For background on copulas, see for instance the textbooks by Durante and Sempi [12, Chaps. 1–2] or Nelsen [36, Chaps. 1–2].

For any nonempty compact set  $K \subseteq [0, 1]^d$ , we can define a pseudo-metric  $d_K$  on  $\mathbf{C}_d$  through

$$d_K(C_1, C_2) := \sup_{u \in K} |C_1(u) - C_2(u)|.$$

Since the elements of  $\mathbf{C}_d$  are all Lipschitz-continuous with Lipschitz constant 1 on  $[0, 1]^d$ , the set  $\mathbf{C}_d$  is uniformly equicontinuous. This implies that convergence of a sequence  $(C_n)_{n \in \mathbb{N}} \in \mathbf{C}_d^{\mathbb{N}}$  to some  $C \in \mathbf{C}_d$  with respect to  $d_K$  is equivalent to pointwise convergence of  $(C_n)_{n \in \mathbb{N}}$  to  $C$  on  $K$ . The topology on  $\mathbf{C}_d$  generated by  $d_K$  is denoted by  $\mathcal{O}_K$ . For any  $\mu_1, \dots, \mu_d \in \mathcal{M}_1$ , we let

$$d_{\mu_1, \dots, \mu_d} := d_K \quad \text{and} \quad \mathcal{O}_{\mu_1, \dots, \mu_d} := \mathcal{O}_K \quad \text{with} \quad K := \overline{\text{ran} F_{\mu_1}} \times \dots \times \overline{\text{ran} F_{\mu_d}}. \quad (2.2)$$

For  $K = [0, 1]^d$ , the pseudo-metric  $d_K$  is even a metric, and the topology  $\mathcal{O}_K$  is the standard topology on  $\mathbf{C}_d$  (and the counterpart of the weak topology on the set of all Borel probability measures on  $[0, 1]^d$  whose distribution functions are  $d$ -variate copulas). In particular, if  $F_{\mu_1}, \dots, F_{\mu_d}$  are all continuous, then  $d_{\mu_1, \dots, \mu_d} = d_{[0, 1]^d}$  and  $\mathcal{O}_{\mu_1, \dots, \mu_d} = \mathcal{O}_{[0, 1]^d}$ .

Convergence of copulas with respect to  $\mathcal{O}_{[0, 1]^d}$  has been addressed in the literature several times, for instance by Charpentier and Segers [7] and Trutschnig [51]. Metrics inducing topologies that are at least as fine as  $\mathcal{O}_{[0, 1]^d}$  have been studied for instance by Li et al. [30], Trutschnig [50], Fernández Sánchez and Trutschnig [19] and Kasper et al. [24]. On the other hand, the (pseudo-) metric  $d_{\mu_1, \dots, \mu_d}$  defined by (2.2) generates a topology that is at most as fine as  $\mathcal{O}_{[0, 1]^d}$ . It is finally worth mentioning that the metric on the set of bivariate subcopulas that was recently introduced by Rachasingho and Tasena [40] basically differs from the metric  $d_{\mu_1, \mu_2}$  and from its variant  $d_{\mu_1, \mu_2}^{\sim}$  introduced in the following Remark 2.1; for details, see Appendix B.

**Remark 2.1** For any fixed  $\mu_1, \dots, \mu_d \in \mathcal{M}_d$ , one can regard the pseudo-metric  $d_{\mu_1, \dots, \mu_d}$  as a metric when changing from  $\mathbf{C}_d$  to the quotient set  $\mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d}$  with respect to the equivalence relation

$$\sim_{\mu_1, \dots, \mu_d} := \{(C, C') \in \mathbf{C}_d \times \mathbf{C}_d : C = C' \text{ on } \overline{\text{ran} F_{\mu_1}} \times \dots \times \overline{\text{ran} F_{\mu_d}}\}.$$

On the resulting quotient set  $\mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d}$ , one may then define a metric through  $d_{\mu_1, \dots, \mu_d}^{\sim}(\mathbb{C}, \mathbb{C}') := d_{\mu_1, \dots, \mu_d}(C, C')$ , where  $C, C'$  are (arbitrary) representatives of the equivalence classes  $\mathbb{C}, \mathbb{C}' \in \mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d}$ . The topology on  $\mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d}$  generated

by  $d_{\mu_1, \dots, \mu_d}^\sim$ , henceforth denoted by  $\mathcal{O}_{\mu_1, \dots, \mu_d}^\sim$ , preserves the topological structure in the sense that a set  $G \subseteq \mathbf{C}_d$  lies in  $\mathcal{O}_{\mu_1, \dots, \mu_d}^\sim$  if and only if the set

$$\{\mathbb{C} \in \mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d} : \text{there exists a } C \in \mathbb{C} \text{ with } C \in G\}$$

lies in  $\mathcal{O}_{\mu_1, \dots, \mu_d}^\sim$ .

## 2.2 The set $\mathcal{M}_d^p$ and the $p$ -weak topology

Fix  $p \in \mathbb{R}_+$  and let  $\mathcal{M}_d^p$  be the set of all  $\mu \in \mathcal{M}_d$  for which  $\int_{\mathbb{R}^d} |x|^p \mu(dx) < \infty$ . Note that  $\mathcal{M}_d = \mathcal{M}_d^0 \supseteq \mathcal{M}_d^{p_1} \supseteq \mathcal{M}_d^{p_2}$  for any  $p_1, p_2 \in \mathbb{R}_+$  with  $p_1 \leq p_2$ . The  $p$ -weak topology on  $\mathcal{M}_d^p$ , henceforth denoted by  $\mathcal{O}_d^p$ , is defined as the coarsest topology for which all mappings  $\mu \mapsto \int f d\mu$ ,  $f \in \mathcal{C}_d^p$ , are continuous, where  $\mathcal{C}_d^p$  is the space of all continuous functions  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  with  $\sup_{x \in \mathbb{R}^d} |f(x)| / (1 + |x|^p) < \infty$ . The 0-weak topology on  $\mathcal{M}_d^0 (= \mathcal{M}_d)$  is just the classical *weak topology*, and the  $p$ -weak topology  $\mathcal{O}_d^p$  is finer than the relative weak topology  $\mathcal{O}_d^0 \cap \mathcal{M}_d^p$  when  $p > 0$ .

It is known from Krättschmer et al. [28, Lemma 2.1] that  $(\mathcal{M}_d^p, \mathcal{O}_d^p)$  is a Polish space and that  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^p$  if and only if both  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^0 \cap \mathcal{M}_d^p$  and

$$\int_{\mathbb{R}^d} |x|^p \mu_n(dx) \longrightarrow \int_{\mathbb{R}^d} |x|^p \mu(dx).$$

In particular,  $\mathcal{O}_d^p$  is metrised by

$$d(\mu, \nu) := d_{\text{weak}}(\mu, \nu) + \left| \int_{\mathbb{R}^d} |x|^p \mu(dx) - \int_{\mathbb{R}^d} |x|^p \nu(dx) \right|$$

for any metric  $d_{\text{weak}}$  which metrises  $\mathcal{O}_d^0$ . Already in the 1980s, Bickel and Freedman [5, Lemma 8.3] proved for  $p \in [1, \infty)$  that  $\mathcal{O}_d^p$  is also metrisable by the  $L^p$ -Wasserstein metric. The following proposition is a sort of continuous mapping theorem.

**Proposition 2.2** *Let  $d, d' \in \mathbb{N}$  and  $p, p' \in \mathbb{R}_+$ . Let  $h: \mathbb{R}^d \rightarrow \mathbb{R}^{d'}$  be a continuous function such that  $\sup_{x \in \mathbb{R}^d} |h(x)|^{p'} / (1 + |x|^p) < \infty$ . Then  $\mathfrak{h}(\mu) := \mu \circ h^{-1}$  lies in  $\mathcal{M}_{d'}^{p'}$  for any  $\mu \in \mathcal{M}_d^p$ , and the map  $\mathfrak{h}: \mathcal{M}_d^p \rightarrow \mathcal{M}_{d'}^{p'}$  is  $(\mathcal{O}_d^p, \mathcal{O}_{d'}^{p'})$ -continuous.*

## 2.3 A generalisation of Deheuvels' copula convergence theorem

Deheuvels' convergence theorem [10, Théorème 2.3, Lemma 4.1] says that given a  $d$ -variate distribution  $\mu \in \mathcal{M}_d$  whose marginal distributions  $\mu_1, \dots, \mu_d \in \mathcal{M}_1$  all possess continuous distribution functions  $F_{\mu_1}, \dots, F_{\mu_d}$ , a sequence  $(\mu_n)_{n \in \mathbb{N}} \in \mathcal{M}_d^{\mathbb{N}}$  converges to  $\mu$  in  $\mathcal{O}_d^0$  if and only if  $(\mu_{n,i})$  converges to  $\mu_i$  in  $\mathcal{O}_1^0$ ,  $i = 1, \dots, d$ , and  $d_{[0,1]^d}(C_n, C) \rightarrow 0$ . Here  $C$  is the unique copula of  $\mu$ ,  $C_n$  is any copula of  $\mu_n$ , and  $\mu_{n,i}$  is the  $i$ th marginal distribution of  $\mu_n$ . Sempi [47] and Lindner and Szimayer [31] extended Deheuvels' result to the general case where the marginal distribution functions  $F_{\mu_1}, \dots, F_{\mu_d}$  might be discontinuous. Theorem 2.1 in [31] shows

that  $(\mu_n)_{n \in \mathbb{N}} \in \mathcal{M}_d^{\mathbb{N}}$  converges to  $\mu$  in  $\mathcal{O}_d^0$  if and only if  $(\mu_{n,i})$  converges to  $\mu_i$  in  $\mathcal{O}_1^0$ ,  $i = 1, \dots, d$ , and  $d_{\mu_1, \dots, \mu_d}(C_n, C) \rightarrow 0$  (a similar result was proved earlier for  $d = 2$  in [47, Theorems 2 and 3]). Recall that  $C$  is uniquely determined only on  $\text{ran} F_{\mu_1} \times \dots \times \text{ran} F_{\mu_d}$ . The results of [47, Example 2] and [31, Example 2.2] show that convergence of the copula on the whole unit cube  $[0, 1]^d$  can indeed fail. Theorem 2.3 below is a version of the Sempí–Lindner–Szimayer result where the weak topologies are replaced by  $p$ -weak topologies.

Consider the map  $\mathfrak{P}_d : \mathbf{C}_d \times \mathcal{M}_1 \times \dots \times \mathcal{M}_1 \rightarrow \mathcal{M}_d$  defined by

$$\mathfrak{P}_d(C, \mu_1, \dots, \mu_d) := \mathfrak{p}_d(C(F_{\mu_1}, \dots, F_{\mu_d})), \quad (2.3)$$

where  $\mathfrak{p}_d$  assigns to a  $d$ -variate distribution function its corresponding Borel probability measure on  $\mathbb{R}^d$ . Note that  $\mathfrak{P}_d(C, \mu_1, \dots, \mu_d)$  remains unchanged when  $C$  is modified outside  $\text{ran} F_{\mu_1} \times \dots \times \text{ran} F_{\mu_d}$ . It is easily seen (see Appendix A.2) that for any  $p \in \mathbb{R}_+$ , the univariate distributions  $\mu_1, \dots, \mu_d$  lie in  $\mathcal{M}_1^p$  if and only if the  $d$ -variate distribution  $\mathfrak{P}_d(C, \mu_1, \dots, \mu_d)$  lies in  $\mathcal{M}_d^p$ , regardless of the copula  $C$ . In particular, the restriction of  $\mathfrak{P}_d$  to  $\mathbf{C}_d \times \mathcal{M}_1^p \times \dots \times \mathcal{M}_1^p$  can be regarded as an  $\mathcal{M}_d^p$ -valued map.

**Theorem 2.3** Fix  $p \in \mathbb{R}_+$  and let  $(C, \mu_1, \dots, \mu_d)$  and  $(C_n, \mu_{n,1}, \dots, \mu_{n,d})$ ,  $n \in \mathbb{N}$ , be elements of  $\mathbf{C}_d \times \mathcal{M}_1^p \times \dots \times \mathcal{M}_1^p$ . Then

$$\mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d}) \longrightarrow \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)$$

in  $\mathcal{O}_d^p$  if and only if  $\mu_{n,i} \rightarrow \mu_i$  in  $\mathcal{O}_1^p$ ,  $i = 1, \dots, d$ , and  $d_{\mu_1, \dots, \mu_d}(C_n, C) \rightarrow 0$ .

Recall that a sequence in a product space converges in the product topology if and only if for each projection, the corresponding marginal sequence converges. As a direct consequence, we can obtain from Theorem 2.3 the following corollary, taking into account that each of the involved topologies is metrisable, or at least pseudo-metrisable, and that  $d_{\mu_1, \dots, \mu_d}(C, C) = 0$  for any  $C \in \mathbf{C}_d$  and  $\mu_1, \dots, \mu_d \in \mathcal{M}_1$ .

**Corollary 2.4** For any  $p \in \mathbb{R}_+$ ,  $C \in \mathbf{C}_d$  and  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ , the following two assertions hold:

- (i) The map  $\mathfrak{P}_d(\cdot, \mu_1, \dots, \mu_d) : \mathbf{C}_d \rightarrow \mathcal{M}_d^p$  is  $(\mathcal{O}_{\mu_1, \dots, \mu_d}, \mathcal{O}_d^p)$ -continuous.
- (ii) The map  $\mathfrak{P}_d(C, \cdot, \dots, \cdot) : \mathcal{M}_1^p \times \dots \times \mathcal{M}_1^p \rightarrow \mathcal{M}_d^p$  is continuous for the pair  $(\mathcal{O}_1^p \times \dots \times \mathcal{O}_1^p, \mathcal{O}_d^p)$ .

## 2.4 Characterisation of $(p)$ -weak convergence in Fréchet classes

For any  $p \in \mathbb{R}_+$  and  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ , the image of the map

$$\mathfrak{P}_d(\cdot, \mu_1, \dots, \mu_d) : \mathbf{C}_d \rightarrow \mathcal{M}_d^p$$

is the Fréchet class  $\mathcal{M}_d(\mu_1, \dots, \mu_d)$ , and we have  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^{p'}$  as well as  $\mathcal{M}_d(\mu_1, \dots, \mu_d) \subseteq \mathcal{M}_d^{p'}$  for any  $p' \in [0, p]$ . Therefore, Corollary 2.4 (i) immediately yields the following result.

**Corollary 2.5** For any  $p \in \mathbb{R}_+$  and  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ , the map

$$\mathfrak{P}_d(\cdot, \mu_1, \dots, \mu_d) : \mathbf{C}_d \rightarrow \mathcal{M}_d(\mu_1, \dots, \mu_d)$$

is  $(\mathcal{O}_{\mu_1, \dots, \mu_d}, \mathcal{O}_d^{p'} \cap \mathcal{M}_d(\mu_1, \dots, \mu_d))$ -continuous for any  $p' \in [0, p]$ .

As a simple consequence of Theorem 2.3, we obtain the following corollary (see Appendix A.3). The result is already known from Krätschmer et al. [28, Proposition 3.9] (with  $A_d$  chosen to be the identity on  $\mathbb{R}^d$ ), where other arguments have been used for the proof.

**Corollary 2.6** For any  $p \in \mathbb{R}_+$  and  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ , we have that

$$\mathcal{O}_d^p \cap \mathcal{M}_d(\mu_1, \dots, \mu_d) = \mathcal{O}_d^{p'} \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$$

for any  $p' \in [0, p]$ .

For any fixed  $\mu_1, \dots, \mu_d \in \mathcal{M}_d^p$ , we use as before  $\mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d}$  to denote the quotient set of  $\mathbf{C}_d$  with respect to the equivalence relation  $\sim_{\mu_1, \dots, \mu_d}$  of identity on  $\text{ran} F_{\mu_1} \times \dots \times \text{ran} F_{\mu_d}$ . Recall from Remark 2.1 that we denote by  $\mathcal{O}_{\mu_1, \dots, \mu_d}^\sim$  the topology on  $\mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d}$  generated by the metric  $d_{\mu_1, \dots, \mu_d}^\sim$  corresponding to the pseudometric  $d_{\mu_1, \dots, \mu_d}$ , and that  $\mathcal{O}_{\mu_1, \dots, \mu_d}^\sim$  preserves the topological structure of  $\mathcal{O}_{\mu_1, \dots, \mu_d}$ .

Let us denote by  $\mathfrak{P}_{\mu_1, \dots, \mu_d} : \mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d} \rightarrow \mathcal{M}_d(\mu_1, \dots, \mu_d)$  the map that assigns to each equivalence class  $\mathbb{C} \in \mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d}$  the unique probability measure  $\mu_{\mathbb{C}} \in \mathcal{M}_d(\mu_1, \dots, \mu_d)$  that satisfies  $\mu_{\mathbb{C}} = \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)$  for all representatives  $C \in \mathbb{C}$ . Then Corollary 2.5 can be reformulated as follows.

**Corollary 2.7** For any  $p \in \mathbb{R}_+$  and  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ , the map

$$\mathfrak{P}_{\mu_1, \dots, \mu_d} : \mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d} \rightarrow \mathcal{M}_d(\mu_1, \dots, \mu_d)$$

is  $(\mathcal{O}_{\mu_1, \dots, \mu_d}^\sim, \mathcal{O}_d^{p'} \cap \mathcal{M}_d(\mu_1, \dots, \mu_d))$ -continuous for any  $p' \in [0, p]$ .

Let  $\mathfrak{C}_{\mu_1, \dots, \mu_d} : \mathcal{M}_d(\mu_1, \dots, \mu_d) \rightarrow \mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d}$  be the map that assigns to each  $\mu \in \mathcal{M}_d(\mu_1, \dots, \mu_d)$  the unique equivalence class  $\mathbb{C}_\mu \in \mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d}$  whose representatives are copulas of  $\mu$ . Then we have the following converse of Corollary 2.7.

**Corollary 2.8** For any  $p \in \mathbb{R}_+$  and  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ , the map

$$\mathfrak{C}_{\mu_1, \dots, \mu_d} : \mathcal{M}_d(\mu_1, \dots, \mu_d) \rightarrow \mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d}$$

is  $(\mathcal{O}_d^{p'} \cap \mathcal{M}_d(\mu_1, \dots, \mu_d), \mathcal{O}_{\mu_1, \dots, \mu_d}^\sim)$ -continuous for any  $p' \in [0, p]$ .

As an immediate consequence of Corollaries 2.7 and 2.8, we obtain the following result. Note that the equivalence (a)  $\Leftrightarrow$  (b) also follows from Corollary 2.6, and that condition (c) is equivalent with  $d_{\mu_1, \dots, \mu_d}(C_n, C) \rightarrow 0$  for any copulas  $C$  and  $C_n$ ,  $n \in \mathbb{N}$ , of  $\mu$  and  $\mu_n$ ,  $n \in \mathbb{N}$ , respectively.



**Corollary 2.9** Fix  $p \in \mathbb{R}_+$  and  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ . Then the following assertions are equivalent for any  $(\mu_n)_{n \in \mathbb{N}} \in \mathcal{M}_d(\mu_1, \dots, \mu_d)^{\mathbb{N}}$  and  $\mu \in \mathcal{M}_d(\mu_1, \dots, \mu_d)$ :

- (a)  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^0 \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$ .
- (b)  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^p \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$ .
- (c)  $\mathfrak{C}_{\mu_1, \dots, \mu_d}(\mu_n) \rightarrow \mathfrak{C}_{\mu_1, \dots, \mu_d}(\mu)$  in  $\mathcal{O}_{\mu_1, \dots, \mu_d}^{\sim}$ .

## 3 Copula robustness

### 3.1 Definition of copula robustness

Let  $\mathcal{M}'_d \subseteq \mathcal{M}_d$  and  $\mathcal{T}_d: \mathcal{M}'_d \rightarrow \mathbf{E}$  be any map taking values in some topological space  $(\mathbf{E}, \mathcal{O}_{\mathbf{E}})$ . As before, let the map  $\mathfrak{P}_d: \mathbf{C}_d \times \mathcal{M}_1 \times \dots \times \mathcal{M}_1 \rightarrow \mathcal{M}_d$  be defined by (2.3). Let  $\mathfrak{D}'_d$  be the set of all  $(C, \mu_1, \dots, \mu_d) \in \mathbf{C}_d \times \mathcal{M}_1 \times \dots \times \mathcal{M}_1$  for which  $\mathfrak{P}_d(C, \mu_1, \dots, \mu_d)$  lies in  $\mathcal{M}'_d$ . Then we can associate with  $\mathcal{T}_d$  a functional  $\mathfrak{T}_d: \mathfrak{D}'_d \rightarrow \mathbf{E}$  through

$$\mathfrak{T}_d(C, \mu_1, \dots, \mu_d) := \mathcal{T}_d(\mathfrak{P}_d(C, \mu_1, \dots, \mu_d)). \quad (3.1)$$

Let  $\mathfrak{M}'_d$  be the set of all  $d$ -tuples  $(\mu_1, \dots, \mu_d) \in \mathcal{M}_1 \times \dots \times \mathcal{M}_1$  for which there exists a copula  $C \in \mathbf{C}_d$  such that  $(C, \mu_1, \dots, \mu_d) \in \mathfrak{D}'_d$ . Moreover, for any fixed  $(\mu_1, \dots, \mu_d) \in \mathfrak{M}'_d$ , let the set  $\mathbf{C}'_d(\mu_1, \dots, \mu_d)$  consist of all those copulas  $C \in \mathbf{C}_d$  for which  $(C, \mu_1, \dots, \mu_d) \in \mathfrak{D}'_d$ .

**Definition 3.1** The map  $\mathcal{T}_d$  is *copula robust* if for any fixed  $(\mu_1, \dots, \mu_d) \in \mathfrak{M}'_d$ , the map  $\mathfrak{T}_d(\cdot, \mu_1, \dots, \mu_d): \mathbf{C}'_d(\mu_1, \dots, \mu_d) \rightarrow \mathbf{E}$  is continuous for the pair

$$(\mathcal{O}_{\mu_1, \dots, \mu_d} \cap \mathbf{C}'_d(\mu_1, \dots, \mu_d), \mathcal{O}_{\mathbf{E}}).$$

The sets  $\mathfrak{M}'_d$  and  $\mathbf{C}'_d(\mu_1, \dots, \mu_d)$  are illustrated by Examples 3.2 and 3.3 below. The examples show in particular that the set  $\mathbf{C}'_d(\mu_1, \dots, \mu_d)$  can be quite different from case to case. In Example 3.2, and in the further course, let  $\mathcal{N}_1$  be the set of all non-degenerate univariate normal distributions and for  $d \geq 2$ , let  $\mathcal{N}_d$  be the set of all (possibly degenerate)  $d$ -variate normal distributions with continuous marginals. In Example 3.2, we also need the notion of a Gaussian copula. Recall that, by definition, a  $d$ -variate *Gaussian copula* is an element  $C \in \mathbf{C}_d$  given through

$$C(u_1, \dots, u_d) := \Phi_{\mathbf{0}, R}(\Phi_{0,1}^{-1}(u_1), \dots, \Phi_{0,1}^{-1}(u_d)) \quad (3.2)$$

for some correlation matrix  $R$ , i.e., for some symmetric and positive semi-definite matrix  $R \in [-1, 1]^{d \times d}$  which has entries 1 on the diagonal. Here  $\Phi_{0,1}$  and  $\Phi_{\mathbf{0}, R}$  are respectively the distribution function of the univariate standard normal distribution and the distribution function of the centered  $d$ -variate normal distribution  $\mathbf{N}_{\mathbf{0}, R}$  with covariance matrix equal to  $R$ , and we set  $\Phi_{0,1}^{-1}(0) := -\infty$  and  $\Phi_{0,1}^{-1}(1) := +\infty$  as well as

$$\Phi_{\mathbf{0}, R}(x_1, \dots, x_d) := \mathbf{N}_{\mathbf{0}, R} \left[ \bigtimes_{i=1}^d (-\infty, x_i] \cap \mathbb{R}^d \right]$$

for any  $x_1, \dots, x_d \in \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ . The set of all Gaussian copulas is denoted by  $\mathbf{C}_d^{\text{Ga}}$ .

**Example 3.2** If  $\mathcal{M}'_d = \mathcal{N}_d$ , then  $\mathfrak{D}'_d = \mathbf{C}_d^{\text{Ga}} \times \mathcal{N}_1 \times \dots \times \mathcal{N}_1$  (see Appendix A.5). In particular,  $\mathfrak{M}'_d = \mathcal{N}_1 \times \dots \times \mathcal{N}_1$  and  $\mathbf{C}'_d(\mu_1, \dots, \mu_d) = \mathbf{C}_d^{\text{Ga}}$  for any  $(\mu_1, \dots, \mu_d) \in \mathfrak{M}'_d$ .

**Example 3.3** If  $\mathcal{M}'_d = \mathcal{M}_d^p$  for some  $p \in \mathbb{R}_+$ , then  $\mathfrak{D}'_d = \mathbf{C}_d \times \mathcal{M}_1^p \times \dots \times \mathcal{M}_1^p$  (see Appendix A.6). In particular,  $\mathfrak{M}'_d = \mathcal{M}_1^p \times \dots \times \mathcal{M}_1^p$  and  $\mathbf{C}'_d(\mu_1, \dots, \mu_d) = \mathbf{C}_d$  for any  $(\mu_1, \dots, \mu_d) \in \mathfrak{M}'_d$ .

The following lemma is trivial, but, nevertheless, worth to be written down. In the lemma,  $(\mathbf{E}', \mathcal{O}'_{\mathbf{E}})$  is another topological space.

**Lemma 3.4** If  $\mathcal{T}_d$  is copula robust and  $\mathcal{U} : \mathbf{E} \rightarrow \mathbf{E}'$  is any  $(\mathcal{O}_{\mathbf{E}}, \mathcal{O}'_{\mathbf{E}})$ -continuous map, then the composition  $\mathcal{T}'_d := \mathcal{U} \circ \mathcal{T}_d$  is copula robust.

### 3.2 Copula robustness of functionals on $\mathcal{N}_d$

In this section, let specifically  $\mathcal{M}'_d = \mathcal{N}_d$ . That is, let  $\mathcal{T}_d : \mathcal{N}_d \rightarrow \mathbf{E}$  be any map taking values in some topological space  $(\mathbf{E}, \mathcal{O}_{\mathbf{E}})$ . In view of Example 3.2, the definition of copula robustness of  $\mathcal{T}_d$  (Definition 3.1) can then be reformulated as follows.

**Definition 3.5** The map  $\mathcal{T}_d$  on  $\mathcal{N}_d$  is *copula robust* if for any fixed  $\mu_1, \dots, \mu_d \in \mathcal{N}_1$ , the map  $\mathfrak{T}_d(\cdot, \mu_1, \dots, \mu_d) : \mathbf{C}_d^{\text{Ga}} \rightarrow \mathbf{E}$  is  $(\mathcal{O}_{[0,1]^d} \cap \mathbf{C}_d^{\text{Ga}}, \mathcal{O}_{\mathbf{E}})$ -continuous.

**Remark 3.6** Convergence in  $(\mathbf{C}_d^{\text{Ga}}, \mathcal{O}_{[0,1]^d} \cap \mathbf{C}_d^{\text{Ga}})$  is nothing but pointwise (or uniform) convergence in  $\mathbf{C}_d^{\text{Ga}}$ . This sort of convergence is therefore equivalent to convergence of the respective correlation matrices in any matrix norm; for details, see Appendix A.7.

**Example 3.7** The identity map  $\mathcal{P}_d : \mathcal{N}_d \rightarrow \mathcal{N}_d$  is copula robust in the sense of Definition 3.5 when the role of  $(\mathbf{E}, \mathcal{O}_{\mathbf{E}})$  is played by  $(\mathcal{N}_d, \mathcal{O}_d^p \cap \mathcal{N}_d)$  for arbitrary (but fixed)  $p \in \mathbb{R}_+$ . For details, see Appendix A.8.

Example 3.7 and Lemma 3.4 (applied to  $\mathcal{T}'_d := \mathcal{T}_d$ ,  $\mathcal{T}_d := \mathcal{P}_d$ ,  $\mathcal{U} := \mathcal{T}_d$ ) immediately yield the following result.

**Theorem 3.8** If  $\mathcal{T}_d$  is  $(\mathcal{O}_d^p \cap \mathcal{N}_d, \mathcal{O}_{\mathbf{E}})$ -continuous for some  $p \in \mathbb{R}_+$ , then it is copula robust.

Of course, Theorem 3.8 can be generalised to larger sets of parametric distributions as for instance to the set  $\mathcal{S}_d$  of all  $d$ -variate (Student)  $t$ -distributions with continuous marginals; see for instance Demarta and McNeil [11] for the definitions of  $d$ -variate  $t$ -distributions and  $t$ -copulas. However, for the sake of clarity and ease, the exposition here is restricted to the Gaussian setting. A perhaps more interesting setting is addressed in the next section.

### 3.3 Copula robustness of functionals on $\mathcal{M}_d^p$

In this section, let specifically  $\mathcal{M}'_d = \mathcal{M}_d^p$  for some  $p \in \mathbb{R}_+$ . That is, let  $\mathcal{T}_d : \mathcal{M}_d^p \rightarrow \mathbf{E}$  be any map taking values in some topological space  $(\mathbf{E}, \mathcal{O}_{\mathbf{E}})$ . In view of Example 3.3, the definition of copula robustness of  $\mathcal{T}_d$  (Definition 3.1) can then be reformulated as follows.

**Definition 3.9** The map  $\mathcal{T}_d$  on  $\mathcal{M}_d^p$  is *copula robust* if for any fixed  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ , the map  $\mathfrak{T}_d(\cdot, \mu_1, \dots, \mu_d) : \mathbf{C}_d \rightarrow \mathbf{E}$  is  $(\mathcal{O}_{\mu_1, \dots, \mu_d}, \mathcal{O}_{\mathbf{E}})$ -continuous.

With the help of Corollaries 2.5 and 2.8, we can derive the following characterisation of copula robustness of  $\mathcal{T}_d$ . For details, see Appendix A.9.

**Theorem 3.10** Let  $\mathcal{T}_d : \mathcal{M}_d^p \rightarrow \mathbf{E}$  be any map. Then  $\mathcal{T}_d$  is copula robust if and only if for any fixed  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ , its restriction  $\mathcal{T}_d|_{\mathcal{M}_d(\mu_1, \dots, \mu_d)}$  to the Fréchet class  $\mathcal{M}_d(\mu_1, \dots, \mu_d)$  is continuous for the pair  $(\mathcal{O}_d^0 \cap \mathcal{M}_d(\mu_1, \dots, \mu_d), \mathcal{O}_{\mathbf{E}})$ .

**Example 3.11** Corollary 2.4 (i) shows that the identity map  $\mathcal{P}_d : \mathcal{M}_d^p \rightarrow \mathcal{M}_d^p$  is copula robust in the sense of Definition 3.9 when the role of  $(\mathbf{E}, \mathcal{O}_{\mathbf{E}})$  is played by  $(\mathcal{M}_d^p, \mathcal{O}_{\mathcal{M}_d^p})$ .

Example 3.11 and Lemma 3.4 (applied to  $\mathcal{T}'_d := \mathcal{T}_d$ ,  $\mathcal{T}_d := \mathcal{P}_d$ ,  $\mathcal{U} := \mathcal{T}_d$ ) immediately yield the following result.

**Theorem 3.12** If  $\mathcal{T}_d$  is  $(\mathcal{O}_d^p, \mathcal{O}_{\mathbf{E}})$ -continuous, then it is copula robust.

Now fix  $d' \in \mathbb{N}$  and  $p' \in \mathbb{R}_+$ . Proposition 2.2 ensures that in the scope of the following corollary, we have  $\mu \circ h^{-1} \in \mathcal{M}_{d'}^{p'}$  for any  $\mu \in \mathcal{M}_d^p$ .

**Corollary 3.13** Let  $\mathcal{T}_{d'} : \mathcal{M}_{d'}^{p'} \rightarrow \mathbf{E}$  be an  $(\mathcal{O}_{p'}^{d'}, \mathcal{O}_{\mathbf{E}})$ -continuous map and suppose  $h : \mathbb{R}^d \rightarrow \mathbb{R}^{d'}$  is a continuous map with  $\sup_{x \in \mathbb{R}^d} |h(x)|^{p'} / (1 + |x|^p) < \infty$ . Then the map  $\mathcal{T}'_d : \mathcal{M}_d^p \rightarrow \mathbf{E}$  defined by  $\mathcal{T}'_d(\mu) := \mathcal{T}_{d'}(\mu \circ h^{-1})$  is copula robust.

## 4 Example 1: risk measures of aggregate risks

### 4.1 Foundations of risk measures

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be an atomless probability space and denote by  $L^0 := L^0(\Omega, \mathcal{F}, \mathbb{P})$  the usual class of all finite-valued random variables modulo the equivalence relation of  $\mathbb{P}$ -a.s. identity. Moreover, let  $L^p = L^p(\Omega, \mathcal{F}, \mathbb{P})$  be the usual  $L^p$ -space,  $p > 0$ . For any  $p \in \mathbb{R}_+$ , we say that a map  $\rho : L^p \rightarrow \mathbb{R}$  is a *risk measure* when the following three conditions are satisfied:

- (i) (monotonicity)  $\rho(X) \leq \rho(Y)$  for  $X, Y \in L^p$  with  $X \leq Y$ ;
- (ii) (cash-additivity)  $\rho(X + m) = \rho(X) + m$  for  $X \in L^p$  and  $m \in \mathbb{R}$ ;
- (iii) (distribution-invariance)  $\rho(X) = \rho(Y)$  for  $X, Y \in L^p$  with  $\mathbb{P}_X = \mathbb{P}_Y$ .

In this context, the elements of  $L^p$  should be seen as payoff profiles where positive realisations correspond to losses. Following Föllmer and Schied [22], [23, Chap. 4], a risk measure  $\rho : L^p \rightarrow \mathbb{R}$  is said to be *convex* if it satisfies the following condition:

(iv) (convexity)  $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$  for all  $X, Y \in L^p$  and  $\lambda \in [0, 1]$ .

The following example recalls three risk measures which are popular in practice and/or among academics. For background, see Emmer et al. [18] and references cited therein.

**Example 4.1** Fix  $\alpha \in (0, 1)$ .

(i) The *value at risk* at level  $\alpha$  is the risk measure  $\text{VaR}_\alpha : L^0 \rightarrow \mathbb{R}$  defined by  $\text{VaR}_\alpha(X) := F_X^\leftarrow(\alpha)$ , where  $F_X^\leftarrow(\alpha) := \inf\{x \in \mathbb{R} : F_X(x) \geq \alpha\}$  is the lower  $\alpha$ -quantile of  $\mathbb{P}_X$ . It is not convex.

(ii) The *average value at risk* at level  $\alpha$  is the risk measure  $\text{AVaR}_\alpha : L^1 \rightarrow \mathbb{R}$  defined by  $\text{AVaR}_\alpha(X) := \frac{1}{1-\alpha} \int_\alpha^1 F_X^\leftarrow(s) ds$  and known to be convex; see for instance the work of Wang and Dhaene [53].

(iii) The  $\alpha$ -*expectile* at level  $\alpha$  is the risk measure  $\text{Ept}_\alpha : L^1 \rightarrow \mathbb{R}$  defined by  $\text{Ept}_\alpha(X) := \mathbb{U}_\alpha(X)^{-1}(0)$ , where  $\mathbb{U}_\alpha(X)^{-1}$  denotes the inverse of the function  $\mathbb{U}_\alpha(X)(m) := \mathbb{E}[U_\alpha(X - m)]$  with  $U_\alpha(x) := \alpha x$  or  $(1 - \alpha)x$  depending on whether  $x \geq 0$  or  $x < 0$ . It is well defined, and known to be convex if and only if  $\alpha \geq 1/2$ ; see the work of Bellini et al. [3].

For any risk measure  $\rho : L^p \rightarrow \mathbb{R}$ , we may define a functional  $\mathcal{R}_\rho : \mathcal{M}_1^p \rightarrow \mathbb{R}$  through

$$\mathcal{R}_\rho(\mu) := \rho(X_\mu), \quad (4.1)$$

where  $X_\mu$  is any random variable on  $(\Omega, \mathcal{F}, \mathbb{P})$  with distribution  $\mu$ . We refer to  $\mathcal{R}_\rho$  as the *risk functional* associated with  $\rho$ . The assertion of the following result is a direct consequence of Cheridito and Li [8, Theorem 4.1] combined with the representation theorem of Krättschmer et al. [27, Theorem 3.5]. Here  $\mathcal{O}_{\mathbb{R}}$  refers to the natural topology on  $\mathbb{R}$ .

**Theorem 4.2** Let  $p \in \mathbb{R}_+$ . For any convex risk measure  $\rho : L^p \rightarrow \mathbb{R}$ , the corresponding risk functional  $\mathcal{R}_\rho : \mathcal{M}_1^p \rightarrow \mathbb{R}$  is  $(\mathcal{O}_1^p, \mathcal{O}_{\mathbb{R}})$ -continuous.

## 4.2 Copula robustness of risk measures of aggregate risks

Let  $\rho : L^{p'} \rightarrow \mathbb{R}$  be a risk measure for some  $p' \in \mathbb{R}_+$ . Let  $A_d : \mathbb{R}^d \rightarrow \mathbb{R}$  be any continuous map, regarded as an aggregation map in the spirit of McNeil et al. [33, Sect. 6.2.1]. Assume that for some  $p \in \mathbb{R}_+$  and any  $X_1, \dots, X_d \in L^p$ , the random variable  $A_d(X_1, \dots, X_d)$  lies in  $L^{p'}$ . Then we can define a map  $\mathcal{R}_{\rho, A_d} : \mathcal{M}_d^p \rightarrow \mathbb{R}$  through

$$\mathcal{R}_{\rho, A_d}(\mu) := \mathcal{R}_\rho(\mu \circ A_d^{-1}). \quad (4.2)$$

We refer to  $\mathcal{R}_{\rho, A_d}$  as *aggregation risk functional* associated with  $\rho$  and  $A_d$ . Note that the right-hand side in (4.2) equals  $\rho(A_d(X_1, \dots, X_d))$  when  $(X_1, \dots, X_d)$  is an  $\mathbb{R}^d$ -valued random variable with distribution  $\mu$ . As a direct consequence of Corollary 3.13 (applied to  $\mathcal{T}_1 := \mathcal{R}_\rho$ ,  $h := A_d$ ,  $\mathcal{T}'_d := \mathcal{R}_{\rho, A_d}$ ) and Theorem 4.2, we obtain the following result.

**Corollary 4.3** *Take  $p, p' \in \mathbb{R}_+$ , a convex risk measure  $\rho : L^{p'} \rightarrow \mathbb{R}$  and a continuous map  $A_d : \mathbb{R}^d \rightarrow \mathbb{R}$  satisfying  $\sup_{x \in \mathbb{R}^d} |A_d(x)|^{p'}/(1 + |x|)^p < \infty$ . Then the aggregation risk functional  $\mathcal{R}_{\rho, A_d} : \mathcal{M}_d^p \rightarrow \mathbb{R}$  defined by (4.2) is copula robust.*

**Example 4.4** In risk management,  $A_d$  is frequently chosen as one of the following maps; see for instance the textbook by McNeil et al. [33, Sect. 6.2]:

- (i)  $A_d(x_1, \dots, x_d) := \sum_{i=1}^d x_i$ ;
- (ii)  $A_d(x_1, \dots, x_d) := \max\{x_1, \dots, x_d\}$ ;
- (iii)  $A_d(x_1, \dots, x_d) := \sum_{i=1}^d (x_i - t_i)^+$  for thresholds  $t_1, \dots, t_d > 0$ ;
- (iv)  $A_d(x_1, \dots, x_d) := (\sum_{i=1}^d x_i - t)^+$  for a threshold  $t > 0$ .

It is easily seen that for each of these four maps,  $\sup_{x \in \mathbb{R}^d} |A_d(x)|^p/(1 + |x|)^p < \infty$  holds for any  $p \in \mathbb{R}_+$ . That is, all these maps satisfy the assumptions of Corollary 4.3 for  $p' = p$  (and thus for any  $p \in \mathbb{R}_+$  and  $p' \in [0, p]$ ). In particular, for each of these four maps  $A_d$  and for any convex risk measure  $\rho : L^p \rightarrow \mathbb{R}$ , the corresponding aggregation risk functional  $\mathcal{R}_{\rho, A_d} : \mathcal{M}_d^p \rightarrow \mathbb{R}$  defined by (4.2) is copula robust for any  $d \in \mathbb{N}$ .

**Remark 4.5** Of course, the assertion of Corollary 4.3 and the last assertion in Example 4.4 also hold true for any other risk measure  $\rho$  for which the corresponding risk functional  $\mathcal{R}_\rho : \mathcal{M}_1^{p'} \rightarrow \mathbb{R}$  is  $(\mathcal{O}_1^{p'}, \mathcal{O}_{\mathbb{R}})$ -continuous.

Remark 4.5 indicates that in the setting of Corollary 4.3, the assumed convexity of  $\rho$  is not necessary. To give an example that shows that this is indeed true, let  $\rho$  be the  $\alpha$ -expectile  $\text{Ept}_\alpha$  with  $\alpha < 1/2$  (see Example 4.1 (iii)). Then  $\rho$  is not convex (see Bellini et al. [3, Proposition 7(b–c)]), but the corresponding risk functional  $\mathcal{R}_\rho : \mathcal{M}_1^1 \rightarrow \mathbb{R}$  is  $(\mathcal{O}_1^1, \mathcal{O}_{\mathbb{R}})$ -continuous (see Krätschmer and Zähle [29, Theorem 2.1]), and the latter implies that the aggregation risk functional  $\mathcal{R}_{\rho, A_d} : \mathcal{M}_d^1 \rightarrow \mathbb{R}$  is copula robust.

On the other hand, if the risk functional  $\mathcal{R}_\rho : \mathcal{M}_1^{p'} \rightarrow \mathbb{R}$  corresponding to some  $\rho$  is not  $(\mathcal{O}_1^{p'}, \mathcal{O}_{\mathbb{R}})$ -continuous, then copula robustness of  $\mathcal{R}_{\rho, A_d} : \mathcal{M}_d^{p'} \rightarrow \mathbb{R}$  can indeed fail to hold. For instance, Example 4.7 below shows that  $\mathcal{R}_{\rho, A_2} : \mathcal{M}_2^0 \rightarrow \mathbb{R}$  is not copula robust when  $\rho := \text{VaR}_\alpha$  (see Example 4.1 (i)) and  $A_2(x_1, x_2) := x_1 + x_2$ . Note here that the risk functional  $\mathcal{R}_\rho : \mathcal{M}_1^0 \rightarrow \mathbb{R}$  associated with  $\rho := \text{VaR}_\alpha$  is known not to be weakly continuous, and that weak continuity is just  $(\mathcal{O}_1^0, \mathcal{O}_{\mathbb{R}})$ -continuity.

It is further known that the risk functional  $\mathcal{R}_\rho : \mathcal{M}_1^0 \rightarrow \mathbb{R}$  associated with  $\rho := \text{VaR}_\alpha$  can be made weakly continuous when restricting it to the set  $\mathcal{M}_1^{(\alpha)}$  of all those Borel probability measures on  $\mathbb{R}$  that possess a unique  $\alpha$ -quantile (see e.g. van

der Vaart [52, Lemma 21.2]), or even to the set  $\mathcal{M}_1^{\mathcal{L}}$  of all  $\mu \in \bigcap_{s \in (0,1)} \mathcal{M}_1^{(s)}$  that possess a Lebesgue density. Nonetheless, the corresponding aggregation risk functional  $\mathcal{R}_{\rho, A_2}$ , with  $A_2(x_1, x_2) := x_1 + x_2$ , defined on the set  $\mathcal{M}_2^{\mathcal{L}}$  of all Borel probability measures on  $\mathbb{R}^2$  with marginal distributions in  $\mathcal{M}_1^{\mathcal{L}}$ , is still not copula robust. This is also a consequence of Example 4.7. The lack of copula robustness of  $\mathcal{R}_{\rho, A_2}$  on  $\mathcal{M}_2^{\mathcal{L}}$  is not immediately obvious. Note, however, that for  $\mu \in \mathcal{M}_2^{\mathcal{L}}$ , the image measure  $\mu \circ A_2^{-1}$  can be purely discrete (see Example 4.7), i.e.,  $\mu \circ A_2^{-1}$  can lie outside the set  $\mathcal{M}_1^{(\alpha)}$  on which  $\mathcal{R}_{\rho}$  is weakly continuous.

When restricting  $\mathcal{R}_{\rho, A_2}$ , with  $\rho := \text{VaR}_{\alpha}$  and  $A_2(x_1, x_2) := x_1 + x_2$ , to the much smaller set  $\mathcal{N}_2$  introduced before (3.2), then copula robustness holds true. Note that  $\mu \circ A_2^{-1} \in \mathcal{N}'_1 \subseteq \mathcal{M}_1^{(\alpha)}$  for all  $\mu \in \mathcal{N}_2$ , where  $\mathcal{N}'_1 (\supseteq \mathcal{N}_1)$  is the set of all (possibly degenerate) univariate normal distributions. The copula robustness follows from Theorem 3.8 since the restriction of  $\mathcal{R}_{\rho, A_2}$  to  $\mathcal{N}_2$  is  $(\mathcal{O}_2^0 \cap \mathcal{N}_2, \mathcal{O}_{\mathbb{R}})$ -continuous. The latter follows from the  $(\mathcal{O}_2^0 \cap \mathcal{N}_2, \mathcal{O}_1^0 \cap \mathcal{N}'_1)$ -continuity of the map  $\mathfrak{h} : \mathcal{N}_2 \rightarrow \mathcal{N}'_1$  defined by  $\mathfrak{h}(\mu) := \mu \circ A_2^{-1}$  and the  $(\mathcal{O}_1^0 \cap \mathcal{N}'_1, \mathcal{O}_{\mathbb{R}})$ -continuity of the restriction of  $\mathcal{R}_{\rho}$  to  $\mathcal{N}'_1 (\subseteq \mathcal{M}_1^{(\alpha)})$ .

### 4.3 Relation to aggregation robustness of risk measures

In [17], Embrechts et al. consider the special case where  $A_d$  is defined as in (i) of Example 4.4 and  $\rho$  is a coherent distortion risk measure defined on a subset of  $L^1$ . In this case, they obtain an analogue of Corollary 4.3 and refer to it as aggregation robustness. In fact, they do not explicitly consider continuity in the copula, but rather weak continuity of the analogous functional defined on the corresponding Fréchet class. However, as seen in Theorem 3.10, this is the same. A generalisation to more general risk measures and more general aggregation maps is given in the work of Krätschmer et al. [28, Sect. 4.2.4].

The following definition is a reformulation of the definition of aggregation robustness of a risk measure  $\rho : L^p \rightarrow \mathbb{R}$  (i.e., of [17, Definition 2.1]). As before, the aggregation risk functional  $\mathcal{R}_{\rho, A_d}$  associated with  $\rho$  and  $A_d(x_1, \dots, x_d) := \sum_{i=1}^d x_i$  is defined by (4.2).

**Definition 4.6** Let  $p \in \mathbb{R}_+$ . A risk measure  $\rho : L^p \rightarrow \mathbb{R}$  is said to be *aggregation robust* if the corresponding aggregation risk functionals  $\mathcal{R}_{\rho, A_d} : \mathcal{M}_d^p \rightarrow \mathbb{R}$ ,  $d \geq 2$ , are copula robust.

In view of Corollary 4.3 and Example 4.4, any convex risk measure  $\rho : L^p \rightarrow \mathbb{R}$  is aggregation robust. This assertion remains true when replacing in Definition 4.6 the map  $A_d(x_1, \dots, x_d) := \sum_{i=1}^d x_i$  by any other of the maps introduced in Example 4.4.

In their Example 2.2, Embrechts et al. [17] demonstrated that for any  $\alpha \in (0, 1)$ , the value at risk  $\text{VaR}_{\alpha} : L^0 \rightarrow \mathbb{R}$  is not aggregation robust. The following example extends the first part of that example (from  $\alpha = 1/2$  to general  $\alpha \in (0, 1)$ ) and shows that for any  $\alpha \in (0, 1)$  and  $p \in \mathbb{R}_+$ , the aggregation risk functional  $\mathcal{R}_{\text{VaR}_{\alpha}, A_2} : \mathcal{M}_2^p \rightarrow \mathbb{R}$  is not copula robust. The example is in particular interesting in that it shows that even if the marginal distributions  $\mu_1, \dots, \mu_d$  possess Lebesgue densities and unique quantiles, the map  $C \mapsto \mathfrak{R}_{\text{VaR}_{\alpha}, A_d}(C, \mu_1, \dots, \mu_d)$  need not be continuous when choosing

$A_d(x_1, \dots, x_d) := \sum_{i=1}^d x_i$  (here  $\mathfrak{R}_{\text{VaR}_{\alpha}, A_d}$  is derived from  $\mathcal{R}_{\rho, A_d}$  as  $\mathfrak{T}_d$  is derived from  $\mathcal{T}_d$  in (3.1)). The point here is that for random variables  $X_1, \dots, X_d$  with distributions  $\mu_1, \dots, \mu_d$ , the distribution of  $\sum_{i=1}^d X_i$  can be discrete even if  $\mu_1, \dots, \mu_d$  possess Lebesgue densities. This fact has already been pointed out in [17].

**Example 4.7** Generalising the first part of Embrechts et al. [17, Example 2.2], define for  $\alpha \in (0, 1/2]$  a bivariate copula  $C_0^{(\alpha)}$  through

$$\begin{aligned} C_0^{(\alpha)}(u_1, u_2) \\ := \max \{ \min\{u_1, \alpha\} + \min\{u_2, \alpha\} - \alpha, 0 \} + \max\{u_1 + u_2 - (1 + \alpha), 0\}, \end{aligned}$$

let  $C_1$  be the bivariate independence copula, i.e.,  $C_1(u_1, u_2) := u_1 u_2$ , and define for any  $t \in [0, 1]$  the copula  $C_t^{(\alpha)}$  as a mixture of  $C_0^{(\alpha)}$  and  $C_1$  via

$$C_t^{(\alpha)}(u_1, u_2) := (1 - t) C_0^{(\alpha)}(u_1, u_2) + t C_1(u_1, u_2).$$

Moreover, for any  $t \in [0, 1]$ , let  $\hat{C}_t^{(\alpha)}$  be the survival copula of  $C_t^{(\alpha)}$  which is defined by  $\hat{C}_t^{(\alpha)}(u_1, u_2) := u_1 + u_2 - 1 + C_t^{(\alpha)}(1 - u_1, 1 - u_2)$ . Finally, let  $\mu_1 := \mu_2 := U_{[0,1]}$  as well as  $\hat{\mu}_1 := \hat{\mu}_2 := U_{[-1,0]}$ , where  $U_I$  is used to denote the uniform distribution on  $I$ . Then the following two assertions are valid:

(i)  $\mathfrak{R}_{\text{VaR}_{\alpha}, A_2}(C_0^{(\alpha)}, \mu_1, \mu_2) = \alpha$  and  $\mathfrak{R}_{\text{VaR}_{\alpha}, A_2}(C_t^{(\alpha)}, \mu_1, \mu_2) = \sqrt{2\alpha}$  for any  $t \in (0, 1]$ . Therefore we have  $\lim_{t \searrow 0} C_t^{(\alpha)} = C_0^{(\alpha)}$  uniformly, but

$$\lim_{t \searrow 0} \mathfrak{R}_{\text{VaR}_{\alpha}, A_2}(C_t^{(\alpha)}, \mu_1, \mu_2) \neq \mathfrak{R}_{\text{VaR}_{\alpha}, A_2}(C_0^{(\alpha)}, \mu_1, \mu_2).$$

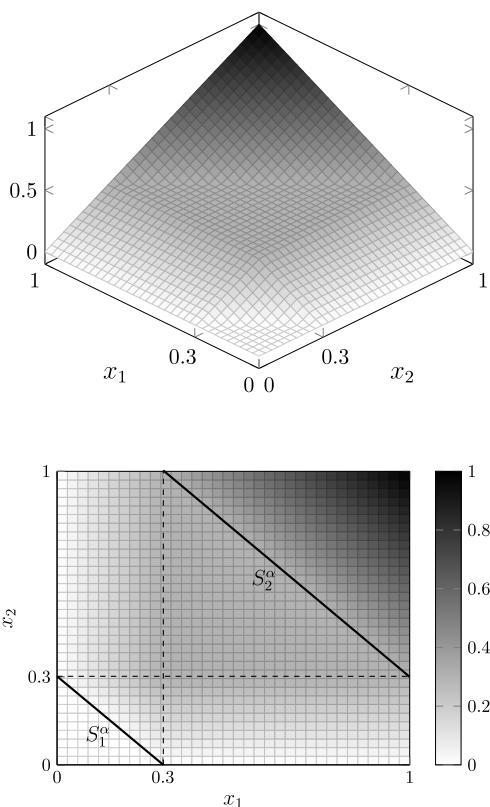
(ii)  $\mathfrak{R}_{\text{VaR}_{1-\alpha}, A_2}(\hat{C}_0^{(\alpha)}, \hat{\mu}_1, \hat{\mu}_2) = -1 - \alpha$  and  $\mathfrak{R}_{\text{VaR}_{1-\alpha}, A_2}(\hat{C}_t^{(\alpha)}, \hat{\mu}_1, \hat{\mu}_2) = -\sqrt{2\alpha}$  for any  $t \in (0, 1]$ . Therefore we have that  $\lim_{t \searrow 0} \hat{C}_t^{(\alpha)} = \hat{C}_0^{(\alpha)}$  uniformly, but

$$\lim_{t \searrow 0} \mathfrak{R}_{\text{VaR}_{1-\alpha}, A_2}(\hat{C}_t^{(\alpha)}, \hat{\mu}_1, \hat{\mu}_2) \neq \mathfrak{R}_{\text{VaR}_{1-\alpha}, A_2}(\hat{C}_0^{(\alpha)}, \hat{\mu}_1, \hat{\mu}_2).$$

For details, see Appendix A.11.

It is worth commenting on the copulas  $C_0^{(\alpha)}$ ,  $C_1$  and  $C_t^{(\alpha)}$  in the preceding example. The copula  $C_1$  is well known; it is simply the distribution function of the uniform distribution on  $[0, 1]^2$ . The copula  $C_0^{(\alpha)}$  is the distribution function of the ‘uniform distribution’ on the union of the two disjoint line segments  $S_1^\alpha$  and  $S_2^\alpha$  with endpoints  $(\alpha, 0)$ ,  $(0, \alpha)$  and  $(1, \alpha)$ ,  $(\alpha, 1)$ , respectively (see Appendix A.11 for the precise definition). Thus  $C_t^{(\alpha)}$  is the distribution function of the Borel probability measure on  $[0, 1]^2$  that is defined as a convex combination (with coefficients  $t$  and  $1 - t$ ) of the uniform distribution on  $[0, 1]^2$  and the ‘uniform distribution’ on  $S_1^\alpha \cup S_2^\alpha$ . For a visualisation of  $C_0^{(\alpha)}$ , see Fig. 1, and note that the distribution of the sum of two  $U_{[0,1]}$ -distributed random variables coupled via  $C_0^{(\alpha)}$  is the two-point distribution  $\alpha\delta_\alpha + (1 - \alpha)\delta_{1+\alpha}$ .

**Fig. 1** Visualisation of the copula  $C_0^{(\alpha)}$  and of the sets  $S_1^\alpha$  and  $S_2^\alpha$  for  $\alpha = 0.3$



#### 4.4 Application to optimal capital and risk allocations

Let  $p \in [1, \infty)$ . As in the work of Filipović and Svindland [21], consider  $d$  agents, or business units, with endowments  $X_1, \dots, X_d \in L^p$ . We then assume that these agents assess the riskiness of their positions by means of some convex risk measures  $\rho_1, \dots, \rho_d : L^p \rightarrow \mathbb{R}$  (in the sense of Sect. 4.1). In order to minimise the total and individual risk, the agents redistribute the aggregate endowment  $X := \sum_{i=1}^d X_i$  among themselves. By a *redistribution* of  $X$ , we mean any  $d$ -tuple  $(Y_1, \dots, Y_d)$  of random variables (payoffs) in  $L^p$  such that  $X = \sum_{i=1}^d Y_i$ . A redistribution  $(X_1^*, \dots, X_d^*)$  is called an *optimal capital and risk allocation* of  $X$  if

$$\sum_{i=1}^n \rho_i(X_i^*) = \inf \left\{ \sum_{i=1}^n \rho_i(Y_i) : Y_1, \dots, Y_d \in L^p \text{ and } \sum_{i=1}^n Y_i = X \right\}. \quad (4.3)$$

Here it is assumed that the redistribution is not subject to frictions, i.e., that every redistribution of  $X$  is admissible, even if this is not always the case (as pointed out by Filipović and Kupper [20]).

Note that an optimal capital and risk allocation  $(X_1^*, \dots, X_d^*)$  as a redistribution must satisfy  $X = \sum_{i=1}^d X_i^*$ . An optimal capital and risk allocation of  $X$  need not



exist. If it exists, it coincides with the *inf-convolution* of  $\rho_1, \dots, \rho_d$  at  $X$ , denoted by  $\square_{i=1}^d \rho_i(X)$ , which is defined by the right-hand side of (4.3). Using the convention  $\inf \emptyset = \infty$ , the inf-convolution can be seen as a map  $\square_{i=1}^d \rho_i : L^p \rightarrow (-\infty, \infty]$ . For background, see [21] and references cited therein.

In the above economic setting, the inf-convolution can also be regarded as a map  $\boxtimes_{i=1}^d \rho_i : L^p \times \dots \times L^p \rightarrow (-\infty, \infty]$  through

$$\boxtimes_{i=1}^d \rho_i(X_1, \dots, X_d) := \square_{i=1}^d \rho_i \left( \sum_{i=1}^d X_i \right).$$

Since we assumed  $\rho_1, \dots, \rho_d$  to be convex risk measures on  $L^p$ , a result of Filipović and Svindland [21, Corollary 2.7] ensures that the inf-convolution  $\square_{i=1}^d \rho_i$  is a convex risk measure on  $L^p$ , too (note that in [21, Corollary 2.7],  $\square_{i=1}^d \rho_i$  is exact, and hence it is  $\mathbb{R}$ -valued if  $\rho_1, \dots, \rho_d$  are  $\mathbb{R}$ -valued). As a convex risk measure,  $\square_{i=1}^d \rho_i$  is distribution-invariant, and so is  $\boxtimes_{i=1}^d \rho_i$ . Thus we may associate with  $\boxtimes_{i=1}^d \rho_i$  a corresponding functional  $\mathcal{R}_{\boxtimes_{i=1}^d \rho_i} : \mathcal{M}_d^p \rightarrow \mathbb{R}$  through

$$\mathcal{R}_{\boxtimes_{i=1}^d \rho_i}(\mu) := \mathcal{R}_{\square_{i=1}^d \rho_i, A_d}(\mu) = \mathcal{R}_{\square_{i=1}^d \rho_i}(\mu \circ A_d^{-1})$$

with  $A_d(x_1, \dots, x_d) = \sum_{i=1}^d x_i$ , where  $\mathcal{R}_{\square_{i=1}^d \rho_i}$  and  $\mathcal{R}_{\square_{i=1}^d \rho_i, A_d}$  are defined as in (4.1) and (4.2), respectively.

It is worth mentioning that [21, Corollary 2.7] even ensures that for any  $X \in L^p$ , there exists a comonotone optimal capital and risk allocation  $(X_1^*, \dots, X_d^*)$ . This implies that whenever  $\rho_1 = \dots = \rho_d$  and  $\rho := \rho_1$  is comonotonic (i.e., finitely additive for all comonotone risks), we have

$$\mathcal{R}_{\boxtimes_{i=1}^d \rho}(\mu) = \mathcal{R}_{\rho}(\mu \circ A_d^{-1}) \quad (4.4)$$

for any  $\mu \in \mathcal{M}_d^p$  (see Appendix A.12). Of course, for convex risk measures  $\rho$  that are not comonotonic, the representation (4.4) need not apply. An example for a comonotonic convex risk measure is the average value at risk at level  $\alpha \in (0, 1)$ . A counterexample is the  $\alpha$ -expectile at level  $\alpha \in [1/2, 1)$ ; see Emmer et al. [18].

The following result is a direct consequence of Corollary 4.3 and Example 4.4 since we have seen above that  $\square_{i=1}^d \rho_i$  is a convex risk measure on  $L^p$  if  $\rho_1, \dots, \rho_d$  are.

**Corollary 4.8** *Let  $p \in [1, \infty)$  and  $\rho_1, \dots, \rho_d : L^p \rightarrow \mathbb{R}$  be convex risk measures. Then  $\mathcal{R}_{\boxtimes_{i=1}^d \rho_i} : \mathcal{M}_d^p \rightarrow \mathbb{R}$  is copula robust.*

The following example shows that if the risk measures  $\rho_1, \dots, \rho_d$  are not assumed to be convex, copula robustness of  $\mathcal{R}_{\boxtimes_{i=1}^d \rho_i}$  may fail; recall that  $\text{VaR}_{\alpha}$  is not convex.

**Example 4.9** It is known from the work of Embrechts et al. [13, Corollary 2] that  $\square_{i=1}^2 \text{VaR}_{\alpha} = \text{VaR}_{2\alpha}$  on  $L^1$  when  $\alpha \in (0, 1/2)$ . Therefore

$$\mathcal{R}_{\boxtimes_{i=1}^2 \text{VaR}_{\alpha}}(\mu) = \mathcal{R}_{\square_{i=1}^2 \text{VaR}_{\alpha}, A_2}(\mu) = \mathcal{R}_{\text{VaR}_{2\alpha}, A_2}(\mu) = \mathcal{R}_{\text{VaR}_{2\alpha}}(\mu \circ A_2^{-1})$$

for any  $\mu \in \mathcal{M}_2^1$  and  $\alpha \in (0, 1/2)$ . Thus it follows from Example 4.7 that

$$\mathcal{R}_{\boxtimes_{i=1}^2 \text{VaR}_\alpha} : \mathcal{M}_2^1 \rightarrow \mathbb{R}$$

is not copula robust for any  $\alpha \in (0, 1/2)$ .

## 5 Example 2: stochastic programming problems

### 5.1 A class of stochastic programming problems

Adopting the framework of Claus et al. [9], let  $\mathcal{E}$  be a nonempty and compact subset of  $\mathbb{R}^k$ ,  $h : \mathcal{E} \times \mathbb{R}^d \rightarrow \mathbb{R}$  a Borel-measurable function and  $Z$  an  $\mathbb{R}^d$ -valued random variable on an atomless probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Let  $p \in [1, \infty)$  and assume that  $h(\xi, Z)$  is contained in  $L^p = L^p(\Omega, \mathcal{F}, \mathbb{P})$  for any  $\xi \in \mathcal{E}$ . Consider the optimisation problem

$$\min \{ \rho(h(\xi, Z)) : \xi \in \mathcal{E} \}, \quad (5.1)$$

where  $\rho : L^p \rightarrow \mathbb{R}$  is any map. A classical example for  $\rho$  is the expectation, i.e.,  $\rho(Y) = \mathbb{E}[Y]$ , where  $p = 1$ . In Sect. 5.2, we consider another example where  $\rho$  is a more general monotone, distribution-invariant and convex function on  $L^p$ . Problem (5.1) can be written as  $\min \{ \mathcal{R}_\rho(\mathbb{P} \circ h(\xi, Z)^{-1}) : \xi \in \mathcal{E} \}$  or, equivalently, as

$$\min \{ \mathcal{R}_\rho((\delta_\xi \otimes \mu) \circ h^{-1}) : \xi \in \mathcal{E} \}, \quad (5.2)$$

where  $\mathcal{R}_\rho$  is derived from  $\rho$  as in (4.1) and  $\mu$  denotes the distribution of  $Z$ .

Lemma 5.1 below assumes the following three conditions, where monotonicity, distribution-invariance and convexity are defined as in (i), (iii) and (iv) in Sect. 4.1. Recall that  $(\Omega, \mathcal{F}, \mathbb{P})$  is assumed to be atomless.

(a)  $\rho : L^p \rightarrow \mathbb{R}$ , for some  $p \in [1, \infty)$ , is monotone, distribution-invariant and convex.

(b)  $h : \mathcal{E} \times \mathbb{R}^d \rightarrow \mathbb{R}$  is Borel-measurable and limited by an exponent  $\gamma \in \mathbb{R}_{++}$ .

(c)  $(\delta_\xi \otimes \mu)[D_h] = 0$  for any  $\xi \in \mathcal{E}$  and  $\mu \in \mathcal{M}_d^{\gamma p}$ .

The second requirement in (b) means that there exists some locally bounded map  $\eta : \mathcal{E} \rightarrow (0, \infty)$  such that  $|h(\xi, z)| \leq \eta(\xi)(1 + |z|)^\gamma$  for all  $(\xi, z) \in \mathcal{E} \times \mathbb{R}^d$ . In (c), the set  $D_h$  is the set of all discontinuity points of  $h$ . Under conditions (a) and (b), the map  $\mathcal{Q}_{\rho, h} : \mathcal{E} \times \mathcal{M}_d^{\gamma p} \rightarrow \mathbb{R}$  given by

$$\mathcal{Q}_{\rho, h}(\xi, \mu) := \mathcal{R}_\rho((\delta_\xi \otimes \mu) \circ h^{-1})$$

is well defined. The following lemma is known from Claus et al. [9, Theorem 5.2].

**Lemma 5.1** *If conditions (a)–(c) hold true, then the map  $\mathcal{Q}_{\rho, h} : \mathcal{E} \times \mathcal{M}_d^{\gamma p} \rightarrow \mathbb{R}$  is  $((\mathcal{O}_{\mathbb{R}^k} \cap \mathcal{E}) \times \mathcal{O}_d^{\gamma p}, \mathcal{O}_{\mathbb{R}})$ -continuous.*

Lemma 5.1 can be used to obtain the following result on the map

$$\begin{aligned}\mathcal{R}_{\rho,h} : \mathcal{M}_d^{\gamma p} &\rightarrow \mathbb{R} \cup \{-\infty\}, \\ \mathcal{R}_{\rho,h}(\mu) &:= \inf\{\mathcal{Q}_{\rho,h}(\xi, \mu) : \xi \in \mathcal{E}\}.\end{aligned}\quad (5.3)$$

Recall that the set  $\mathcal{E}$  was assumed to be compact.

**Theorem 5.2** *If conditions (a)–(c) hold true, then the infimum in (5.3) is attained for any  $\mu \in \mathcal{M}_d^{\gamma p}$ , and the map  $\mathcal{R}_{\rho,h} : \mathcal{M}_d^{\gamma p} \rightarrow \mathbb{R}$  is  $(\mathcal{O}_d^{\gamma p}, \mathcal{O}_{\mathbb{R}})$ -continuous.*

Note here that if the infimum in (5.3) is attained, then  $\mathcal{R}_{\rho,h}(\mu)$  is a solution to (5.2). Theorem 5.2 is a variant of Claus et al. [9, Corollary 2.4].

**Remark 5.3** The  $(\mathcal{O}_d^{\gamma p}, \mathcal{O}_{\mathbb{R}})$ -continuity of  $\mathcal{R}_{\rho,h} : \mathcal{M}_d^{\gamma p} \rightarrow \mathbb{R}$  obtained in the preceding theorem can be seen as robustness of  $\rho$  relative to  $(\mathcal{G}, Z, \pi_d^{\gamma p})$  in the sense of Embrechts et al. [16, Definition 1], where  $\mathcal{G} := \{h(\xi, \cdot) : \xi \in \mathcal{E}\}$  and  $\pi_d^{\gamma p}$  is any metric metrising the  $(p\gamma)$ -weak topology  $\mathcal{O}_d^{\gamma p}$ .

## 5.2 Example: one-period mean–risk portfolio optimisation

Consider a one-period financial market consisting of one riskless bond and  $d$  risky assets with prices per unit  $S_0^0 := 1$  and  $S_0^1, \dots, S_0^d \in \mathbb{R}_{++}$  at time 0. In between time 0 and time 1, the prices change to  $S_1^0, S_1^1, \dots, S_1^d$  according to  $S_1^i = Z^i S_0^i$ ,  $i = 0, \dots, d$ , where the bond's relative price change  $Z^0$  is deterministic ( $\in \mathbb{R}_{++}$ ) and known at time 0 and the assets' relative price changes  $Z^1, \dots, Z^d$  are  $\mathbb{R}_+$ -valued random variables on a common atomless probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and are unobservable at time 0. Let  $x_0 \in \mathbb{R}_{++}$  be an amount of capital to be invested in the bond and in the  $d$  assets at time 0. If for any  $i = 1, \dots, d$ , the amount of capital invested in the asset  $i$  is denoted by  $\xi_i$ , then the amount of capital invested in the bond is  $\xi_0 := x_0 - \langle \xi, \mathbf{1} \rangle$ , where  $\xi := (\xi_1, \dots, \xi_d)$  and  $\mathbf{1} := (1, \dots, 1) \in \mathbb{R}^d$ . When identifying a portfolio with the corresponding amounts of capital  $\xi_1, \dots, \xi_d$  and assuming that taking loans and short selling are banned, the set

$$\mathcal{E} := \{(\xi_1, \dots, \xi_d) \in \mathbb{R}_+^d : \langle \xi, \mathbf{1} \rangle \leq x_0\} \quad (5.4)$$

can be seen as the set of all admissible portfolios. The realised loss at time 1 of a portfolio  $\xi = (\xi_1, \dots, \xi_d) \in \mathcal{E}$  is given by

$$h(\xi, z) := (x_0 - \langle \xi, \mathbf{1} \rangle)(1 - Z^0) + \langle \xi, \mathbf{1} - z \rangle, \quad (5.5)$$

when  $z = (z_1, \dots, z_d)$  is the vector of the assets' realised relative price changes, i.e., the realisation of  $Z := (Z^1, \dots, Z^d)$ .

Of course, the portfolio  $\xi = (\xi_1, \dots, \xi_d) \in \mathcal{E}$  should be chosen such that the expected profit is as high as possible, i.e., such that the expected loss  $\mathbb{E}[h(\xi, Z)]$  is as small as possible. Simultaneously the portfolio's downside risk should be as small as

possible, where the downside risk can be measured by  $\sigma(h(\xi, Z))$  for a suitable given ‘downside’ risk measure  $\sigma : L^p \rightarrow \mathbb{R}$ . This leads to the *mean–risk model*

$$\min \{ \mathbb{E}[h(\xi, Z)] + \kappa \sigma(h(\xi, Z)) : \xi \in \mathcal{E} \}, \quad (5.6)$$

where  $\kappa \in \mathbb{R}_{++}$  is a risk aversion parameter. Note that the model (5.6) aims at minimising the weighted sum of two competing objects and is in line with Markowitz’ [32] classical mean–variance optimisation theory (where  $\sigma(h(\xi, Z)) = \sqrt{\text{Var}[h(\xi, Z)]}$ ). It is also worth mentioning that mean–risk models are related to the corresponding multiobjective optimisation problems; see for instance the works of Ogryczak and Ruszczyński [37, 38] and Schultz and Tiedemann [46].

The mean–risk model (5.6) coincides with problem (5.2) when  $Z$  is distributed according to  $\mu$  and  $\rho : L^p \rightarrow \mathbb{R}$  is defined by

$$\rho(Y) := \mathbb{E}[Y] + \kappa \sigma(Y), \quad (5.7)$$

where one should note that  $\rho$  is monotone, distribution-invariant and convex if  $\sigma$  is. For any fixed  $p \in [1, \infty)$ , the following corollary is a simple consequence of Theorem 5.2; see Appendix A.14.

**Corollary 5.4** *Let  $\sigma : L^p \rightarrow \mathbb{R}$  be monotone, distribution-invariant and convex. For any  $\mu \in \mathcal{M}_d^p$ , let  $\mathcal{R}_{\rho, h}(\mu)$  be defined by (5.3) (and (5.7)). Then the infimum on the right-hand side of (5.3) is attained (and thus finite) for any  $\mu \in \mathcal{M}_d^p$ , and the map  $\mathcal{R}_{\rho, h} : \mathcal{M}_d^p \rightarrow \mathbb{R}$  is  $(\mathcal{O}_d^p, \mathcal{O}_{\mathbb{R}})$ -continuous.*

**Remark 5.5** If we use  $\mathfrak{A}_d$  to denote the set of all functions  $h(\xi, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $\xi \in \mathcal{E}$ , then  $\mathcal{R}_{\rho, h} = \mathcal{R}_{\rho, \mathfrak{A}_d}$  for the functional  $\mathcal{R}_{\rho, \mathfrak{A}_d}$  defined by (1.3), i.e., by

$$\mathcal{R}_{\rho, \mathfrak{A}_d}(\mu) := \inf \{ \mathcal{R}_{\rho}(\mu \circ A_d^{-1}) : A_d \in \mathfrak{A}_d \} = \inf \{ \mathcal{R}_{\rho, A_d}(\mu) : A_d \in \mathfrak{A}_d \}.$$

Here  $\mathcal{R}_{\rho}$  is derived from  $\rho$  as in (4.1), and  $\mathcal{R}_{\rho, A_d}$  is derived from  $\mathcal{R}_{\rho}$  and  $A_d$  as in (4.2).

### 5.3 Copula robustness of stochastic programming problems

In the setting of Sect. 5.1, assume that conditions (a)–(c) are satisfied and recall that  $\mathcal{E}$  was assumed to be compact. Then by Theorem 5.2, the map  $\mathcal{R}_{\rho, h} : \mathcal{M}_d^{\gamma p} \rightarrow \mathbb{R}$  is  $(\mathcal{O}_d^{\gamma p}, \mathcal{O}_{\mathbb{R}})$ -continuous. Together with Theorem 3.12, this leads to the following result.

**Corollary 5.6** *If conditions (a)–(c) hold true, then the map  $\mathcal{R}_{\rho, h} : \mathcal{M}_d^{\gamma p} \rightarrow \mathbb{R}$  is copula robust.*

Corollary 5.6 shows that under conditions (a)–(c), the minimal value of problem (5.2) is robust with respect to slight changes in the copula of  $\mu$ .

**Example 5.7** Let us return to the specific setting of Sect. 5.2 (mean–risk portfolio optimisation), where  $\mu$  played the role of the joint distribution of the relative price changes  $(Z_1, \dots, Z_d)$ . In this framework, it can be seen in the proof of Corollary 5.4 that conditions (a)–(c) are satisfied for  $p = 1$  when  $\sigma : L^p \rightarrow \mathbb{R}$  is monotone, distribution-invariant and convex. Thus under the latter assumptions on  $\sigma$ , Corollary 5.6 ensures that the functional  $\mathcal{R}_{\rho,h} : \mathcal{M}_d^p \rightarrow \mathbb{R}$  is copula robust. Of course, the copula robustness of  $\mathcal{R}_{\rho,h}$  also directly follows from Theorem 3.12 and Corollary 5.4.

## 6 Example 3: multi-period portfolio optimisation

In this section, the objective is to show that the maximal expected utility of the terminal wealth of a portfolio in a multi-period financial market model (see Sect. 6.2) is copula robust if it is regarded as a function of the joint distribution of the assets' relative price changes. The terminal wealth portfolio optimisation problem can be regarded as a Markov decision problem as introduced in the textbook by Bäuerle and Rieder [1, Chaps. 1 and 2] and in other standard monographs. To prove the main result of this section (Corollary 6.8), it is therefore useful to first establish a variant of a result of Müller [35, Theorem 4.2] about the dependence of the value function on the Markov transition probability function. This variant can be found in Theorem 6.2 and is of independent interest. It is worth pointing out that the factor  $1/\psi(x)$  on the right-hand side of (6.3) is essential for our purposes; see the proof of Corollary 6.7.

### 6.1 Groundwork: a class of Markov decision models

#### 6.1.1 Basic notation and terminology

Let  $(E, \mathcal{E})$  be a measurable space, to be regarded as the state space, and  $N \in \mathbb{N}$  the fixed finite planning horizon. For each  $n = 0, \dots, N - 1$  and  $x \in E$ , let  $A_n(x)$  be a nonempty set whose elements are regarded as the admissible actions at time  $n$  in state  $x$ . For each  $n = 0, \dots, N - 1$ , let  $A_n := \bigcup_{x \in E} A_n(x)$  and  $D_n := \{(x, a) \in E \times A_n : a \in A_n(x)\}$ . The elements of  $A_n$  can be seen as the actions that may basically be selected at time  $n$ , whereas the elements of  $D_n$  are the possible state–action combinations at time  $n$ . We equip  $A_n$  with a  $\sigma$ -algebra  $\mathcal{A}_n$  and  $D_n$  with the trace  $\sigma$ -algebra  $\mathcal{D}_n := (\mathcal{E} \otimes \mathcal{A}_n) \cap D_n$ . We use  $\mathbf{A}$  to denote the family that consists of all sets  $A_n(x)$ ,  $n = 0, \dots, N - 1$ ,  $x \in E$ , and of all  $\sigma$ -algebras  $\mathcal{A}_n$ ,  $n = 0, \dots, N - 1$ . All the sets and spaces just introduced are fully determined by  $(E, \mathcal{E})$  and  $\mathbf{A}$ . Although all the objects introduced in what follows depend on  $(E, \mathcal{E})$  and  $\mathbf{A}$ , we suppress this dependence in the notation.

By a (Markov decision) transition function associated with  $(E, \mathcal{E})$  and  $\mathbf{A}$ , we mean an  $N$ -tuple  $P = (P_n)_{n=0}^{N-1}$ , where  $P_n$  is a probability kernel from  $(D_n, \mathcal{D}_n)$  to  $(E, \mathcal{E})$ , to be seen as the one-step transition kernel at time  $n$ . The set of all transition functions is denoted by  $\overline{P}$ . The actions are governed by a so called  $N$ -stage strategy, i.e., by an  $N$ -tuple  $\pi = (f_0, \dots, f_{N-1})$  where  $f_n$  is a decision rule at time  $n$ , i.e., an  $(\mathcal{E}, \mathcal{A}_n)$ -measurable map  $f_n : E \rightarrow A_n$  satisfying  $f_n(x) \in A_n(x)$  for all  $x \in E$ . Let  $F_n$  be a nonempty set of decision rules at time  $n$ , and define the set of all ‘admissible’ strategies by  $\Pi := F_0 \times \dots \times F_{N-1}$ .

For any  $P = (P_n)_{n=0}^{N-1} \in \overline{\mathcal{P}}$ ,  $\pi = (f_n)_{n=0}^{N-1} \in \Pi$  and  $n = 0, \dots, N-1$ , define the probability kernel  $P_n^\pi$  from  $(E, \mathcal{E})$  to  $(E, \mathcal{E})$  by  $P_n^\pi(x, B) := P_n((x, f_n(x)), B)$ ,  $x \in E$ ,  $B \in \mathcal{E}$ . The probability measure  $P_n^\pi(x, \cdot)$  can be seen as the one-step transition probability at time  $n$  given state  $x$  when the actions are chosen according to  $\pi$ . On the measurable space  $(\Omega, \mathcal{F}) := (E^{N+1}, \mathcal{E}^{\otimes(N+1)})$ , we can define for any  $x_0 \in E$  and  $\pi \in \Pi$  the probability measure  $\mathbb{P}^{x_0, P; \pi} := \delta_{x_0} \otimes P_0^\pi \otimes \dots \otimes P_{N-1}^\pi$ , where the right-hand side is the usual product of the probability measure  $\delta_{x_0}$  and the kernels  $P_0^\pi, \dots, P_{N-1}^\pi$ . Under the probability measure  $\mathbb{P}^{x_0, P; \pi}$ , the identity map  $X = (X_n)_{n=0}^N$  on  $\Omega$  is called *Markov decision process (MDP)* associated with initial state  $x_0$ , transition function  $P$  and strategy  $\pi$ .

Let  $r_n : D_n \rightarrow \mathbb{R}$  be a  $(\mathcal{D}_n, \mathcal{B}(\mathbb{R}))$ -measurable map, referred to as *one-stage reward function*, and  $r_N : E \rightarrow \mathbb{R}$  an  $(\mathcal{E}, \mathcal{B}(\mathbb{R}))$ -measurable map, referred to as *terminal reward function*. Here  $r_n(x, a)$  specifies the one-stage reward when action  $a$  is taken at time  $n$  in state  $x$ , and  $r_N(x)$  specifies the reward of being in state  $x$  at the terminal time  $N$ . Finally, set  $\vec{r} := (r_n)_{n=0}^N$ .

For any fixed subset  $\mathcal{P} \subseteq \overline{\mathcal{P}}$ , the collection of the objects  $(E, \mathcal{E})$ ,  $\mathbf{A}$ ,  $\Pi$ ,  $\mathcal{P}$ ,  $\{\mathbb{P}^{x_0, P; \pi} : x_0 \in E, P \in \mathcal{P}, \pi \in \Pi\}$ ,  $X$  and  $\vec{r}$  introduced so far are often referred to as *Markov decision model*. In fact, in the standard literature, the set  $\mathcal{P}$  is typically a singleton. If, however, there is uncertainty with respect to the ‘true’ transition function, then one should allow a whole bundle of transition functions in the model.

### 6.1.2 Intrinsic optimisation problem

We assume that  $r_k(X_k, f_k(X_k))$ ,  $k = 0, \dots, N-1$ , and  $r_N(X_N)$  are  $\mathbb{P}^{x_0, P; \pi}$ -integrable for any  $x_0 \in E$ ,  $P \in \mathcal{P}$ ,  $\pi \in \Pi$  (for a sufficient condition, see Lemma 6.1 below). As a consequence, we can define for any  $P \in \mathcal{P}$  and  $\pi = (f_n)_{n=0}^{N-1} \in \Pi$  an  $(\mathcal{E}, \mathcal{B}(\mathbb{R}))$ -measurable map  $V_0^{P; \pi} : E \rightarrow \mathbb{R}$  by

$$V_0^{P; \pi}(x_0) := \mathbb{E}^{x_0, P; \pi} \left[ \sum_{k=0}^{N-1} r_k(X_k, f_k(X_k)) + r_N(X_N) \right].$$

The value  $V_n^{P; \pi}(x_0)$  specifies the *expected total reward* of  $X$  under  $\mathbb{P}^{x_0, P; \pi}$ . Here ‘under  $\mathbb{P}^{x_0, P; \pi}$ ’ means that  $X$  starts in  $x_0$  and that the random transitions of  $X$  are governed by  $P$  and  $\pi$ . For fixed  $P \in \mathcal{P}$ , it is natural to look for those strategies  $\pi \in \Pi$  for which the expected total reward from time 0 to  $N$  is maximal for a given initial states  $x_0 \in E$ . This results in the optimisation problem

$$\max\{V_0^{P; \pi}(x_0) : \pi \in \Pi\}. \quad (6.1)$$

We assume that  $\sup_{\pi \in \Pi} V_0^{P; \pi}(x_0) < \infty$  for any  $x_0 \in E$ , which means that it is impossible to gain an arbitrarily high reward. A strategy  $\pi^P \in \Pi$  is said to be *optimal for (6.1)* if  $V_0^{P; \pi^P}(x_0) = V_0^P(x_0)$  for any  $x_0 \in E$ , where the map  $V_0^P : E \rightarrow \mathbb{R}$  is defined by  $V_0^P(x_0) := \sup_{\pi \in \Pi} V_0^{P; \pi}(x_0)$ . The map  $V_0^P$  is referred to as *value function*.

Some known facts about the existence of optimal strategies are recalled in Appendix C. Part (i) of Theorem C.1 shows that under some assumptions, the value

function can be obtained by the Bellman iteration scheme. The latter involves the time- $n$  value functions  $V_n^P$ ,  $n = 1, \dots, N$ , defined by  $V_n^P(x) := \sup_{\pi \in \Pi} V_n^{P;\pi}(x)$ , where for any  $\pi = (f_n)_{n=0}^{N-1} \in \Pi$  the  $(\mathcal{E}, \mathcal{B}(\mathbb{R}))$ -measurable map  $V_n^{P;\pi} : E \rightarrow \mathbb{R}$  is defined by  $V_n^{P;\pi}(x) := \mathbb{E}^{x_0, P; \pi} [\sum_{k=n}^{N-1} r_k(X_k, f_k(X_k)) + r_N(X_N) | X_n = x]$  (note that the right-hand side is independent of  $x_0 \in E$ ). Here and in the following, we use the convention  $\sum_{n=N}^{N-1} := 0$ . The maps  $V_n^{P;\pi}(\cdot)$ ,  $\pi \in \Pi$ , are sometimes called policy value functions and appear in Theorem 6.2.

### 6.1.3 Bounding function

For the Markov decision model introduced above and  $P \in \mathcal{P}$ , an  $(\mathcal{E}, \mathcal{B}([1, \infty)))$ -measurable function  $\psi : E \rightarrow [1, \infty)$  is called a *bounding function for  $P$*  if there exist constants  $K_1, K_2, K_3 \in \mathbb{R}_+$  such that the following three assertions hold:

- (a)  $|r_n(x, a)| \leq K_1 \psi(x)$  for any  $n = 0, \dots, N-1$  and  $(x, a) \in D_n$ ;
- (b)  $|r_N(x)| \leq K_2 \psi(x)$  for any  $x \in E$ ;
- (c)  $\int_E \psi(y) P_n((x, a), dy) \leq K_3 \psi(x)$  for any  $n = 0, \dots, N-1$  and  $(x, a) \in D_n$ .

This terminology is adapted from the work of Müller [34, Definition 2.4] and the textbook by Bäuerle and Rieder [1, Definition 2.4.1]. Denote by  $\mathbb{M}(E)$  the set of all  $(\mathcal{E}, \mathcal{B}(\mathbb{R}))$ -measurable maps  $v : E \rightarrow \mathbb{R}$ , and by  $\mathbb{M}_\psi(E)$  the set of all  $v \in \mathbb{M}(E)$  satisfying  $\|v\|_\psi < \infty$ , where  $\|v\|_\psi := \sup_{x \in E} |v(x)|/\psi(x)$ .

**Lemma 6.1** *Let  $P \in \mathcal{P}$ . If there exists a bounding function  $\psi$  for  $P$ , then the random variables  $r_k(X_k, f_k(X_k))$ ,  $k = 0, \dots, N-1$ , and  $r_N(X_N)$  are  $\mathbb{P}^{x_0, P; \pi}$ -integrable for any  $x_0 \in E$  and  $\pi \in \Pi$ , and moreover,  $\|V_n^P\|_\psi < \infty$  (in particular,  $V_n^{P;\pi} \in \mathbb{M}_\psi(E)$  for any  $\pi \in \Pi$ ) for any  $n = 0, \dots, N-1$ .*

### 6.1.4 Continuous dependence of the optimal value on the transition function

Let  $\psi : E \rightarrow [1, \infty)$  be an  $(\mathcal{E}, \mathcal{B}([1, \infty)))$ -measurable function, and note that the integral  $\int_E v \, d\mathbf{m}$  exists and is finite for any  $v \in \mathbb{M}_\psi(E)$  and  $\mathbf{m} \in \mathcal{M}_1^\psi(E)$ , the set of all probability measures on  $(E, \mathcal{E})$  with  $\int \psi \, d\mathbf{m} < \infty$ . For any fixed subset  $\mathbb{M} \subseteq \mathbb{M}_\psi(E)$ , the distance between  $\mathbf{m}_1$  and  $\mathbf{m}_2$  from  $\mathcal{M}_1^\psi(E)$  can be measured by

$$d_{\mathbb{M}}(\mathbf{m}_1, \mathbf{m}_2) := \sup_{v \in \mathbb{M}} \left| \int_E v \, d\mathbf{m}_1 - \int_E v \, d\mathbf{m}_2 \right|. \quad (6.2)$$

Note that (6.2) defines a probability pseudo-metric (in the sense of Rachev [41, Sect. 2.3]), i.e., a map  $d_{\mathbb{M}} : \mathcal{M}_1^\psi(E) \times \mathcal{M}_1^\psi(E) \rightarrow \overline{\mathbb{R}}_+$  which is symmetric and fulfils the triangle inequality. If  $\mathbb{M}$  separates points in  $\mathcal{M}_1^\psi(E)$  (i.e., if any two  $\mathbf{m}_1, \mathbf{m}_2 \in \mathcal{M}_1^\psi(E)$  coincide when  $\int_E v \, d\mathbf{m}_1 = \int_E v \, d\mathbf{m}_2$  for all  $v \in \mathbb{M}$ ), then  $d_{\mathbb{M}}$  is even a probability metric. It is sometimes called integral probability metric or probability metric with a  $\zeta$ -structure; see Müller [35] and Zolotarev [54].

In some situations, the (pseudo-)metric  $d_{\mathbb{M}}$  (with  $\mathbb{M} \subseteq \mathbb{M}_\psi(E)$  fixed) can be represented by the right-hand side of (6.2) with  $\mathbb{M}$  replaced by a different subset  $\mathbb{M}'$

of  $\mathbb{M}_\psi(E)$ . Each such set  $\mathbb{M}'$  is said to be a *generator* of  $d_{\mathbb{M}}$ . The largest generator of  $d_{\mathbb{M}}$  is called the *maximal generator* of  $d_{\mathbb{M}}$  and will be denoted by  $\overline{\mathbb{M}}$ . That is,  $\overline{\mathbb{M}}$  is the set of all  $v \in \mathbb{M}_\psi(E)$  for which  $|\int_E v d\mathbf{m}_1 - \int_E v d\mathbf{m}_2| \leq d_{\mathbb{M}}(\mathbf{m}_1, \mathbf{m}_2)$  for all  $\mathbf{m}_1, \mathbf{m}_2 \in \mathcal{M}_1^\psi(E)$ ; see [35, Definition 3.1]. Examples for  $d_{\mathbb{M}}$  and  $\overline{\mathbb{M}}$  are discussed in Kern et al. [26] and Müller [34, 35].

Now denote by  $\mathcal{P}_\psi$  the set of all transition functions  $P = (P_n)_{n=0}^{N-1} \in \mathcal{P}$  with  $P_n((x, a), \cdot) \in \mathcal{M}_1^\psi(E)$  for all  $(x, a) \in D_n$  and  $n = 0, \dots, N-1$ . For any  $P \in \mathcal{P}_\psi$ , the integrals  $\int_E v(y) P_n((x, a), dy)$ ,  $v \in \mathbb{M}_\psi(E)$ ,  $(x, a) \in D_n$ ,  $n = 0, \dots, N-1$ , exist and are finite. For any  $\mathbb{M} \subseteq \mathbb{M}_\psi(E)$ , we may define the distance between two transition functions  $P = (P_n)_{n=0}^{N-1}$  and  $Q = (Q_n)_{n=0}^{N-1}$  from  $\mathcal{P}_\psi$  by

$$d_{\mathbb{M}, \psi}(P, Q) := \max_{n=0, \dots, N-1} \sup_{(x, a) \in D_n} d_{\mathbb{M}}(P_n((x, a), \cdot), Q_n((x, a), \cdot)) / \psi(x). \quad (6.3)$$

For any  $\mathbb{M} \subseteq \mathbb{M}_\psi(E)$ , the *Minkowski functional*  $\varrho_{\mathbb{M}} : \mathbb{M}_\psi(E) \rightarrow \overline{\mathbb{R}}_+$  (in the sense of Rudin [42, paragraph after Definition 1.33]) is defined by

$$\varrho_{\mathbb{M}}(v) := \inf\{\lambda \in \mathbb{R}_{++} : v/\lambda \in \mathbb{M}\},$$

where we set  $\inf \emptyset := \infty$ . Examples for  $\mathbb{M}$  and  $\varrho_{\mathbb{M}}$  are discussed in Kern et al. [26] and Müller [34]. In the following result, we assume that  $\psi$  is a bounding function for any  $Q \in \mathcal{P}_\psi$ . By Lemma 6.1, it then follows that  $V_n^Q(x) < \infty$  for any  $n = 0, \dots, N$ ,  $Q \in \mathcal{P}_\psi$  and  $x \in E$ . In particular, we can define a functional  $\overline{V}_n^x : \mathcal{P}_\psi \rightarrow \mathbb{R}$  by  $\overline{V}_n^x(Q) := V_n^Q(x)$ . Note that Theorem 6.2 is a refinement of Kern's PhD thesis [25, Theorem 2.2.8] and that a related result was proved earlier by Müller [34, Theorem 4.2]. We use  $K_{3,P}$  to denote the constant in condition (c) of a bounding function for  $P$ .

**Theorem 6.2** *We assume that  $\psi$  is a bounding function for any  $Q \in \mathcal{P}_\psi$ , and we let  $\mathbb{M} \subseteq \mathbb{M}_\psi(E)$  and  $\mathbb{M}'$  be a generator of  $d_{\mathbb{M}}$ . Then for any  $n = 0, \dots, N-1$ ,  $x_n \in E$  and  $Q, P \in \mathcal{P}_\psi$ , we have*

$$\begin{aligned} & |\overline{V}_n^{x_n}(Q) - \overline{V}_n^{x_n}(P)| \\ & \leq \sum_{j=n}^{N-1} \sup_{\pi \in \Pi} \varrho_{\mathbb{M}'}(V_{j+1}^{P; \pi})(K_{3,P} + \varrho_{\mathbb{M}'}(\psi) d_{\mathbb{M}, \psi}(Q, P))^{n-j} \psi(x_n) d_{\mathbb{M}, \psi}(Q, P). \end{aligned}$$

As a direct consequence of Theorem 6.2, we get the following result.

**Corollary 6.3** *Assume that  $\psi$  is a bounding function for any  $Q \in \mathcal{P}_\psi$  and let  $P \in \mathcal{P}_\psi$ . Let  $\mathbb{M} \subseteq \mathbb{M}_\psi(E)$  and  $\mathbb{M}'$  be a generator of  $d_{\mathbb{M}}$ . If  $\varrho_{\mathbb{M}'}(\psi) < \infty$  and  $\sup_{\pi \in \Pi} \varrho_{\mathbb{M}'}(V_{n+1}^{P; \pi}) < \infty$  for any  $n = 0, \dots, N-1$ , then  $\overline{V}_n^{x_n}$  is  $(d_{\mathbb{M}, \psi}, |\cdot|)$ -continuous at  $P$  for any  $n = 0, \dots, N-1$  and  $x_n \in E$ .*



## 6.2 A utility-based portfolio optimisation problem

### 6.2.1 Financial market model and a terminal wealth optimisation problem

Consider an  $N$ -period financial market consisting of one riskless bond  $S^0 = (S_n^0)_{n=0}^N$  and  $d$  risky assets  $S^i = (S_n^i)_{n=0}^N$ ,  $i = 1, \dots, d$ , for some fixed  $d \in \mathbb{N}$ . Assume that the value of the bond evolves deterministically according to

$$S_0^0 = 1 \quad \text{and} \quad S_{n+1}^0 = Z_{n+1}^0 S_n^0, \quad n = 0, \dots, N-1,$$

for some fixed constants  $Z_1^0, \dots, Z_N^0 \in [1, \infty)$ , and that the value of the  $i$ th asset evolves stochastically according to

$$S_0^i = s_0^i \quad \text{and} \quad S_{n+1}^i = Z_{n+1}^i S_n^i, \quad n = 0, \dots, N-1,$$

for a constant  $s_0^i \in \mathbb{R}_{++}$  and independent  $\mathbb{R}_+$ -valued random variables  $Z_1^i, \dots, Z_N^i$  on a common probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . For  $n = 0, \dots, N$ , set  $S_n := (S_n^1, \dots, S_n^d)$  and  $Z_n := (Z_n^1, \dots, Z_n^d)$  and denote by  $\mu_n$  the distribution of  $Z_n$ . We also define  $\mathbf{1} := (1, \dots, 1) \in \mathbb{R}^d$ ,  $\mathcal{F}_0 := \{\emptyset, \Omega\}$ ,  $\mathcal{F}_n := \sigma(S_0, \dots, S_n) = \sigma(Z_1, \dots, Z_n)$ ,  $n = 1, \dots, N$ , and  $\mathbb{F} := (\mathcal{F}_n)_{n=0}^N$ .

Now, an agent invests a given amount of capital  $x_0 \in \mathbb{R}_{++}$  in the bond and the assets according to some self-financing trading strategy. By a *trading strategy*, we mean an  $\mathbb{F}$ -adapted  $\mathbb{R}_+^{d+1}$ -valued stochastic process  $\xi = (\xi_n^0, \xi_n)_{n=0}^{N-1}$  with  $\xi_n = (\xi_n^1, \dots, \xi_n^d)$ , where  $\xi_n^0$  and  $\xi_n^i$  specify the amounts of capital invested in the bond and in the  $i$ th asset, respectively, during the time interval  $[n, n+1)$ . The non-negativity of  $\xi_n^0, \xi_n^1, \dots, \xi_n^d$ ,  $n = 0, \dots, N-1$ , means that taking loans and short selling of the assets are excluded. The corresponding ( $\mathbb{F}$ -adapted) *portfolio process*  $X^\xi = (X_n^\xi)_{n=0}^N$  associated with  $\xi = (\xi_n^0, \xi_n)_{n=0}^{N-1}$  is defined by

$$X_0^\xi := \xi_0^0 + \langle \xi_0, \mathbf{1} \rangle, \quad X_{n+1}^\xi := \xi_{n+1}^0 Z_{n+1}^0 + \langle \xi_n, Z_{n+1} \rangle, \quad n = 0, \dots, N-1. \quad (6.4)$$

A trading strategy  $\xi = (\xi_n^0, \xi_n)_{n=0}^{N-1}$  is called *self-financing with respect to the initial capital*  $x_0$  if  $x_0 = \xi_0^0 + \langle \xi_0, \mathbf{1} \rangle$  and  $X_n^\xi = \xi_n^0 + \langle \xi_n, \mathbf{1} \rangle$  for any  $n = 1, \dots, N$ . Note that  $\xi_n^0$  and  $\langle \xi_n, \mathbf{1} \rangle$  specify the amounts of capital invested during the time interval  $[n, n+1)$  in the bond and in the  $d$  assets, respectively. For any self-financing trading strategy  $\xi = (\xi_n^0, \xi_n)_{n=0}^{N-1}$  with respect to  $x_0$ , we have  $\xi_n^0 = X_n^\xi - \langle \xi_n, \mathbf{1} \rangle$  for any  $n = 0, \dots, N-1$ , and therefore the corresponding portfolio process admits the representation

$$X_0^\xi = x_0, \quad X_{n+1}^\xi = Z_{n+1}^0 X_n^\xi + \langle \xi_n, Z_{n+1} - Z_{n+1}^0 \mathbf{1} \rangle \quad \text{for } n = 0, \dots, N-1. \quad (6.5)$$

In view of (6.5), we identify a self-financing trading strategy with respect to  $x_0$  with an  $\mathbb{F}$ -adapted  $\mathbb{R}_+^d$ -valued stochastic process  $\xi = (\xi_n)_{n=0}^{N-1}$  with  $\xi_n = (\xi_n^1, \dots, \xi_n^d)$  such that  $\langle \xi_0, \mathbf{1} \rangle \in [0, x_0]$  and  $\langle \xi_n, \mathbf{1} \rangle \in [0, X_n^\xi]$  for any  $n = 1, \dots, N-1$ . We restrict ourselves to *Markovian* self-financing trading strategies  $\xi = (\xi_n)_{n=0}^{N-1}$  with respect to  $x_0$  which means that  $\xi_n$  only depends on  $n$  and  $X_n^\xi$ . To put it another way, we assume

that for any  $n = 0, \dots, N-1$ , there exists a Borel-measurable map  $f_n : \mathbb{R}_+ \rightarrow \mathbb{R}_+^d$  such that  $\xi_n = f_n(X_n^\xi)$ . Then in particular,  $X^\xi$  is an  $\mathbb{R}_+$ -valued  $\mathbb{F}$ -Markov process whose one-step transition probability at time  $n \in \{0, \dots, N-1\}$  given state  $x \in \mathbb{R}_+$  and strategy  $\xi = (\xi_n)_{n=0}^{N-1}$  (resp.  $\pi := (f_n)_{n=0}^{N-1}$ ) is given by  $\mu_{n+1} \circ \eta_{n,(x,f_n(x))}^{-1}$ , where

$$\eta_{n,(x,a)}(z) := Z_{n+1}^0 x + \langle a, z - Z_{n+1}^0 \mathbf{1} \rangle, \quad z \in \mathbb{R}_+^d. \quad (6.6)$$

The agent's aim is to find a self-financing trading strategy  $\xi = (\xi_n)_{n=0}^{N-1}$  (resp.  $\pi = (f_n)_{n=0}^{N-1}$ ) with respect to  $x_0$  for which her expected utility of the relative terminal wealth is maximised. We assume that the agent is risk-averse and that her attitude towards risk is set via the *power utility* function  $u_\alpha : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  defined by

$$u_\alpha(x) := x^\alpha \quad (6.7)$$

for some fixed  $\alpha \in (0, 1)$ . Hence the agent is interested in those self-financing trading strategies  $\xi = (\xi_n)_{n=0}^{N-1}$  (resp.  $\pi = (f_n)_{n=0}^{N-1}$ ) with respect to  $x_0$  for which the expectation of  $u_\alpha(X_N^\xi / (x_0 S_N^0))$  is maximised. Since  $u_\alpha(X_N^\xi / (x_0 S_N^0)) = u_\alpha(X_N^\xi) / (x_0 S_N^0)^\alpha$ , this is equivalent to maximising the expectation of  $u_\alpha(X_N^\xi)$ . For notational simplicity, we consider the terminal wealth optimisation problem in the latter form. We assume that  $Z_n^1, \dots, Z_n^d$  are  $\mathbb{P}$ -a.s. strictly positive and  $\mathbb{E}[u_\alpha(\langle Z_n, \mathbf{1} \rangle)] < \infty$  for any  $n = 1, \dots, N$ .

**Example 6.4** Assume that the bond and the  $d$  assets evolve according to the 1-dimensional ordinary (Itô stochastic) differential equations

$$\begin{aligned} ds_t^0 &= \delta_0 s_t^0 dt, \quad s_0^0 = 1, \\ ds_t^i &= \delta_i s_t^i dt + \sigma_i s_t^i dB_t^i, \quad s_0^i = s_0^i, \quad i = 1, \dots, d, \end{aligned}$$

where  $\delta_0, \delta_1, \dots, \delta_d, \sigma_1, \dots, \sigma_d \in \mathbb{R}_{++}$  are constants and  $B^1, \dots, B^d$  are (jointly Gaussian) correlated 1-dimensional standard Brownian motions which satisfy for any  $t \in \mathbb{R}_+$  that  $\text{Cov}(B_t^i, B_t^j) = R_{i,j}t$ , where  $R = (R_{i,j})_{1 \leq i, j \leq d} \in \mathbb{R}^{d \times d}$  is a fixed correlation matrix (i.e.,  $R$  is symmetric and positive semi-definite with entries 1 on the diagonal). This is a multivariate version of the classical *Black–Scholes–Merton model*. Choose the trading period to be the unit interval  $[0, 1]$  and assume that the bond and the assets can be traded only at  $N$  equidistant time points in  $[0, 1]$ , namely at  $t_{N,n} := n/N$ ,  $n = 0, \dots, N-1$ . Then the relative price changes  $Z_{n+1}^0 := S_{n+1}^0 / S_n^0 = s_{t_{N,n+1}}^0 / s_{t_{N,n}}^0$  and  $Z_{n+1}^i := S_{n+1}^i / S_n^i = s_{t_{N,n+1}}^i / s_{t_{N,n}}^i$  are given by, respectively,  $e^{\delta_0(t_{N,n+1} - t_{N,n})}$  and  $e^{(\delta_i - \sigma_i^2/2)(t_{N,n+1} - t_{N,n}) + \sigma_i(B_{t_{N,n+1}}^i - B_{t_{N,n}}^i)}$ , i.e., for  $n = 0, \dots, N-1$ ,

$$Z_{n+1}^0 = e^{\delta_0/N} \quad \text{and} \quad Z_{n+1}^i = e^{(\delta_i - \sigma_i^2/2)/N + \sigma_i(B_{t_{N,n+1}}^i - B_{t_{N,n}}^i)}, \quad i = 1, \dots, d.$$

That is, we have  $Z_{n+1} = (e^{G_1}, \dots, e^{G_d})$  for a  $d$ -variate random variable  $(G_1, \dots, G_d)$  which has a  $d$ -variate normal distribution  $\mathbf{N}_{\delta, \Gamma}$  with  $\delta := ((\delta_i - \sigma_i^2/2)/N)_{i=1}^d$  and  $\Gamma := (\sigma_i R_{i,j} \sigma_j / N)_{1 \leq i, j \leq d}$ . Thus we have  $\mu_1 = \dots = \mu_N = \text{LN}_{\delta, \Gamma}$ , where  $\text{LN}_{\delta, \Gamma}$  is a  $d$ -variate log-normal distribution with parameters  $\delta$  and  $\Gamma$ .

## 6.2.2 Interpretation as a Markov decision problem

The terminal wealth optimisation problem just introduced can be embedded in the framework of Sect. 6.1 as follows. Let  $Z_1^0, \dots, Z_N^0 \in [1, \infty)$  be a priori fixed and choose  $(E, \mathcal{E}) := (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ . For any  $x \in \mathbb{R}_+$  and  $n = 0, \dots, N - 1$ , let

$$A_n(x) := A(x) := \{a \in \mathbb{R}_+^d : \langle a, \mathbf{1} \rangle \leq x\}.$$

Hence  $A_n = \mathbb{R}_+^d$  and  $D_n = D := \{(x, a) \in \mathbb{R}_+^{d+1} : a \in A(x)\}$  for  $n = 0, \dots, N - 1$ . Let  $\mathcal{A}_n := \mathcal{B}(\mathbb{R}_+^d)$  and  $\mathcal{D}_n := \mathcal{B}(\mathbb{R}_+^{d+1}) \cap D$  for any  $n = 0, \dots, N - 1$ , and let the set  $F$  consist of all those Borel-measurable maps  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+^d$  that satisfy  $\langle f(x), \mathbf{1} \rangle \in [0, x]$  for any  $x \in \mathbb{R}_+$ . Finally, let  $F_n := F$  for  $n = 0, \dots, N - 1$  and  $\Pi := F_0 \times \dots \times F_{N-1} = F^N$ .

Let  $\mathcal{M}_1^\alpha(\mathbb{R}_{++}^d)$  be the set of all Borel probability measures on  $\mathbb{R}_{++}^d$  for which  $\int_{\mathbb{R}_+^d} |z|^\alpha \mu(dz) < \infty$ . The latter condition is equivalent to  $\int_{\mathbb{R}_{++}^d} \langle z, \mathbf{1} \rangle^\alpha \mu(dz) < \infty$ , which can be shown by using arguments as at the beginning of Appendix A.2. For any  $\vec{\mu} = (\mu_n)_{n=1}^N \in \mathcal{M}_1^\alpha(\mathbb{R}_{++}^d)^N$ , we define a transition function  $P^{\vec{\mu}} = (P_n^{\vec{\mu}})_{n=0}^{N-1}$  by

$$P_n^{\vec{\mu}}((x, a), \cdot) = (\mu_{n+1} \circ \eta_{n,(x,a)}^{-1})[\cdot], \quad (x, a) \in D_n, n = 0, \dots, N - 1,$$

where the map  $\eta_{n,(x,a)} : \mathbb{R}_+^d \rightarrow \mathbb{R}_+$  is defined by (6.6). The set of all such transition functions is denoted by  $\mathcal{P}_\alpha$ , i.e.,  $\mathcal{P}_\alpha := \{P^{\vec{\mu}} : \vec{\mu} \in \mathcal{M}_1^\alpha(\mathbb{R}_{++}^d)^N\}$ , and plays the role of  $\mathcal{P}$ .

Let  $r_n := 0, n = 0, \dots, N - 1$ , and  $r_N(x) := u_\alpha(x), x \in \mathbb{R}_+$ . Then

$$V_0^{P;\pi}(x_0) = \mathbb{E}^{x_0, P; \pi}[r_N(X_N)] = \mathbb{E}^{x_0, P; \pi}[u_\alpha(X_N)]$$

for any  $x_0 \in \mathbb{R}_+$ ,  $P \in \mathcal{P}_\alpha$  and  $\pi \in \Pi$ , and the terminal wealth problem introduced subsequent to (6.7) can be identified with the optimisation problem (6.1), i.e., with

$$\max\{\mathbb{E}^{x_0, P; \pi}[u_\alpha(X_N)] : \pi \in \Pi\} \quad (6.8)$$

for any  $x_0 \in \mathbb{R}_+$  and  $P \in \mathcal{P}_\alpha$ . A strategy  $\pi^P \in \Pi$  is called an *optimal (self-financing) trading strategy for  $P$*  if it solves the maximisation problem (6.8) for any  $x_0 \in \mathbb{R}_+$ . Note that the coordinate process  $X$  plays the role of the portfolio process  $X^\xi$  introduced in (6.4), and that for each  $x_0 \in \mathbb{R}_+$ , any self-financing trading strategy  $\xi = (\xi_n)_{n=0}^{N-1}$  with respect to  $x_0$  may be identified with some  $\pi = (f_n)_{n=0}^{N-1} \in \Pi$  through  $\xi_n = f_n(X_n^\xi)$ . Theorem C.3 ensures that optimal trading strategies exist.

## 6.2.3 Continuous dependence of the optimal value on $P^{\vec{\mu}}$

Let the function  $\psi_\alpha : \mathbb{R}_+ \rightarrow [1, \infty)$  be defined by  $\psi_\alpha(x) := 1 + u_\alpha(x)$ . Moreover, let  $\mathcal{P}_{\psi_\alpha}$  be derived from  $\mathcal{P}_\alpha$  as  $\mathcal{P}_\psi$  is derived from  $\mathcal{P}$  in Sect. 6.1.

**Lemma 6.5**  $\psi_\alpha$  is a bounding function for any  $P \in \mathcal{P}_\alpha$ , and we have  $\mathcal{P}_{\psi_\alpha} = \mathcal{P}_\alpha$ .

Let  $\mathbb{M} := \mathbb{M}_{\text{Hö},\alpha} := \{v \in \mathbb{R}^{\mathbb{R}_+} : \|v\|_{\text{Hö},\alpha} \leq 1\}$ , where the Hölder- $\alpha$  norm is defined by  $\|v\|_{\text{Hö},\alpha} := \sup_{x,y \in \mathbb{R}_+ : x \neq y} |v(x) - v(y)|/|x - y|^\alpha$ . We obviously have  $\mathbb{M}_{\text{Hö},\alpha} \subseteq \mathbb{M}_{\psi_\alpha}(\mathbb{R}_+)$ , and in view of Lemmas 6.5 and 6.1, we can therefore define a functional  $\bar{\mathcal{V}}_n^x : \mathcal{P}_{\psi_\alpha} \rightarrow \mathbb{R}$  through  $\bar{\mathcal{V}}_n^x(P) := V_n^P(x)$ . The set  $\mathbb{M}_{\text{Hö},\alpha}$  separates points in  $\mathcal{M}_1^{\psi_\alpha}(\mathbb{R}_+)$ , implying that  $d_{\mathbb{M}_{\text{Hö},\alpha}}$  (defined by (6.2) with  $\mathbb{M} := \mathbb{M}_{\text{Hö},\alpha}$ ) provides a metric on  $\mathcal{M}_1^{\psi_\alpha}(\mathbb{R}_+)$ ; see Kern et al. [26] for details. Let  $d_{\mathbb{M}_{\text{Hö},\alpha},\psi_\alpha}$  be defined by (6.3) with  $\mathbb{M} := \mathbb{M}_{\text{Hö},\alpha}$  and  $\psi := \psi_\alpha$ .

**Theorem 6.6** *For any  $n = 0, \dots, N - 1$  and  $x \in \mathbb{R}_+$ , the map  $\bar{\mathcal{V}}_n^x : \mathcal{P}_{\psi_\alpha} \rightarrow \mathbb{R}$  is  $(d_{\mathbb{M}_{\text{Hö},\alpha},\psi_\alpha}, |\cdot|)$ -continuous.*

Recall that the elements of  $\mathcal{P}_\alpha (= \mathcal{P}_{\psi_\alpha})$  are parametrised by the elements of the set  $\mathcal{M}_1^\alpha(\mathbb{R}_{++})^N$ . For any  $\mu \in \mathcal{M}_1^\alpha(\mathbb{R}_{++})^N$ , denote by  $\bar{\mu}$  the element of  $\mathcal{M}_1^\alpha(\mathbb{R}_{++})^N$  whose  $N$  entries are all equal to  $\mu$ , i.e.,  $\bar{\mu} := (\mu)_{n=1}^N$ . Then we can define a functional  $\mathcal{V}_n^x : \mathcal{M}_1^\alpha(\mathbb{R}_{++})^d \rightarrow \mathbb{R}$  by

$$\mathcal{V}_n^x(\mu) := \bar{\mathcal{V}}_n^x(\bar{\mu}) = V_n^{P^{\bar{\mu}}}(x). \quad (6.9)$$

Since we used  $\mathcal{O}_d^\alpha$  to denote the  $\alpha$ -weak topology on  $\mathcal{M}_1^\alpha$  (see Sect. 2.2), we use  $\mathcal{O}_d^\alpha(\mathbb{R}_{++}^d)$  to denote the analogous topology on  $\mathcal{M}_1^\alpha(\mathbb{R}_{++}^d)$ .

**Corollary 6.7** *For any  $n = 0, \dots, N - 1$  and  $x \in \mathbb{R}_+$ , the map  $\mathcal{V}_n^x : \mathcal{M}_1^\alpha(\mathbb{R}_{++}^d) \rightarrow \mathbb{R}$  defined by (6.9) is  $(\mathcal{O}_d^\alpha(\mathbb{R}_{++}^d), \mathcal{O}_{\mathbb{R}})$ -continuous.*

### 6.3 Copula robustness of the maximal expected utility of the terminal wealth

For any  $n = 0, \dots, N - 1$  and  $x \in \mathbb{R}_+$ , let the map  $\mathcal{V}_n^x : \mathcal{M}_1^\alpha(\mathbb{R}_{++}^d) \rightarrow \mathbb{R}$  be defined by (6.9), and note that  $\mathcal{V}_0^{x_0}(\mu)$  corresponds to the maximal expected utility of the terminal wealth in (6.8) with  $P = P^{\bar{\mu}}$ . When regarding each  $\mu \in \mathcal{M}_1^\alpha(\mathbb{R}_{++}^d)$  as a Borel probability measure on the whole Euclidean space  $\mathbb{R}^d$  (with  $\mu[\mathbb{R}_{++}^d] = 1$ ), the set  $\mathcal{M}_1^\alpha(\mathbb{R}_{++}^d)$  can be seen as a subset of  $\mathcal{M}_1^\alpha$ . Thus  $\mathbf{C}_d(\mu_1, \dots, \mu_d) = \mathbf{C}_d$  for any  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^\alpha(\mathbb{R}_{++}^d)$ , and Theorem 3.12 and Corollary 6.7 together imply the following result.

**Corollary 6.8** *For any  $n = 0, \dots, N - 1$  and  $x \in \mathbb{R}_+$ , the map  $\mathcal{V}_n^x : \mathcal{M}_1^\alpha(\mathbb{R}_{++}^d) \rightarrow \mathbb{R}$  defined by (6.9) is copula robust.*

## Appendix A: Proofs

### A.1 Proof of Proposition 2.2

By the assumption imposed on  $h$ , we can find a constant  $c \in \mathbb{R}_{++}$  such that

$$\int_{\mathbb{R}^{d'}} |y|^{p'} h(\mu)(dy) = \int_{\mathbb{R}^d} |h(x)|^{p'} \mu(dx) \leq c \left( 1 + \int_{\mathbb{R}^d} |x|^p \mu(dx) \right) < \infty$$

for any  $\mu \in \mathcal{M}_d^p$ . This gives the first assertion. Since the involved topologies are metrisable, it suffices for the second assertion to show that the map  $\mathfrak{h} : \mathcal{M}_d^p \rightarrow \mathcal{M}_{d'}^{p'}$  is sequentially continuous. Let  $\mu$  and  $\mu_n, n \in \mathbb{N}$ , be elements of  $\mathcal{M}_d^p$  such that  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^p$  and thus in particular in  $\mathcal{O}_d^0 \cap \mathcal{M}_1^p$ . By the classical continuous mapping theorem, we have  $\mathfrak{h}(\mu_n) \rightarrow \mathfrak{h}(\mu)$  in  $\mathcal{O}_{d'}^0 \cap \mathcal{M}_{d'}^{p'}$ . Moreover, by the assumption on  $h$ , the function  $f_h : \mathbb{R}^d \rightarrow \mathbb{R}$  defined by  $f_h(x) := |h(x)|^{p'}$  lies in  $\mathcal{C}_d^p$ . This implies

$$\begin{aligned} \lim_{n \rightarrow \infty} \int_{\mathbb{R}^{d'}} |y|^{p'} \mathfrak{h}(\mu_n)(dy) &= \lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} f_h(x) \mu_n(dx) \\ &= \int_{\mathbb{R}^d} f_h(x) \mu(dx) = \int_{\mathbb{R}^{d'}} |y|^{p'} \mathfrak{h}(\mu)(dy). \end{aligned}$$

Thus  $\mathfrak{h}(\mu_n) \rightarrow \mathfrak{h}(\mu)$  in  $\mathcal{O}_{d'}^{p'}$ . This gives the second assertion.  $\square$

## A.2 Proof of Theorem 2.3

We first prove that  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$  if and only if  $\mathfrak{P}_d(C, \mu_1, \dots, \mu_d) \in \mathcal{M}_d^p$ , regardless of the copula  $C \in \mathbf{C}_d$ . If we have  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ , then

$$\begin{aligned} \int_{\mathbb{R}^d} |x|^p \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)(dx) &\leq \int_{\mathbb{R}^d} c_1 |x|_1^p \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)(dx) \\ &= c_1 \int_{\mathbb{R}^d} \left( \sum_{i=1}^d |\pi_i(x)| \right)^p \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)(dx) \\ &\leq c_p \sum_{i=1}^d \int_{\mathbb{R}} |x_i|^p \mu_i(dx_i) < \infty, \end{aligned}$$

where  $c_p := c_1 2^{\max\{0, p-1\}}$ ,  $|x|_1 := \sum_{i=1}^d |x_i|$  is the 1-norm of  $x = (x_1, \dots, x_d)$  and  $c_1 \in \mathbb{R}_{++}$  is a suitable constant. Thus  $\mathfrak{P}_d(C, \mu_1, \dots, \mu_d) \in \mathcal{M}_d^p$ . Conversely, assume that  $\mathfrak{P}_d(C, \mu_1, \dots, \mu_d) \in \mathcal{M}_d^p$ . Then we have

$$\begin{aligned} \int_{\mathbb{R}} |x_i|^p \mu_i(dx_i) &= \int_{\mathbb{R}^d} |\pi_i(x)|^p \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)(dx) \\ &\leq \int_{\mathbb{R}^d} |x|^p \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)(dx) < \infty, \end{aligned}$$

i.e.,  $\mu_i \in \mathcal{M}_1^p$ , for any  $i = 1, \dots, d$ .

To prove the main assertion of Theorem 2.3, we first let  $(C, \mu_1, \dots, \mu_d)$  and  $(C_n, \mu_{n,1}, \dots, \mu_{n,d})$ ,  $n \in \mathbb{N}$ , be elements of  $\mathbf{C}_d \times \mathcal{M}_1^p \times \dots \times \mathcal{M}_1^p$  such that  $\mu_{n,i} \rightarrow \mu_i$  in  $\mathcal{O}_1^p$ ,  $i = 1, \dots, d$ , and  $d_{\mu_1, \dots, \mu_d}(C_n, C) \rightarrow 0$ . Since  $p$ -weak convergence implies weak convergence, we then have in particular that  $\mu_{n,i} \rightarrow \mu_i$  in  $\mathcal{O}_1^0 \cap \mathcal{M}_1^p$ ,  $i = 1, \dots, d$ . So Lindner and Szimayer [31, Theorem 2.1] implies that

$\mathfrak{P}_d(C_n, \mu_{1,n}, \dots, \mu_{d,n}) \rightarrow \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)$  in  $\mathcal{O}_d^0 \cap \mathcal{M}_d^p$ . For the convergence  $\mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d}) \rightarrow \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)$  in  $\mathcal{O}_d^p$  (when  $p > 0$ ), it remains to show the convergence

$$\int_{\mathbb{R}^d} |x|^p \mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d})(dx) \longrightarrow \int_{\mathbb{R}^d} |x|^p \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)(dx).$$

Clearly,

$$\begin{aligned} & \left| \int_{\mathbb{R}^d} |x|^p \mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d})(dx) - \int_{\mathbb{R}^d} |x|^p \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)(dx) \right| \quad (\text{A.1}) \\ & \leq \left| \int_{\mathbb{R}^d} |x|^p \mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d})(dx) \right. \\ & \quad \left. - \int_{\mathbb{R}^d} (|x|^p \wedge a) \mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d})(dx) \right| \\ & \quad + \left| \int_{\mathbb{R}^d} (|x|^p \wedge a) \mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d})(dx) \right. \\ & \quad \left. - \int_{\mathbb{R}^d} (|x|^p \wedge a) \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)(dx) \right| \\ & \quad + \left| \int_{\mathbb{R}^d} (|x|^p \wedge a) \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)(dx) - \int_{\mathbb{R}^d} |x|^p \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)(dx) \right| \\ & =: S_1(n, a) + S_2(n, a) + S_3(a) \end{aligned}$$

for any  $a \in \mathbb{R}_{++}$ . Since  $\mathfrak{P}_d(C, \mu_1, \dots, \mu_d) \in \mathcal{M}_d^p$  (as seen above), we can choose for every  $\varepsilon > 0$  a suitable constant  $a_3 \in \mathbb{R}_{++}$  such that

$$S_3(a_3) \leq \int_{\mathbb{R}^d} |x|^p \mathbb{1}_{(a_3, \infty)}(|x|) \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)(dx) \leq \varepsilon/3.$$

Furthermore, for any  $a_1 \in \mathbb{R}_{++}$ , we have

$$\begin{aligned} & S_1(n, a_1) \\ & \leq \int_{\mathbb{R}^d} |x|^p \mathbb{1}_{(a_1^{1/p}, \infty)}(|x|) \mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d})(dx) \\ & \leq \int_{\mathbb{R}^d} c_1 |x|_1^p \mathbb{1}_{(a_1^{1/p}, \infty)}(c_\infty |x|_\infty) \mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d})(dx) \\ & \leq c_p \sum_{i=1}^d \int_{\mathbb{R}^d} |\pi_i(x)|^p \mathbb{1}_{(a_1^{1/p}/c_\infty, \infty)}(|x|_\infty) \mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d})(dx) \\ & \leq c_p \sum_{i=1}^d \sum_{j=1}^d \int_{\mathbb{R}^d} |\pi_i(x)|^p \mathbb{1}_{(a_1^{1/p}/c_\infty, \infty)}(|\pi_j(x)|) \mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d})(dx), \end{aligned}$$

where  $c_p := c_1 2^{\max\{0, p-1\}}$ ,  $|x|_1 := \sum_{i=1}^d |x_i|$  and  $|x|_\infty := \max_{i=1, \dots, d} |x_i|$  for  $x = (x_1, \dots, x_n)$ , and  $c_1, c_\infty \in \mathbb{R}_{++}$  are suitable constants. For  $i = j$ , the last integral equals  $\int_{\mathbb{R}} |x_i|^p \mathbb{1}_{(a_1^{1/p}/c_\infty, \infty)}(|x_i|) \mu_{n,i}(dx_i)$ . Since we assumed  $\mu_{n,i} \rightarrow \mu_i$  in  $\mathcal{O}_1^p$  (i.e.,  $\mu_{n,i} \rightarrow \mu_i$  in  $\mathcal{O}_1^0 \cap \mathcal{M}_1^p$  and  $\int_{\mathbb{R}} |x_i|^p \mu_{n,i}(dx_i) \rightarrow \int_{\mathbb{R}} |x_i|^p \mu_i(dx_i)$ ), Krättschmer et al. [28, Theorem 2.3 (5. $\Rightarrow$ 3.)] ensures that we can choose  $a_{1,i,i} \in \mathbb{R}_{++}$  so large so that this expression (with  $a_1 = a_{1,i,i}$ ) is bounded above by  $\varepsilon/(3dc_p)$  uniformly in  $n \in \mathbb{N}$ . For  $i \neq j$  and any  $b \in \mathbb{R}_{++}$ , the summand is bounded above by

$$\begin{aligned} & \int_{\mathbb{R}^d} |\pi_i(x)|^p \mathbb{1}_{[0,b]}(|\pi_i(x)|) \mathbb{1}_{(a_1^{1/p}/c_\infty, \infty)}(|x_j|) \mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d})(dx) \\ & + \int_{\mathbb{R}^d} |\pi_i(x)|^p \mathbb{1}_{(b, \infty)}(|\pi_i(x)|) \mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d})(dx) \\ & =: S_{1,i,j}(n, a_1, b) + S_{1,i}(n, b). \end{aligned}$$

Again by [28, Theorem 2.3] and the assumed convergence  $\mu_{n,i} \rightarrow \mu_i$  in  $\mathcal{O}_1^p$ , we can choose  $b_i$  so large that  $\sup_{n \in \mathbb{N}} S_{1,i}(n, b_i) \leq \varepsilon/(6d^2c_p)$ . Once we have chosen  $b_i$ , we can in view of  $S_{1,i,j}(n, a_1, b) \leq b^p \mu_{n,j}[[ -a_1^{1/p}/c_\infty, a_1^{1/p}/c_\infty ]^c]$  choose  $a_{1,i,j} \in \mathbb{R}_{++}$  so large that  $\sup_{n \in \mathbb{N}} S_{1,1}(n, a_{1,i,j}, b) \leq \varepsilon/(6d(d-1)c_p)$ ; take into account that  $(\mu_{n,j})_{n \in \mathbb{N}}$  as a weakly convergent sequence is tight. That is, we have  $\sup_{n \in \mathbb{N}} S_1(n, a_1) \leq \varepsilon/3$  when we set  $a_1 := \max_{i,j=1, \dots, d} a_{1,i,j}$ . Finally, by the already established convergence  $\mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d}) \rightarrow \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)$  in  $\mathcal{O}_d^0 \cap \mathcal{M}_d^p$ , we can choose  $n_0 \in \mathbb{N}$  such that  $S_2(n, a) \leq \varepsilon/(3c_p)$  for  $a := \max\{a_1, a_3\}$  and all  $n \geq n_0$ . Altogether, we have shown that for any given  $\varepsilon > 0$ , we can find an  $n_0 \in \mathbb{N}$  such that the left-hand side of (A.1) is  $\leq \varepsilon$  for all  $n \geq n_0$ .

Conversely, let  $(C, \mu_1, \dots, \mu_d)$  and  $(C_n, \mu_{n,1}, \dots, \mu_{n,d})$ ,  $n \in \mathbb{N}$ , be elements of  $\mathbf{C}_d \times \mathcal{M}_1^p \times \dots \times \mathcal{M}_1^p$  such that  $\mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d}) \rightarrow \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)$  in  $\mathcal{O}_d^p$ . Since  $p$ -weak convergence implies weak convergence, it is a direct consequence of Lindner and Szymayer [31, Theorem 2.1] that this implies  $d_{\mu_1, \dots, \mu_n}(C_n, C) \rightarrow 0$  and  $\mu_{n,i} \rightarrow \mu_i$  in  $\mathcal{O}_1^0$ ,  $i = 1, \dots, d$ . Moreover, for any  $i = 1, \dots, d$  we have

$$\begin{aligned} & \lim_{n \rightarrow \infty} \left| \int_{\mathbb{R}} |x_i|^p \mu_{n,i}(dx_i) - \int_{\mathbb{R}} |x_i|^p \mu_i(dx_i) \right| \\ & = \lim_{n \rightarrow \infty} \left| \int_{\mathbb{R}^d} |\pi_i(x)|^p \mathfrak{P}_d(C_n, \mu_{n,1}, \dots, \mu_{n,d})(dx) \right. \\ & \quad \left. - \int_{\mathbb{R}^d} |\pi_i(x)|^p \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)(dx) \right| = 0 \end{aligned}$$

since the function  $|\pi_i|^p : \mathbb{R}^d \rightarrow \mathbb{R}$  lies in  $\mathcal{C}_d^p$ . Thus we even have  $\mu_{n,i} \rightarrow \mu_i$  in  $\mathcal{O}_1^p$ ,  $i = 1, \dots, d$ .  $\square$

### A.3 Proof of Corollary 2.6

Of course, it suffices to show that  $\mathcal{O}_d^p \cap \mathcal{M}_d(\mu_1, \dots, \mu_d) = \mathcal{O}_d^0 \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$ . Recall that the  $p$ -weak topology is metrisable. Therefore it suffices to show that

for any sequence  $(\mu_n)_{n \in \mathbb{N}} \in \mathcal{M}_d(\mu_1, \dots, \mu_d)^{\mathbb{N}}$  and any  $\mu \in \mathcal{M}_d(\mu_1, \dots, \mu_d)$ , it holds that  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^p \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$  if and only if  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^0 \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$ . If  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^p \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$ , then  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^0 \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$  because  $\mathcal{O}_d^p \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$  is finer than  $\mathcal{O}_d^0 \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$ . Conversely, if  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^0 \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$ , we obtain by Theorem 2.3 (with  $p = 0$ ) that  $(C_n)_{n \in \mathbb{N}}$  converges to  $C$  in  $\mathcal{O}_{\mu_1, \dots, \mu_d}$ , where  $C_n$  and  $C$  are (arbitrary) copulas of  $\mu_n$  and  $\mu$ , respectively. Again with Theorem 2.3, we conclude that  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^p \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$ .  $\square$

#### A.4 Proof of Corollary 2.8

Since the involved topologies are both metrisable, it suffices to show that the map  $\mathfrak{C}_{\mu_1, \dots, \mu_d} : \mathcal{M}_d(\mu_1, \dots, \mu_d) \rightarrow \mathbf{C}_d / \sim_{\mu_1, \dots, \mu_d}$  is sequentially continuous for the pair  $(\mathcal{O}_d^{p'} \cap \mathcal{M}_d(\mu_1, \dots, \mu_d), \mathcal{O}_{\mu_1, \dots, \mu_d}^{\sim})$  for any  $p' \in [0, p]$ . Let  $\mu$  and  $\mu_n$ ,  $n \in \mathbb{N}$ , be elements of  $\mathcal{M}_d(\mu_1, \dots, \mu_d)$  such that  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^{p'} \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$  for some  $p' \in [0, p]$ . By Corollary 2.6, we obtain that  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^p \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$ , and by Theorem 2.3, it follows that  $\lim_{n \rightarrow \infty} d_{\mu_1, \dots, \mu_d}(C_n, C) = 0$  for any copulas  $C_n$  and  $C$  of  $\mu_n$  and  $\mu$ , respectively. So we arrive at

$$\lim_{n \rightarrow \infty} d_{\mu_1, \dots, \mu_d}^{\sim}(\mathfrak{C}_{\mu_1, \dots, \mu_d}(\mu_n), \mathfrak{C}_{\mu_1, \dots, \mu_d}(\mu)) = 0,$$

i.e.,  $\mathfrak{C}_{\mu_1, \dots, \mu_d}(\mu_n) \rightarrow \mathfrak{C}_{\mu_1, \dots, \mu_d}(\mu)$  in  $\mathcal{O}_{\mu_1, \dots, \mu_d}^{\sim}$ .  $\square$

#### A.5 Proof of Example 3.2

We here show that if  $\mathcal{M}'_d = \mathcal{N}_d$ , then  $\mathfrak{D}'_d = \mathbf{C}_d^{\text{Ga}} \times \mathcal{N}_1 \times \dots \times \mathcal{N}_1$ .

“ $\supseteq$ ” Let  $(C, \mu_1, \dots, \mu_d) \in \mathbf{C}_d^{\text{Ga}} \times \mathcal{N}_1 \times \dots \times \mathcal{N}_1$ , which means that we have  $\mu_1 = N_{m_1, s_1^2}, \dots, \mu_d = N_{m_d, s_d^2}$  for some  $m_1, \dots, m_d \in \mathbb{R}$  and  $s_1, \dots, s_d \in \mathbb{R}_{++}$ , and  $C$  is given by (3.2) for some correlation matrix  $R$ . Recall that quantiles are translation-equivariant and positively homogeneous (on the level of univariate random variables). For the distribution function of the Borel probability measure  $\mu := \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)$ , we thus obtain

$$\begin{aligned} F_{\mu}(x_1, \dots, x_d) &= C(F_{\mu_1}(x_1), \dots, F_{\mu_d}(x_d)) \\ &= \Phi_{\mathbf{0}, R} \left( \Phi_{0,1}^{-1}(\Phi_{m_1, s_1^2}(x_1)), \dots, \Phi_{0,1}^{-1}(\Phi_{m_d, s_d^2}(x_d)) \right) \\ &= \Phi_{\mathbf{0}, R} \left( \frac{\Phi_{m_1, s_1^2}^{-1}(\Phi_{m_1, s_1^2}(x_1)) - m_1}{s_1}, \dots, \frac{\Phi_{m_d, s_d^2}^{-1}(\Phi_{m_d, s_d^2}(x_d)) - m_d}{s_d} \right) \\ &= \Phi_{\mathbf{0}, R} \left( \frac{x_1 - m_1}{s_1}, \dots, \frac{x_d - m_d}{s_d} \right) = \Phi_{m, SRS}(x_1, \dots, x_d), \end{aligned}$$



where  $m := (m_1, \dots, m_d)^\top$  and  $S$  is the  $d \times d$  diagonal matrix with entries  $s_1, \dots, s_d$  on the diagonal. Note that the matrix  $SRS$  is again symmetric and positive semi-definite, implying that  $\mu$  is the  $d$ -variate normal distribution  $N_{m, SRS}$ . Since the entries on the diagonal of the matrix  $SRS$  are the strictly positive numbers  $s_1^2, \dots, s_d^2$ , the marginal distributions of  $\mu = N_{m, SRS}$  are  $N_{m_1, s_1^2}, \dots, N_{m_d, s_d^2}$ , i.e., elements of  $\mathcal{N}_1$ . In particular,  $\mu$  is a (possibly degenerate)  $d$ -variate normal distribution with continuous marginals, i.e.,  $\mu = \mathfrak{P}_d(C, \mu_1, \dots, \mu_d)$  lies in  $\mathcal{M}'_d = \mathcal{N}_d$ . Thus we obtain that  $(C, \mu_1, \dots, \mu_d) \in \mathfrak{D}'_d$ .

“ $\subseteq$ ” Let  $(C, \mu_1, \dots, \mu_d) \in \mathfrak{D}'_d$ , i.e.,  $\mu_1, \dots, \mu_d \in \mathcal{M}_1$  and  $C \in \mathbf{C}_d$  are such that  $\mu := \mathfrak{P}_d(C, \mu_1, \dots, \mu_d) \in \mathcal{M}'_d = \mathcal{N}_d$ . Then  $\mu_1 = N_{m_1, s_1^2}, \dots, \mu_d = N_{m_d, s_d^2}$  for some  $m_1, \dots, m_d \in \mathbb{R}$  and  $s_1, \dots, s_d \in \mathbb{R}_{++}$ , and  $\mu = N_{m, V}$  for the vector  $m := (m_1, \dots, m_d)^\top$  and a symmetric and positive semi-definite matrix  $V \in \mathbb{R}^{d \times d}$  with entries  $s_1^2, \dots, s_d^2$  on the diagonal. By (2.1) and the translation-equivariance and positive homogeneity of quantiles (on the level of univariate random variables), we have

$$\begin{aligned} C(u_1, \dots, u_d) &= \Phi_{m, V}(\Phi_{m_1, s_1^2}^{-1}(u_1), \dots, \Phi_{m_d, s_d^2}^{-1}(u_d)) \\ &= \Phi_{m, V}(s_1 \Phi_{0,1}^{-1}(u_1) + m_1, \dots, s_d \Phi_{0,1}^{-1}(u_d) + m_d) \\ &= \Phi_{0, S^{-1}VS^{-1}}(\Phi_{0,1}^{-1}(u_1), \dots, \Phi_{0,1}^{-1}(u_d)), \end{aligned}$$

where  $S^{-1}$  is the  $d \times d$  diagonal matrix with entries  $s_1^{-1}, \dots, s_d^{-1}$  on the diagonal. The matrix  $R := S^{-1}VS^{-1}$  is again symmetric and positive semi-definite, and for any  $i, j \in \{1, \dots, d\}$ , its entry at  $(i, j)$  is equal to  $v_{i,j}/(s_i s_j)$ , where  $v_{i,j}$  is the entry at  $(i, j)$  of the matrix  $V$ . Since  $v_{i,i} = s_i^2$  for any  $i = 1, \dots, d$ , it follows that  $R$  is a correlation matrix. Thus  $C \in \mathbf{C}_d^{\text{Ga}}$  and therefore  $(C, \mu_1, \dots, \mu_d) \in \mathbf{C}_d^{\text{Ga}} \times \mathcal{N}_1 \times \dots \times \mathcal{N}_1$ .  $\square$

## A.6 Proof of Example 3.3

We here show that if  $\mathcal{M}'_d = \mathcal{M}_d^p$  for some  $p \in \mathbb{R}_+$ , then  $\mathfrak{D}'_d = \mathbf{C}_d \times \mathcal{M}_1^p \times \dots \times \mathcal{M}_1^p$ .

“ $\supseteq$ ” Let  $(C, \mu_1, \dots, \mu_d) \in \mathbf{C}_d \times \mathcal{M}_1^p \times \dots \times \mathcal{M}_1^p$ . Then, as shown in the first paragraph of Sect. A.2,  $\mathfrak{P}_d(C, \mu_1, \dots, \mu_d) \in \mathcal{M}_d^p = \mathcal{M}'_d$ . Thus  $(C, \mu_1, \dots, \mu_d)$  lies in  $\mathfrak{D}'_d$ .

“ $\subseteq$ ” Let  $(C, \mu_1, \dots, \mu_d) \in \mathfrak{D}'_d$ , i.e.,  $\mu_1, \dots, \mu_d \in \mathcal{M}_1$  and  $C \in \mathbf{C}_d$  are such that  $\mathfrak{P}_d(C, \mu_1, \dots, \mu_d) \in \mathcal{M}_d^p = \mathcal{M}'_d$ . Then, as shown in the first paragraph of Sect. A.2,  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ . Thus  $(C, \mu_1, \dots, \mu_d) \in \mathbf{C}_d \times \mathcal{M}_1^p \times \dots \times \mathcal{M}_1^p$ .  $\square$

## A.7 Proof of Remark 3.6

The univariate standard normal distribution function  $\Phi_{0,1}$  is continuous and strictly increasing. This implies that pointwise convergence of a sequence of  $d$ -variate Gauss-

ian copulas  $(\Phi_{\mathbf{0}, R_n}(\Phi_{0,1}^{-1}(\cdot), \dots, \Phi_{0,1}^{-1}(\cdot)))_{n \in \mathbb{N}}$  to a  $d$ -variate Gaussian copula

$$\Phi_{\mathbf{0}, R}(\Phi_{0,1}^{-1}(\cdot), \dots, \Phi_{0,1}^{-1}(\cdot))$$

is the same as pointwise convergence of  $\Phi_{\mathbf{0}, R_n}(\cdot, \dots, \cdot)$  to  $\Phi_{\mathbf{0}, R}(\cdot, \dots, \cdot)$ . The  $d$ -variate distribution function  $\Phi_{\mathbf{0}, R}(\cdot, \dots, \cdot)$  is continuous since it can be represented as

$$\Phi_{\mathbf{0}, R}(\cdot, \dots, \cdot) = \Phi_{\mathbf{0}, R}(\Phi_{0,1}^{-1}(\Phi_{0,1}(\cdot)), \dots, \Phi_{0,1}^{-1}(\Phi_{0,1}(\cdot)))$$

and  $\Phi_{\mathbf{0}, R}(\Phi_{0,1}^{-1}(\cdot), \dots, \Phi_{0,1}^{-1}(\cdot))$  as a (Gaussian) copula is Lipschitz-continuous with respect to  $\|\cdot\|_1$  and the univariate standard normal distribution function  $\Phi_{0,1}(\cdot)$  is continuous. Therefore (see Shiryaev [48, Sect. III.1]) the latter pointwise convergence is equivalent to weak convergence of the  $d$ -variate normal distribution  $\mathbf{N}_{\mathbf{0}, R_n}$  to the  $d$ -variate normal distribution  $\mathbf{N}_{\mathbf{0}, R}$ , and by Lévy's continuity theorem, this is the same as

$$e^{-\langle R_n t, t \rangle / 2} = \varphi_{\mathbf{N}_{\mathbf{0}, R_n}}(t) \longrightarrow \varphi_{\mathbf{N}_{\mathbf{0}, R}}(t) = e^{-\langle R t, t \rangle / 2}$$

for any  $t \in \mathbb{R}^d$ . Obviously, the latter holds if and only if  $t^\top (R_n - R)t \rightarrow 0$  for any  $t \in \mathbb{R}^d$ .

If  $\|R_n - R\|_{\text{Mat}} \rightarrow 0$  for some matrix norm  $\|\cdot\|_{\text{Mat}}$ , then  $\|R_n - R\|_{\text{max}} \rightarrow 0$  for the maximum norm  $\|\cdot\|_{\text{max}}$ , and thus  $t^\top (R_n - R)t \rightarrow 0$  for any  $t \in \mathbb{R}^d$ .

Conversely, assume that  $t^\top (R_n - R)t \rightarrow 0$  for any  $t \in \mathbb{R}^d$ . Then for any  $i = 1, \dots, d$ , we can conclude by choosing  $t$  as the  $i$ th unit vector  $e_i$  that the entry  $(i, i)$  of the matrix  $R_n - R$  converges to 0. For  $t := e_i + e_j$ , the expression  $t^\top (R_n - R)t$  is twice the entry  $(i, j)$  plus entries  $(i, i)$  and  $(j, j)$  of the symmetric matrix  $R_n - R$ . It follows that also the entry  $(i, j)$  of the matrix  $R_n - R$  converges to 0 for any  $i, j = 1, \dots, d$  with  $i \neq j$ . Thus  $\|R_n - R\|_{\text{max}} \rightarrow 0$ , and consequently  $\|R_n - R\|_{\text{Mat}} \rightarrow 0$  for any matrix norm  $\|\cdot\|_{\text{Mat}}$ .  $\square$

## A.8 Proof of Example 3.7

For any correlation matrix  $R$ , use  $C^R$  to denote the Gaussian copula associated with  $R$ . By definition, the set  $\mathbf{C}_d^{\text{Ga}}$  is parametrised by the set of all correlation matrices. Now fix  $N_{m_1, s_1^2}, \dots, N_{m_d, s_d^2} \in \mathcal{N}_1$  and let  $R, R_1, R_2, \dots$  be correlation matrices such that  $C^{R_n}$  converges to  $C^R$  pointwise, i.e.,  $C^{R_n}(u_1, \dots, u_d) \rightarrow C^R(u_1, \dots, u_d)$  for any  $u_1, \dots, u_d \in [0, 1]$ . In particular,

$$C^{R_n}(\Phi_{m_1, s_1^2}(x_1), \dots, \Phi_{m_d, s_d^2}(x_d)) \longrightarrow C^R(\Phi_{m_1, s_1^2}(x_1), \dots, \Phi_{m_d, s_d^2}(x_d)) \quad (\text{A.2})$$

for all  $x_1, \dots, x_d \in \mathbb{R}$ . Now  $F_n(x_1, \dots, x_d) := C^{R_n}(\Phi_{m_1, s_1^2}(x_1), \dots, \Phi_{m_d, s_d^2}(x_d))$  and  $F(x_1, \dots, x_d) := C^R(\Phi_{m_1, s_1^2}(x_1), \dots, \Phi_{m_d, s_d^2}(x_d))$  are the distribution functions of  $\mathfrak{P}_d(C^{R_n}, N_{m_1, s_1^2}, \dots, N_{m_d, s_d^2})$  and  $\mathfrak{P}_d(C^R, N_{m_1, s_1^2}, \dots, N_{m_d, s_d^2})$ , respectively. More-

over,  $F$  is continuous since  $C^R$  as a copula is Lipschitz-continuous with respect to  $|\cdot|_1$  and the univariate normal distribution functions  $\Phi_{m_1, s_1^2}(\cdot), \dots, \Phi_{m_d, s_d^2}(\cdot)$  are continuous. Therefore (see Shiryaev [48, Sect. III.1]) the convergence in (A.2) is equivalent to

$$\mathfrak{P}_d(C^{R_n}, N_{m_1, s_1^2}, \dots, N_{m_d, s_d^2}) \longrightarrow \mathfrak{P}_d(C^R, N_{m_1, s_1^2}, \dots, N_{m_d, s_d^2}) \quad \text{in } \mathcal{O}_d^0.$$

Moreover, as in (A.1) (with  $\mu_{n,i} = \mu_i = N_{m_i, s_i^2}$ ,  $n \in \mathbb{N}$ ,  $i = 1, \dots, d$ ), we obtain

$$\begin{aligned} & \int_{\mathbb{R}^d} |x|^p \mathfrak{P}_d(C^{R_n}, N_{m_1, s_1^2}, \dots, N_{m_d, s_d^2})(dx) \\ & \longrightarrow \int_{\mathbb{R}^d} |x|^p \mathfrak{P}_d(C^R, N_{m_1, s_1^2}, \dots, N_{m_d, s_d^2})(dx). \end{aligned}$$

Thus  $\mathfrak{P}_d(C^{R_n}, N_{m_1, s_1^2}, \dots, N_{m_d, s_d^2}) \rightarrow \mathfrak{P}_d(C^R, N_{m_1, s_1^2}, \dots, N_{m_d, s_d^2})$  in  $\mathcal{O}_p^d \cap \mathcal{N}_d$ . In view of  $\mathcal{P}_d \circ \mathfrak{P}_d = \mathfrak{P}_d$  and since the involved topologies are metrisable, we arrive at copula robustness of  $\mathcal{P}_d : \mathcal{N}_d \rightarrow \mathcal{N}_d$ , where the image space  $\mathcal{N}_d$  is equipped with  $\mathcal{O}_p^d \cap \mathcal{N}_d$ .  $\square$

## A.9 Proof of Theorem 3.10

Let us first assume that the map  $\mathcal{T}_d|_{\mathcal{M}_d(\mu_1, \dots, \mu_d)} : \mathcal{M}_d(\mu_1, \dots, \mu_d) \rightarrow \mathbf{E}$  is continuous for the pair  $(\mathcal{O}_d^0 \cap \mathcal{M}_d(\mu_1, \dots, \mu_d), \mathcal{O}_{\mathbf{E}})$  for any  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ . Since the map

$$\mathfrak{P}_d(\cdot, \mu_1, \dots, \mu_d) : \mathbf{C}_d \rightarrow \mathcal{M}_d(\mu_1, \dots, \mu_d)$$

is  $(\mathcal{O}_{\mu_1, \dots, \mu_d}, \mathcal{O}_d^0 \cap \mathcal{M}_d(\mu_1, \dots, \mu_d))$ -continuous for any  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$  by Corollary 2.5, it follows that the map

$$\mathfrak{T}_d(\cdot, \mu_1, \dots, \mu_d) = \mathcal{T}_d \circ \mathfrak{P}_d(\cdot, \mu_1, \dots, \mu_d) : \mathbf{C}_d \rightarrow \mathbf{E}$$

is  $(\mathcal{O}_{\mu_1, \dots, \mu_d}, \mathcal{O}_{\mathbf{E}})$ -continuous for any  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ . Thus  $\mathcal{T}_d$  is copula robust.

Conversely, assume  $\mathcal{T}_d$  is copula robust, i.e.,  $\mathfrak{T}_d(\cdot, \mu_1, \dots, \mu_d) : \mathbf{C}_d \rightarrow \mathbf{E}$  is  $(\mathcal{O}_{\mu_1, \dots, \mu_d}, \mathcal{O}_{\mathbf{E}})$ -continuous for any  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ . By way of contradiction, assume that  $\mathcal{T}_d|_{\mathcal{M}_d(\mu_1, \dots, \mu_d)} : \mathcal{M}_d(\mu_1, \dots, \mu_d) \rightarrow \mathbf{E}$  is not continuous for the pair  $(\mathcal{O}_d^0 \cap \mathcal{M}_d(\mu_1, \dots, \mu_d), \mathcal{O}_{\mathbf{E}})$ , for some  $\mu_1, \dots, \mu_d \in \mathcal{M}_1^p$ . Then one can find elements  $\mu_n$ ,  $n \in \mathbb{N}$ , and  $\mu$  of  $\mathcal{M}_d(\mu_1, \dots, \mu_d)^{\mathbb{N}}$  such that we have  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^0 \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$ , but  $\mathcal{T}_d(\mu_n) \not\rightarrow \mathcal{T}_d(\mu)$  in  $\mathcal{O}_{\mathbf{E}}$ . However, in view of Corollary 2.8, the convergence  $\mu_n \rightarrow \mu$  in  $\mathcal{O}_d^0 \cap \mathcal{M}_d(\mu_1, \dots, \mu_d)$  implies that  $C_n \rightarrow C$  in  $\mathcal{O}_{\mu_1, \dots, \mu_d}$  for any copulas  $C_n$  and  $C$  of  $\mu_n$  and  $\mu$ , respectively. Because the map  $\mathfrak{T}_d(\cdot, \mu_1, \dots, \mu_d) : \mathbf{C}_d \rightarrow \mathbf{E}$  is  $(\mathcal{O}_{\mu_1, \dots, \mu_d}, \mathcal{O}_{\mathbf{E}})$ -continuous by assumption, we obtain that  $\mathcal{T}_d(\mu_n) \rightarrow \mathcal{T}_d(\mu)$  in  $\mathcal{O}_{\mathbf{E}}$ . This contradicts  $\mathcal{T}_d(\mu_n) \not\rightarrow \mathcal{T}_d(\mu)$  in  $\mathcal{O}_{\mathbf{E}}$ .  $\square$

### A.10 Proof of Corollary 3.13

The map  $\mathfrak{h} : \mathcal{M}_d^p \rightarrow \mathcal{M}_{d'}^{p'}$  defined by  $\mathfrak{h}(\mu) := \mu \circ h^{-1}$  is  $(\mathcal{O}_d^p, \mathcal{O}_{d'}^{p'})$ -continuous by Proposition 2.2 and therefore copula robust by Theorem 3.12. Because the map  $\mathcal{T}_{d'} : \mathcal{M}_{d'}^{p'} \rightarrow \mathbf{E}$  is  $(\mathcal{O}_{d'}^{p'}, \mathcal{O}_d^p)$ -continuous by assumption, it follows by Lemma 3.4 (with  $\mathcal{U} := \mathcal{T}_{d'}$  and  $\mathcal{T}_d := \mathfrak{h}$ ) that  $\mathcal{T}_d' = \mathcal{T}_{d'} \circ \mathfrak{h} : \mathcal{M}_d^p \rightarrow \mathbf{E}$  is copula robust.  $\square$

### A.11 Proof of Example 4.7

(i) The natural extension  $\overline{C}_0^{(\alpha)}$  of  $C_0^{(\alpha)}$  to  $\mathbb{R}^2$  is the distribution function of the Borel probability measure  $\mathfrak{P}_2(C_0^{(\alpha)}, \mu_1, \mu_2)$  on  $\mathbb{R}^2$ . Thus  $\mathfrak{P}_2(C_0^{(\alpha)}, \mu_1, \mu_2) = \mathcal{H}_{S_1^\alpha \uplus S_2^\alpha}^1 / \sqrt{2}$ , where  $\mathcal{H}_{S_1^\alpha \uplus S_2^\alpha}^1[\cdot] := \mathcal{H}^1[\cdot \cap (S_1^\alpha \uplus S_2^\alpha)]$  is the 1-dimensional (Borel) Hausdorff measure  $\mathcal{H}^1$  on  $\mathbb{R}^2$  restricted to the union of the two disjoint line segments  $S_1^\alpha$  and  $S_2^\alpha$  with endpoints  $(\alpha, 0)$ ,  $(0, \alpha)$  and  $(1, \alpha)$ ,  $(\alpha, 1)$ , respectively. That is, the total mass 1 of  $\mathfrak{P}_2(C_0^{(\alpha)}, \mu_1, \mu_2)$  is uniformly distributed over  $S_1^\alpha \uplus S_2^\alpha$ . In particular,  $\mathcal{H}^1[S_1^\alpha] = \alpha$  and  $\mathcal{H}^1[S_2^\alpha] = 1 - \alpha$ . In view of  $S_1^\alpha = \{(x_1, x_2) \in [0, 1]^2 : x_1 + x_2 = \alpha\}$  and  $S_2^\alpha = \{(x_1, x_2) \in [0, 1]^2 : x_1 + x_2 = 1 + \alpha\}$ , we can conclude that

$$\mathfrak{P}_2(C_0^{(\alpha)}, \mu_1, \mu_2) \circ A_2^{-1} = \alpha \delta_\alpha + (1 - \alpha) \delta_{1+\alpha}.$$

In particular,  $\mathfrak{R}_{\text{VaR}_{\alpha}, A_2}(C_0^{(\alpha)}, \mu_1, \mu_2) = \alpha$ .

The natural extension  $\overline{C}_1$  of  $C_1$  to  $\mathbb{R}^2$  is the distribution function of the Borel probability measure  $\mathfrak{P}_2(C_1, \mu_1, \mu_2)$  on  $\mathbb{R}^2$ . Thus  $\mathfrak{P}_2(C_1, \mu_1, \mu_2) = \mathcal{L}_{[0,1]^2}^{(2)}$ , where  $\mathcal{L}_{[0,1]^2}^{(2)}[\cdot] := \mathcal{L}^{(2)}[\cdot \cap [0, 1]^2]$  is the (Borel) Lebesgue measure  $\mathcal{L}^{(2)}$  on  $\mathbb{R}^2$  restricted to  $[0, 1]^2$ . Then

$$\mathfrak{P}_2(C_1, \mu_1, \mu_2) \circ A_2^{-1} = (\mathcal{L}_{[0,1]}^{(1)} \otimes \mathcal{L}_{[0,1]}^{(1)}) \circ A_2^{-1} = \mathcal{L}_{[0,1]}^{(1)} * \mathcal{L}_{[0,1]}^{(1)} = \Delta_{[0,2]},$$

where  $\mathcal{L}_{[0,1]}^{(1)}[\cdot] := \mathcal{L}^{(1)}[\cdot \cap [0, 1]]$  is the (Borel) Lebesgue measure  $\mathcal{L}^{(1)}$  on  $\mathbb{R}$  restricted to  $[0, 1]$  and  $\Delta_{[0,2]}$  is the symmetric triangular distribution. Therefore

$$\begin{aligned} & \mathfrak{P}_2(C_t^{(\alpha)}, \mu_1, \mu_2) \circ A_2^{-1} \\ &= (1 - t) \mathfrak{P}_2(C_0^{(\alpha)}, \mu_1, \mu_2) \circ A_2^{-1} + t \mathfrak{P}_2(C_1, \mu_1, \mu_2) \circ A_2^{-1} \\ &= (1 - t) (\alpha \delta_\alpha + (1 - \alpha) \delta_{1+\alpha}) + t \Delta_{[0,2]} \end{aligned}$$

for any  $t \in (0, 1]$ . For the distribution function  $F_{\Delta_2}$  of  $\Delta_{[0,2]}$ , we have  $F_{\Delta_2}(x) = 0, = \frac{1}{2}x^2, = 1 - \frac{1}{2}(2 - x)^2, = 1$  according to whether  $x < 0$ ,  $x \in [0, 1]$ ,  $x \in (1, 2]$ ,  $x > 2$ . For the distribution function  $F^{(\alpha)}$  of  $\alpha \delta_\alpha + (1 - \alpha) \delta_{1+\alpha}$ , we have  $F^{(\alpha)}(x) = 0, = \alpha, = 1$  according to whether  $x < \alpha$ ,  $x \in [\alpha, 1 + \alpha]$ ,  $x \geq 1 + \alpha$ . Thus

for any  $t \in (0, 1]$ , the distribution function  $F_t^{(\alpha)}$  of  $\mathfrak{P}_2(C_t^{(\alpha)}, \mu_1, \mu_2) \circ A_2^{-1}$  is given by

$$F_t^{(\alpha)}(x) = \begin{cases} 0, & x \in (-\infty, 0), \\ t\frac{1}{2}x^2, & x \in [0, \alpha), \\ t\frac{1}{2}x^2 + (1-t)\alpha, & x \in [\alpha, 1), \\ t(1 - \frac{1}{2}(2-x)^2) + (1-t)\alpha, & x \in [1, 1+\alpha), \\ t(1 - \frac{1}{2}(2-x)^2) + (1-t), & x \in [1+\alpha, 2), \\ 1, & x \in [2, \infty), \end{cases}$$

so that  $\mathfrak{R}_{\text{VaR}_{\alpha}, A_2}(C_t^{(\alpha)}, \mu_1, \mu_2) = \sqrt{2\alpha}$ .

(ii) For any  $t \in [0, 1]$ , let  $(X_1^{(\alpha),t}, X_2^{(\alpha),t})$  be a bivariate random variable with distribution  $\mathfrak{P}_2(C_t^{(\alpha)}, \mu_1, \mu_2)$ . The distribution of the random variable  $-(X_1^{(\alpha),0} + X_2^{(\alpha),0})$  is then  $\alpha\delta_{-\alpha} + (1-\alpha)\delta_{-(1+\alpha)}$ . Thus  $\text{VaR}_{1-\alpha}(-(X_1^{(\alpha),0} + X_2^{(\alpha),0})) = -(1+\alpha)$ . For any  $t \in (0, 1]$ , the distribution function  $\hat{F}_t^{(\alpha)}$  of  $-(X_1^{(\alpha),t} + X_2^{(\alpha),t})$  satisfies  $\hat{F}_t^{(\alpha)}(x) = 1 - F_t^{(\alpha)}(-x)$  for all those  $x \in \mathbb{R}$  for which  $-x$  is a continuity point of  $F_t^{(\alpha)}$ . Thus for any  $t \in (0, 1]$ , we get  $\text{VaR}_{1-\alpha}(-(X_1^{(\alpha),t} + X_2^{(\alpha),t})) = -\sqrt{2\alpha}$ . Hence,

$$\begin{aligned} \mathfrak{R}_{\text{VaR}_{1-\alpha}, A_2}(C_0^{(\alpha),-}, \hat{\mu}_1, \hat{\mu}_2) &= -(1+\alpha), \\ \mathfrak{R}_{\text{VaR}_{1-\alpha}, A_2}(C_t^{(\alpha),-}, \hat{\mu}_1, \hat{\mu}_2) &= -\sqrt{2\alpha} \quad \text{for any } t \in (0, 1], \end{aligned}$$

where  $C_t^{(\alpha),-}$  denotes the copula of the distribution of  $(-X_1^{(\alpha),t}, -X_2^{(\alpha),t})$ ; here we use that  $\hat{\mu}_j$  is the distribution of  $-X_j^{(\alpha),t}$ ,  $j = 1, 2$ . This gives the assertion since  $C_t^{(\alpha),-} = \hat{C}_t^{(\alpha)}$  for any  $t \in [0, 1]$ . The latter equality holds true since the distribution function  $F^{(\alpha),t,-}$  of  $(-X_1^{(\alpha),t}, -X_2^{(\alpha),t})$  satisfies

$$\begin{aligned} F^{(\alpha),t,-}(x_1, x_2) &= \overline{F}^{(\alpha),t}(-x_1, -x_2) = \hat{C}_t^{(\alpha)}(1 - F_{\mu_1}(-x_1), 1 - F_{\mu_2}(-x_2)) \\ &= \hat{C}_t^{(\alpha)}(F_{\hat{\mu}_1}(x_1), F_{\hat{\mu}_2}(x_2)) \end{aligned}$$

for any  $t \in [0, 1]$ , where  $\overline{F}^{(\alpha),t}$  denotes the survival function of  $(X_1^{(\alpha),t}, X_2^{(\alpha),t})$ .  $\square$

## A.12 Proof of (4.4)

Let  $p \in [1, \infty)$  and  $\rho : L^p \rightarrow \mathbb{R}$  be a comonotonic convex risk measure. Moreover, let  $\mu \in \mathcal{M}_d^p$ . Since  $(\Omega, \mathcal{F}, \mathbb{P})$  is assumed to be atomless, we can choose  $(X_1, \dots, X_d) \in L^p \times \dots \times L^p$  such that  $\mathbb{P} \circ (X_1, \dots, X_d)^{-1} = \mu$ . A result of Filipović and Svindland [21, Corollary 2.7] ensures that there exist a comonotone opti-

mal capital and risk allocation  $(X_1^*, \dots, X_d^*)$  of  $X := \sum_{i=1}^d X_i$  ( $\in L^p$ ). Thus

$$\begin{aligned} \mathcal{R}_{\boxtimes_{i=1}^d \rho}(\mu) &= \mathcal{R}_{\boxtimes_{i=1}^d \rho}(\mathbb{P} \circ (X_1, \dots, X_d)^{-1}) \\ &= \boxtimes_{i=1}^d \rho(X_1, \dots, X_d) = \boxdot_{i=1}^d \rho\left(\sum_{i=1}^d X_i\right) = \sum_{i=1}^d \rho(X_i^*) \\ &= \rho\left(\sum_{i=1}^d X_i^*\right) = \rho\left(\sum_{i=1}^d X_i\right) \\ &= \mathcal{R}_\rho\left(\mathbb{P} \circ \left(\sum_{i=1}^d X_i\right)^{-1}\right) = \mathcal{R}_\rho(\mu \circ A_d^{-1}), \end{aligned}$$

where the fifth equality relies on the comonotonicity of  $\rho$ .  $\square$

### A.13 Proof of Theorem 5.2

We can use arguments similar to those used in the proof of Claus et al. [9, Corollary 2.4]. In view of Lemma 5.1 and the compactness of  $\mathcal{E}$ , we can apply Bonnans and Shapiro [6, Proposition 4.4] to obtain that the map  $\mathcal{R}_{\rho, h} : \mathcal{M}_d^{\gamma p} \rightarrow \mathbb{R}$  is continuous with respect to  $(\mathcal{O}_d^{\gamma p}, \mathcal{O}_{\mathbb{R}})$ . Berge [4, Theorem VI.3.2] ensures that the infimum in (5.3) is attained for any  $\mu \in \mathcal{M}_d^{\gamma p}$ .  $\square$

### A.14 Proof of Corollary 5.4

Conditions (a)–(c) of Sect. 5.1 hold true for  $p = 1$ . Indeed, condition (a) holds true since monotonicity, distribution-invariance and convexity carry over from  $\sigma$  to  $\rho$ . Condition (b) with  $p = 1$  clearly holds true for the function  $h : \mathcal{E} \times \mathbb{R}^d \rightarrow \mathbb{R}$  defined by (5.5), and condition (c) holds true since  $h$  is continuous everywhere. Thus, since the set  $\mathcal{E}$  defined by (5.4) is a compact subset of  $\mathbb{R}^d$ , the assertions follow from Theorem 5.2.  $\square$

### A.15 Proof of Lemma 6.1

Since we assumed that conditions (a) and (c) hold true, we have that

$$\begin{aligned} &\mathbb{E}^{x_0, P; \pi} [ |r_k(X_k, f_k(X_k))| ] \\ &\leq \mathbb{E}^{x_0, P; \pi} [K_1 \psi(X_k)] \\ &= K_1 \int \cdots \int \int \psi(y_k) P_{k-1}((y_{k-1}, f_{k-1}(y_{k-1})), dy_k) \\ &\quad P_{k-2}((y_{k-2}, f_{k-2}(y_{k-2})), dy_{k-1}) \cdots P_0((x_0, f_0(x_0)), dy_1) \\ &\leq K_1 K_3^k \psi(x_0). \end{aligned}$$

Thus  $r_k(X_k, f_k(X_k))$ ,  $k = 0, \dots, N - 1$ , are  $\mathbb{P}^{x_0, P; \pi}$ -integrable for any  $x_0 \in E$ ,  $P \in \mathcal{P}'$  and  $\pi \in \Pi$ . Using (b) and (c), we analogously obtain that  $r_N(X_N)$  is  $\mathbb{P}^{x_0, P; \pi}$ -integrable for any  $x_0 \in E$ ,  $P \in \mathcal{P}'$  and  $\pi \in \Pi$ . For any  $n = 1, \dots, N - 1$ , we get in the same way that

$$\begin{aligned} |V_n^{P, \pi}(x)| &= \left| \mathbb{E}^{x_0, P; \pi} \left[ \sum_{k=n}^{N-1} r_k(X_k, f_k(X_k)) + r_N(X_N) \middle| X_n = x \right] \right| \\ &\leq \sum_{k=n}^{N-1} \mathbb{E}^{x_0, P; \pi} [|r_k(X_k, f_k(X_k))| | X_n = x] + \mathbb{E}^{x_0, P; \pi} [|r_N(X_N)| | X_n = x] \\ &\leq \sum_{k=n}^{N-1} K_1 K_3^{k-n} \psi(x) + K_2 K_3^{N-n} \psi(x) = K_n \psi(x) \end{aligned}$$

holds true for all  $\pi \in \Pi$  and  $x \in E$ , where  $K_n := \sum_{k=n}^{N-1} K_1 K_3^{k-n} + K_2 K_3^{N-n}$ . Therefore, we indeed have that

$$\|V_n^P\|_\psi = \sup_{x \in E} \sup_{\pi \in \Pi} |V_n^{P, \pi}(x)| / \psi(x) \leq \sup_{x \in E} \sup_{\pi \in \Pi} |V_n^{P, \pi}(x)| / \psi(x) < \infty$$

for any  $n = 0, \dots, N - 1$  and  $P \in \mathcal{P}$ . □

## A.16 Proof of Theorem 6.2

Let  $n \in \{0, \dots, N - 1\}$  and  $x_n \in E$ . Since  $\mathbb{P}^{x_0, Q; \pi} = \delta_{x_0} \otimes Q_0^\pi \otimes \dots \otimes Q_{N-1}^\pi$ , we have for any  $Q \in \mathcal{P}_\psi$  that

$$\begin{aligned} V_n^{Q; \pi}(x_n) &= \sum_{k=n}^{N-1} \mathbb{E}^{x_0, Q; \pi} [r_k(X_k, f_k(X_k)) | X_n = x_n] + \mathbb{E}^{x_0, Q; \pi} [r_N(X_N) | X_n = x_n] \\ &= r_n(x_n, f_n(x_n)) \\ &\quad + \sum_{k=n+1}^{N-1} \int_E \dots \int_E r_k(x_k, f_k(x_k)) Q_{k-1}^\pi(x_{k-1}, dy_k) \dots Q_n^\pi(x_n, dx_{n+1}) \\ &\quad + \int_E \dots \int_E r_N(x_N) Q_{N-1}^\pi(x_{N-1}, dx_N) \dots Q_n^\pi(x_n, dx_{n+1}) \end{aligned}$$

and thus

$$\begin{aligned}
& V_n^{Q;\pi}(x_n) - V_n^{P;\pi}(x_n) \\
&= \sum_{k=n+1}^{N-1} \int_E \cdots \int_E r_k(x_k, f_k(x_k)) Q_{k-1}^\pi(x_{k-1}, dx_k) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
&\quad - \sum_{k=n+1}^{N-1} \int_E \cdots \int_E r_k(x_k, f_k(x_k)) P_{k-1}^\pi(x_{k-1}, dx_k) \cdots P_n^\pi(x_n, dx_{n+1}) \\
&\quad + \int_E \cdots \int_E r_N(x_N) Q_{N-1}^\pi(x_{N-1}, dx_N) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
&\quad - \int_E \cdots \int_E r_N(x_N) P_{N-1}^\pi(x_{N-1}, dx_N) \cdots P_n^\pi(x_n, dx_{n+1}) \\
&= \sum_{k=n+1}^{N-1} \sum_{j=n}^{k-1} \int_E \cdots \int_E \int_E \int_E \cdots \int_E r_k(x_k, f_k(x_k)) P_{k-1}^\pi(x_{k-1}, dx_k) \cdots \\
&\quad P_{j+1}^\pi(x_{j+1}, dx_{j+2}) (Q_j^\pi - P_j^\pi)(x_j, dx_{j+1}) \\
&\quad Q_{j-1}^\pi(x_{j-1}, dx_j) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
&\quad + \sum_{j=n}^{N-1} \int_E \cdots \int_E \int_E \int_E \cdots \int_E r_N(x_N) P_{N-1}^\pi(x_{N-1}, dx_N) \cdots \\
&\quad P_{j+1}^\pi(x_{j+1}, dx_{j+2}) (Q_j^\pi - P_j^\pi)(x_j, dx_{j+1}) \\
&\quad Q_{j-1}^\pi(x_{j-1}, dx_j) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
&= \sum_{j=n}^{N-2} \int_E \cdots \int_E \int_E \\
&\quad \left( \sum_{k=j+1}^{N-1} \int_E \cdots \int_E r_k(x_k, f_k(x_k)) P_{k-1}^\pi(x_{k-1}, dx_k) \cdots P_{j+1}^\pi(x_{j+1}, dx_{j+2}) \right. \\
&\quad \left. + \int_E \cdots \int_E r_N(x_N) P_{N-1}^\pi(x_{N-1}, dx_N) \cdots P_{j+1}^\pi(x_{j+1}, dx_{j+2}) \right) \\
&\quad (Q_j^\pi - P_j^\pi)(x_j, dx_{j+1}) Q_{j-1}^\pi(x_{j-1}, dx_j) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
&\quad + \int_E \cdots \int_E \int_E r_N(x_N) (Q_{N-1}^\pi - P_{N-1}^\pi)(x_{N-1}, dx_N) \\
&\quad Q_{N-2}^\pi(x_{N-2}, dx_{N-1}) \cdots Q_n^\pi(x_n, dx_{n+1})
\end{aligned}$$



$$\begin{aligned}
 &= \sum_{j=n}^{N-2} \int_E \cdots \int_E \int_E V_{j+1}^{P;\pi}(x_{j+1}) (Q_j^\pi - P_j^\pi)(x_j, dx_{j+1}) \\
 &\quad Q_{j-1}^\pi(x_{j-1}, dx_j) \cdots Q_n^\pi(x, dx_{n+1}) \\
 &\quad + \int_E \cdots \int_E \int_E r_N(x_N) (Q_{N-1}^\pi - P_{N-1}^\pi)(x_{N-1}, dx_N) \\
 &\quad Q_{N-2}^\pi(x_{N-2}, dx_{N-1}) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
 &= \sum_{j=n}^{N-1} \int_E \cdots \int_E \int_E V_{j+1}^{P;\pi}(x_{j+1}) (Q_j^\pi - P_j^\pi)(x_j, dx_{j+1}) \\
 &\quad Q_{j-1}^\pi(x_{j-1}, dx_j) \cdots Q_n^\pi(x, dx_{n+1}),
 \end{aligned}$$

where we used the conventions

$$\begin{aligned}
 &\int_E \int_E r_N(x_N) P_{N-1}^\pi(x_{N-1}, dx_N) P_N^\pi(x_{j+1}, dx_{N+1}) := r_N(x_N), \\
 &\int_E \cdots \int_E \int_E V_{n+1}^{P;\pi}(x_{n+1}) (Q_n^\pi - P_n^\pi)(x_n, dx_{n+1}) \\
 &\quad Q_{n-1}^\pi(x_{n-1}, dx_n) \cdots Q_n^\pi(x, dx_{n+1}) \\
 &:= \int_E V_{n+1}^{P;\pi}(x_{n+1}) (Q_n^\pi - P_n^\pi)(x_n, dx_{n+1}),
 \end{aligned}$$

and other similar ones. Therefore,

$$\begin{aligned}
 &|\bar{V}_n^{x_n}(Q) - \bar{V}_n^{x_n}(P)| \\
 &= \left| \sup_{\pi \in \Pi} V_n^{Q;\pi}(x_n) - \sup_{\pi \in \Pi} V_n^{P;\pi}(x_n) \right| \\
 &\leq \sup_{\pi \in \Pi} |V_n^{Q;\pi}(x_n) - V_n^{P;\pi}(x_n)| \\
 &\leq \sum_{j=n}^{N-1} \sup_{\pi \in \Pi} \sup_{x \in E} \left( \left| \int_E V_{j+1}^{P;\pi}(y) Q_j^\pi(x, dy) - \int_E V_{j+1}^{P;\pi}(y) P_j^\pi(x, dy) \right| \frac{1}{\psi(x)} \right. \\
 &\quad \left. \times \int_E \cdots \int_E \psi(x_j) Q_{j-1}^\pi(x_{j-1}, dx_j) \cdots Q_n^\pi(x_n, dx_{n+1}) \right)
 \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{j=n}^{N-1} \sup_{\pi \in \Pi} \left( \varrho_{\mathbb{M}'}(V_{j+1}^{P;\pi}) \right. \\
&\quad \times \sup_{(x,a) \in D_j} \left( d_{\mathbb{M}} \left( Q_j((x,a), \cdot), P_j((x,a), \cdot) \right) \frac{1}{\psi(x)} \right) \\
&\quad \times \int_E \cdots \int_E \psi(x_j) Q_{j-1}^\pi(x_{j-1}, dx_j) \cdots Q_n^\pi(x_n, dx_{n+1}) \Big) \\
&\leq \sum_{j=n}^{N-1} \sup_{\pi \in \Pi} \left( \varrho_{\mathbb{M}'}(V_{j+1}^{P;\pi}) d_{\mathbb{M}, \psi}(Q, P) \right. \\
&\quad \times \int_E \cdots \int_E \psi(x_j) Q_{j-1}^\pi(x_{j-1}, dx_j) \cdots Q_n^\pi(x_n, dx_{n+1}) \Big).
\end{aligned}$$

For the latter multiple integral, we have

$$\begin{aligned}
&\int_E \cdots \int_E \psi(x_j) Q_{j-1}^\pi(x_{j-1}, dx_j) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
&\leq \int_E \cdots \int_E \int_E \psi(x_j) P_{j-1}^\pi(x_{j-1}, dx_j) Q_{j-2}^\pi(x_{j-2}, dx_{j-1}) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
&\quad + \int_E \cdots \int_E \left| \int_E \psi(x_j) (Q_{j-1}^\pi(x_{j-1}, dx_j) - P_{j-1}^\pi(x_{j-1}, dx_j)) \right| \\
&\quad \quad Q_{j-2}^\pi(x_{j-2}, dx_{j-1}) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
&\leq \int_E \cdots \int_E K_{3,P} \psi(x_{j-1}) Q_{j-2}^\pi(x_{j-2}, dx_{j-1}) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
&\quad + \int_E \cdots \int_E \varrho_{\mathbb{M}'}(\psi) d_{\mathbb{M}}(Q_{j-1}^\pi(x_{j-1}, \cdot), P_{j-1}^\pi(x_{j-1}, \cdot)) \\
&\quad \quad Q_{j-2}^\pi(x_{j-2}, dx_{j-1}) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
&\leq K_{3,P} \int_E \cdots \int_E \psi(x_{j-1}) Q_{j-2}^\pi(x_{j-2}, dx_{j-1}) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
&\quad + \varrho_{\mathbb{M}'}(\psi) d_{\mathbb{M}, \psi}(Q, P) \\
&\quad \times \int_E \cdots \int_E \psi(x_{j-1}) Q_{j-2}^\pi(x_{j-2}, dx_{j-1}) \cdots Q_n^\pi(x_n, dx_{n+1}) \\
&= (K_{3,P} + \varrho_{\mathbb{M}'}(\psi) d_{\mathbb{M}, \psi}(Q, P)) \\
&\quad \times \int_E \cdots \int_E \psi(x_{j-1}) Q_{j-2}^\pi(x_{j-2}, dx_{j-1}) \cdots Q_n^\pi(x_n, dx_{n+1}).
\end{aligned}$$

Treating the remaining multiple integral analogously and proceeding iteratively in this way, we may continue with

$$\begin{aligned}
 &= (K_{3,P} + \varrho_{\mathbb{M}}(\psi) d_{\mathbb{M},\psi}(Q, P))^{n-j-1} \int_E \psi(x_{j-1}) Q_n^\pi(x_n, dx_{n+1}) \\
 &\leq (K_{3,P} + \varrho_{\mathbb{M}}(\psi) d_{\mathbb{M},\psi}(Q, P))^{n-j-1} \\
 &\quad \times \left( \int_E \psi(x_{j-1}) P_n^\pi(x_n, dx_{n+1}) \right. \\
 &\quad \left. + \int_E \psi(x_{j-1}) (Q_n^\pi(x_n, dx_{n+1}) - P_n^\pi(x_n, dx_{n+1})) \right) \\
 &\leq (K_{3,P} + \varrho_{\mathbb{M}}(\psi) d_{\mathbb{M},\psi}(Q, P))^{n-j-1} \left( K_{3,P} \psi(x_n) + \varrho_{\mathbb{M}}(\psi) d_{\mathbb{M},\psi}(Q, P) \psi(x_n) \right) \\
 &1 = (K_{3,P} + \varrho_{\mathbb{M}}(\psi) d_{\mathbb{M},\psi}(Q, P))^{n-j} \psi(x_n).
 \end{aligned}$$

Altogether, we obtain the asserted inequality.  $\square$

### A.17 Proof of Lemma 6.5

Let  $P \in \mathcal{P}_\alpha$ , i.e.,  $P = P^{\vec{\mu}}$  for some  $\vec{\mu} = (\mu_n)_{n=1}^N \in \mathcal{M}_1^\alpha(\mathbb{R}_{++})^N$ . We have to verify that the defining conditions (a)–(c) of a bounding functions are satisfied. Conditions (a) and (b) are trivially satisfied. Since  $\langle a, \mathbf{1} \rangle \in [0, x]$  for any  $(x, a) \in D$  and  $\langle a, z \rangle \leq \langle a, \langle z, \mathbf{1} \rangle \mathbf{1} \rangle = \langle a, \mathbf{1} \rangle \langle z, \mathbf{1} \rangle$  for any  $a, z \in \mathbb{R}_+^d$ , we also have

$$\begin{aligned}
 &\int_{\mathbb{R}_+} \psi_\alpha(y) P_n^{\vec{\mu}}((x, a), dy) \\
 &= 1 + \int_{\mathbb{R}_+} u_\alpha(\eta_{n,(x,a)}(z)) \mu_{n+1}(dz) \\
 &= 1 + \int_{\mathbb{R}_{++}} (Z_{n+1}^0 x + \langle a, z - Z_{n+1}^0 \mathbf{1} \rangle)^\alpha \mu_{n+1}(dz) \\
 &\leq 1 + (Z_{n+1}^0)^\alpha (x - \langle a, \mathbf{1} \rangle)^\alpha + \int_{\mathbb{R}_{++}} \langle a, z \rangle^\alpha \mu_{n+1}(dz) \\
 &\leq 1 + \left( \max_{k=1,\dots,N} Z_k^0 \right)^\alpha x^\alpha + \int_{\mathbb{R}_{++}} \langle a, \mathbf{1} \rangle^\alpha \langle z, \mathbf{1} \rangle^\alpha \mu_{n+1}(dz) \\
 &\leq 1 + \left( \left( \max_{k=1,\dots,N} Z_k^0 \right)^\alpha + \int_{\mathbb{R}_{++}} \langle z, \mathbf{1} \rangle^\alpha \mu_{n+1}(dz) \right) x^\alpha \leq K_3 \psi_\alpha(x)
 \end{aligned}$$

for any  $n = 0, \dots, N-1$  and  $(x, a) \in D$ , where

$$K_3 := \left( \max_{k=1,\dots,N} Z_k^0 \right)^\alpha + \max_{k=1,\dots,N} \int_{\mathbb{R}_{++}} \langle z, \mathbf{1} \rangle^\alpha \mu_k(dz)$$

is independent of  $n = 0, \dots, N-1$  and  $(x, a) \in D$ . This shows that condition (c) is also satisfied and that  $\mathcal{P}_{\psi_\alpha} = \mathcal{P}_\alpha$ .  $\square$

### A.18 Proof of Theorem 6.6

In view of Theorem C.3 below, we may and do replace without loss of generality the set of all strategies  $\Pi$  by the subset  $\Pi_{\text{lin}}$  of all those  $\pi = (f_n)_{n=0}^{N-1} \in \Pi$  for which

for any  $n = 0, \dots, N-1$ , the decision rule  $f_n : \mathbb{R}_+ \rightarrow \mathbb{R}^d$  admits the representation  $f_n(x) = \kappa_n x$  for some  $\kappa_n \in K$ . For any  $\vec{\kappa} = (\kappa_n)_{n=0}^{N-1} \in K^N$ , we set  $\pi_{\vec{\kappa}} := (f_n^{\kappa_n})_{n=0}^{N-1}$  with  $f_n^{\kappa_n}(x) := \kappa_n x$ . Thus  $K^N$  can be seen as a parameter set for  $\Pi_{\text{lin}}$ .

We now show that the assumptions of Corollary 6.3 are met, so that this result ensures the assertion of Theorem 6.6. Let  $\vec{\mu} \in \mathcal{M}_1^\alpha(\mathbb{R}_{++}^d)^N$ . By Lemma 6.5, we know that  $\psi_\alpha$  is a bounding function for  $Q \in \mathcal{P}_\alpha$ , and obviously  $\varrho_{\mathbb{M}_{\text{HöL},\alpha}}(\psi_\alpha) < \infty$ . So it remains to show that  $\sup_{\vec{\kappa} \in K^N} \varrho_{\mathbb{M}_{\text{HöL},\alpha}}(V_{n+1}^{P^{\vec{\mu}};\pi_{\vec{\kappa}}}) < \infty$  for  $n = 0, \dots, N-1$ . By Lemma C.4, we have for any  $n = 0, \dots, N-1$  and  $\vec{\kappa} \in K^N$  that  $V_n^{P^{\vec{\mu}};\pi_{\vec{\kappa}}}(\cdot) = \phi_n^{\vec{\mu};\pi_{\vec{\kappa}}} u_\alpha(\cdot)$ , where  $\phi_n^{\vec{\mu};\pi_{\vec{\kappa}}} := \prod_{j=n}^{N-1} \gamma_j^{\vec{\mu};\kappa_j}$  is finite, and therefore  $\varrho_{\mathbb{M}_{\text{HöL},\alpha}}(V_{n+1}^{P^{\vec{\mu}};\pi_{\vec{\kappa}}}) = \phi_n^{\vec{\mu};\pi_{\vec{\kappa}}} < \infty$  since  $\varrho_{\mathbb{M}_{\text{HöL},\alpha}}(u_\alpha) = 1$ . Along with Lemma C.2, we conclude that the assumptions of Corollary 6.3 are indeed met.  $\square$

### A.19 Proof of Corollary 6.7

Recall the definition  $\mathbb{M}_{\text{HöL},\alpha}^d := \{h \in \mathbb{R}^{\mathbb{R}_+^d} : \|h\|_{\text{HöL},\alpha}^d \leq 1\}$  with the Hölder- $\alpha$  norm  $\|h\|_{\text{HöL},\alpha}^d := \sup_{z_1, z_2 \in \mathbb{R}_+^d : z_1 \neq z_2} |h(z_1) - h(z_2)| / |z_1 - z_2|^\alpha$ . It is known from Kern et al. [26] that  $d_{\mathbb{M}_{\text{HöL},\alpha}^d}$  (defined analogously to (6.2)) provides a metric on  $\mathcal{M}_1^\alpha(\mathbb{R}_{++}^d)$  that metrises the  $\alpha$ -weak topology  $\mathcal{O}_d^\alpha(\mathbb{R}_{++}^d)$ .

We now show that the mapping  $\mathcal{M}_1^\alpha(\mathbb{R}_{++}) \rightarrow \mathcal{P}_\alpha, \mu \mapsto P^{\bar{\mu}}$ , is continuous for the pair  $(d_{\mathbb{M}_{\text{HöL},\alpha}^d}, d_{\mathbb{M}_{\text{HöL},\alpha}, \psi_\alpha})$ . The assertion of Corollary 6.7 then directly follows from Theorem 6.6. For any  $v, \mu \in \mathcal{M}_1^\alpha(\mathbb{R}_{++})$ , we have

$$\begin{aligned} & d_{\mathbb{M}_{\text{HöL},\alpha}, \psi_\alpha}(P^{\bar{v}}, P^{\bar{\mu}}) \\ &= \max_{n=0, \dots, N-1} \sup_{(x,a) \in D_n} d_{\mathbb{M}_{\text{HöL},\alpha}^d} \left( Q_n^{\bar{v}}((x,a), \cdot), P_n^{\bar{\mu}}((x,a), \cdot) \right) / \psi_\alpha(x) \\ &= \max_{n=0, \dots, N-1} \sup_{(x,a) \in D} d_{\mathbb{M}_{\text{HöL},\alpha}^d} (v \circ \eta_{n,(x,a)}^{-1}[\cdot], \mu \circ \eta_{n,(x,a)}^{-1}[\cdot]) / \psi_\alpha(x) \\ &= \max_{n=0, \dots, N-1} \sup_{(x,a) \in D} \sup_{v \in \mathbb{M}_{\text{HöL},\alpha}^d} \left| \int_{\mathbb{R}_+^d} v(Z_{n+1}^0(x - \langle a, \mathbf{1} \rangle) + \langle a, z \rangle) (v(dz) - \mu(dz)) \right| \frac{1}{\psi_\alpha(x)} \\ &\leq \sup_{(x,a) \in D} \sup_{v \in \mathbb{M}_{\text{HöL},\alpha}^d} \sup_{y \in \mathbb{R}_+^d} \left| \int_{\mathbb{R}_+^d} v(y + \langle a, z \rangle) v(dz) - \int_{\mathbb{R}_+^d} v(y + \langle a, z \rangle) \mu(dz) \right| \frac{1}{\psi_\alpha(x)} \\ &\leq \sup_{(x,a) \in D} \sup_{w \in \mathbb{M}_{\text{HöL},\alpha}^d} \left| \int_{\mathbb{R}_+^d} w(\langle a, z \rangle) v(dz) - \int_{\mathbb{R}_+^d} w(\langle a, z \rangle) \mu(dz) \right| \frac{1}{\psi_\alpha(x)} \\ &= \sup_{(x,a) \in D} \sup_{w \in \mathbb{M}_{\text{HöL},\alpha}^d} \left| \int_{\mathbb{R}_+^d} h_{w,a}(z) v(dz) - \int_{\mathbb{R}_+^d} h_{w,a}(z) \mu(dz) \right| \frac{1}{\psi_\alpha(x)}, \end{aligned}$$

where  $D_n := D := \{(x, a) \in \mathbb{R}_+^{d+1} : \langle a, \mathbf{1} \rangle \leq x\}$  and  $h_{w,a}(z) := w(\langle a, z \rangle)$ . For the map  $h_{w,a} : \mathbb{R}_+^d \rightarrow \mathbb{R}$ , we have

$$\begin{aligned} \|h_{w,a}\|_{\text{Hö},\alpha}^d &= \sup_{z_1, z_2 \in \mathbb{R}_+^d : z_1 \neq z_2} |h_{w,a}(z_1) - h_{w,a}(z_2)| / |z_1 - z_2|^\alpha \\ &= \sup_{z_1, z_2 \in \mathbb{R}_+^d : z_1 \neq z_2} \frac{|\langle a, z_1 \rangle - \langle a, z_2 \rangle|^\alpha}{|z_1 - z_2|^\alpha} \frac{|h_{w,a}(z_1) - h_{w,a}(z_2)|^\alpha}{|\langle a, z_1 \rangle - \langle a, z_2 \rangle|^\alpha} \\ &= \sup_{z_1, z_2 \in \mathbb{R}_+^d : z_1 \neq z_2} \frac{x^\alpha |z_1 - z_2|_\infty^\alpha}{|z_1 - z_2|^\alpha} \|w\|_{\text{Hö},\alpha} \leq c_\infty^\alpha x^\alpha, \end{aligned}$$

where  $c_\infty \in \mathbb{R}_{++}$  is chosen such that  $|\cdot|_\infty \leq c_\infty |\cdot|$  and we used that

$$|\langle a, z_1 - z_2 \rangle| \leq \langle a, \mathbf{1} \rangle |z_1 - z_2|_\infty \leq x |z_1 - z_2|_\infty.$$

Since  $\psi_\alpha(x) = 1 + x^\alpha$ , the above calculation can therefore be continued with

$$\leq \sup_{x \in \mathbb{R}_+} (c_\infty^\alpha x^\alpha d_{\mathbb{M}_{\text{Hö},\alpha}^d}(\nu, \mu) / \psi_\alpha(x)) \leq c_\infty^\alpha d_{\mathbb{M}_{\text{Hö},\alpha}^d}(\nu, \mu). \quad \square$$

## Appendix B: A comment on the relation between $d_{\mu_1, \mu_2}$ and the metric introduced in [40]

Rachasingho and Tasena [40] recently defined a distance  $d$  on the set of bivariate subcopulas as follows. For two bivariate subcopulas  $C_0$  and  $C'_0$ , they put

$$d(C_0, C'_0) := \mathfrak{h}_{d_{[0,1]^2}}([C_0], [C'_0]) + \mathfrak{h}_{|\cdot|}(\text{dom}(C_0), \text{dom}(C'_0)),$$

where the summand  $\mathfrak{h}_{d_{[0,1]^2}}([C_0], [C'_0])$  is the Hausdorff distance (with respect to  $d_{[0,1]^2}$ ) between the sets of bivariate copulas  $[C_0]$  and  $[C'_0]$  induced by  $C_0$  and  $C'_0$ , respectively, and the summand  $\mathfrak{h}_{|\cdot|}(\text{dom}(C_0), \text{dom}(C'_0))$  is the Hausdorff distance between the domains  $\text{dom}(C_0)$  and  $\text{dom}(C'_0)$  of  $C_0$  and  $C'_0$ , respectively.

The distance  $d_{\mu_1, \mu_2}$  defined by (2.2) basically differs from  $d$  for the following reasons. First of all,  $d$  is a metric on the set of bivariate subcopulas, whereas  $d_{\mu_1, \mu_2}$  is a pseudo-metric on the set of bivariate copulas. Moreover,  $d_{\mu_1, \mu_2}$  is designed to be a reasonable distance measure on the set of copulas associated with probability measures from the Fréchet class  $\mathcal{M}_2(\mu_1, \mu_2)$ ; it defines the distance between two such copulas  $C$  and  $C'$  by the maximal pointwise distance between the corresponding subcopulas with domain  $K := \text{ran} F_{\mu_1} \times \text{ran} F_{\mu_2}$ . On the other hand,  $d$  is designed to be a reasonable distance measure on the set of arbitrary subcopulas. Even when restricting  $d$  to the set of those subcopulas with domain  $K$  (and  $d_{\mu_1, \mu_2}$  to the set of copulas associated with probability measures from  $\mathcal{M}_2(\mu_1, \mu_2)$ ), the resulting distance measures are different. Indeed, if  $C_0$  and  $C'_0$  are two subcopulas with domain  $K$ , then

$$d_{\mu_1, \mu_2}(C, C') = \sup_{u \in K} |C_0(u) - C'_0(u)|,$$

but  $d(C_0, C'_0) \geq \sup_{u \in [0,1]^2} |C(u) - C'(u)|$  for any copulas  $C$  and  $C'$  induced by  $C_0$  and  $C'_0$ , respectively. Note here that  $d_{\mu_1, \mu_2}$  is defined in such a way that it does not take into account the behaviour of copulas outside  $K = \text{ran} F_{\mu_1} \times \text{ran} F_{\mu_2}$ , which is motivated by the fact that the behaviour of a copula  $C$  outside  $K$  is irrelevant for a probability measure from the Fréchet class  $\mathcal{M}_2(\mu_1, \mu_2)$  with copula  $C$ .

## Appendix C: Supplements to Sect. 6

Here we discuss the existence of optimal strategies in the Markov decision model considered in Sect. 6. In Sect. C.1, we first consider the general model introduced in Sect. 6.1. Thereafter, in Sect. C.2, we study in detail the special case of the multi-period portfolio optimisation problem considered in Sect. 6.2.

### C.1 Existence of optimal strategies in the general model

For any  $n = 0, \dots, N-1$  and  $P \in \mathcal{P}$ , denote by  $\mathbb{M}_n^P(E)$  the set of all  $v \in \mathbb{M}(E)$  for which  $\int_E |v(y)| P_n((x, f_n(x)), dy) < \infty$  for any  $x \in E$  and  $f_n \in F_n$ . Recall that  $\mathbb{M}(E)$  is the set of all  $(\mathcal{E}, \mathcal{B}(\mathbb{R}))$ -measurable maps  $v : E \rightarrow \mathbb{R}$ . For any  $n = 0, \dots, N-1$ ,  $f_n \in F_n$  and  $v \in \mathbb{M}_n^P(E)$ , we define maps  $T_{n, f_n}^P v : E \rightarrow \mathbb{R}$  and  $T_n^P v : E \rightarrow \mathbb{R} \cup \{\infty\}$  by

$$T_{n, f_n}^P v(x) := r_n(x, f_n(x)) + \int_E v(x') P_n((x, f_n(x)), dx'),$$

$$T_n^P v(x) := \sup_{f_n \in F_n} T_{n, f_n}^P v(x).$$

Note that  $T_{n, f_n}^P$  and  $T_n^P$  can be seen as maps from  $\mathbb{M}_n^P(E)$  to  $\mathbb{M}(E)$  and from  $\mathbb{M}_n^P(E)$  to  $(\mathbb{R} \cup \{\infty\})^E$ , respectively.

For any  $n = 0, \dots, N-1$ ,  $P \in \mathcal{P}$  and  $v \in \mathbb{M}_n^P(E)$ , a decision rule  $f_n^P \in F_n$  is said to be a *maximiser of  $v$*  if  $T_{n, f_n^P}^P v(x) = T_n^P v(x)$  for all  $x \in E$ . The following result is known from Bäuerle and Rieder [1, Theorem 2.3.8].

**Theorem C.1** *Let  $P \in \mathcal{P}$  and assume that for any  $n = 0, \dots, N-1$ , there exist sets  $\mathbb{M}_n^P \subseteq \mathbb{M}_n^P(E)$  and  $F_n^P \subseteq F_n$  such that the following three conditions hold:*

- (a)  $r_N \in \mathbb{M}_{N-1}^P$ .
- (b)  $T_n^P v \in \mathbb{M}_{n-1}^P$  for any  $v \in \mathbb{M}_n^P$  and  $n = 1, \dots, N-1$ .
- (c) For any  $n = 0, \dots, N-1$  and  $v \in \mathbb{M}_n^P$ , there exists an  $f_n^P \in F_n^P$  that is a maximiser of  $v$ .

*Then the following three assertions hold true:*

- (i)  $V_0^P \in \mathbb{M}(E)$  and  $V_{n+1}^P \in \mathbb{M}_n^P$ ,  $n = 0, \dots, N-1$ . Moreover, the Bellman iteration scheme holds true, i.e.,  $V_N^P = r_N$  and  $V_n^P = T_n^P V_{n+1}^P$ ,  $n = 0, \dots, N-1$ .
- (ii)  $V_n^P = T_n^P T_{n+1}^P \cdots T_{N-1}^P r_N$  for any  $n = 0, \dots, N-1$ .
- (iii) For any  $n = 0, \dots, N-1$ , there exists an  $f_n^P \in F_n^P$  that is a maximiser of  $V_{n+1}^P$ . Any such maximisers  $f_0^P, \dots, f_{N-1}^P$  form a strategy  $\pi^P := (f_n^P)_{n=0}^{N-1} \in \Pi$  that is optimal for the optimisation problem (6.1).

## C.2 Existence of optimal trading strategies in the setting of Sect. 6.2

We now focus on the specific setting of Sect. 6.2, i.e., we discuss the existence of optimal trading strategies for the multi-period portfolio optimisation problem considered there. Let  $K := \{\kappa \in [0, 1]^d : \langle \kappa, \mathbf{1} \rangle \leq 1\}$  and note that  $K$  is compact. For any  $\bar{\mu} = (\mu_n)_{n=1}^N \in \mathcal{M}_1^\alpha(\mathbb{R}_{++})^N$ ,  $\kappa \in K$  and  $n = 0, \dots, N-1$ , set  $\gamma_n^{\bar{\mu}; \kappa} := \int_{\mathbb{R}_+^d} u_\alpha(Z_{n+1}^0 + \langle \kappa, z - Z_{n+1}^0 \mathbf{1} \rangle) \mu_{n+1}(dz)$ . Moreover, set  $\gamma_n^{\bar{\mu}} := \sup_{\kappa \in K} \gamma_n^{\bar{\mu}; \kappa}$  for any  $n = 0, \dots, N-1$ .

**Lemma C.2** *For any  $\bar{\mu} = (\mu_n)_{n=1}^N \in \mathcal{M}_1^\alpha(\mathbb{R}_{++})^N$  and  $n = 0, \dots, N-1$ , there exists at least one solution  $\kappa_n^{\bar{\mu}} \in K$  to the optimisation problem  $\max\{\gamma_n^{\bar{\mu}; \kappa} : \kappa \in K\}$ . In particular, the maximal value  $\gamma_n^{\bar{\mu}} = \gamma_n^{\bar{\mu}; \kappa_n^{\bar{\mu}}}$  is finite.*

**Proof** Let  $\bar{\mu} = (\mu_n)_{n=1}^N \in \mathcal{M}_1^\alpha(\mathbb{R}_{++})^N$  and  $n \in \{0, \dots, N-1\}$ . Define a map  $g_n : \mathbb{R}_+^d \times K \rightarrow \mathbb{R}_+$  by  $g_n(z, \kappa) := u_\alpha(Z_{n+1}^0 + \langle \kappa, z - Z_{n+1}^0 \mathbf{1} \rangle)$ . The map  $g_n(\cdot, \kappa)$  is Borel-measurable for any fixed  $\kappa \in K$ , and we have

$$|g_n(z, \kappa)| = u_\alpha(Z_{n+1}^0 + \langle \kappa, z - Z_{n+1}^0 \mathbf{1} \rangle) \leq u_\alpha(Z_{n+1}^0 + \langle \mathbf{1}, z \rangle)$$

for any  $(z, \kappa) \in \mathbb{R}_+^d \times K$ . Therefore,  $g_n$  is dominated by the Borel-measurable function  $h_n : \mathbb{R}_+^d \rightarrow \mathbb{R}_+$  defined by  $h_n(z) := u_\alpha(Z_{n+1}^0 + \langle z, \mathbf{1} \rangle)$ . This function is  $\mu_{n+1}$ -integrable since  $\int_{\mathbb{R}_+^d} h_n d\mu_{n+1} \leq (Z_{n+1}^0)^\alpha + \int_{\mathbb{R}_+^d} \langle \mathbf{1}, z \rangle^\alpha \mu_{n+1}(dz) < \infty$  (take into account the definition of  $\mathcal{M}_1^\alpha(\mathbb{R}_{++})$ ), and  $g_n(z, \cdot)$  is continuous on  $K$  for any  $z \in \mathbb{R}_+^d$ . So we can apply the continuity lemma (in the form of Bauer [2, Lemma 16.1]) to obtain that the map  $G_n : K \rightarrow \mathbb{R}_{++}$  defined by  $G_n(\kappa) := \int_{\mathbb{R}_+^d} g_n(z, \kappa) \mu_{n+1}(dz)$  is continuous. Since  $K$  is compact, we can infer that there exists a solution  $\kappa_n^{\bar{\mu}} \in K$  to the optimisation problem  $\max\{\gamma_n^{\bar{\mu}; \kappa} : \kappa \in K\}$ .  $\square$

Part (ii) of the following result shows in particular that an optimal trading strategy can be found in the subset  $\Pi_{\text{lin}}$  of all those  $\pi = (f_n)_{n=0}^{N-1} \in \Pi$  for which for any  $n = 0, \dots, N-1$ , the decision rule  $f_n : \mathbb{R}_+ \rightarrow \mathbb{R}^d$  admits the representation  $f_n(x) = \kappa_n x$  for some  $\kappa_n \in K$ . For any  $n = 0, \dots, N-1$ , let  $\kappa_n^{\bar{\mu}} \in K$  be any solution to the optimisation problem  $\max\{\gamma_n^{\bar{\mu}; \kappa} : \kappa \in K\}$  (see Lemma C.2).

**Theorem C.3** *For any  $\bar{\mu} \in \mathcal{M}_1^\alpha(\mathbb{R}_{++})^N$ , the following two assertions hold true:*

(i) *For any  $n = 0, \dots, N-1$ , the time- $n$  value function  $V_n^{P^{\bar{\mu}}} : \mathbb{R}_+ \rightarrow \mathbb{R}$  admits the representation  $V_n^{P^{\bar{\mu}}}(\cdot) = \phi_n^{\bar{\mu}} u_\alpha(\cdot)$ , where  $\phi_n^{\bar{\mu}} := \prod_{j=n}^{N-1} \gamma_j^{\bar{\mu}}$ .*

(ii) *If for every  $n = 0, \dots, N-1$ , a decision rule  $f_n^{\bar{\mu}} : \mathbb{R}_+ \rightarrow \mathbb{R}_+^d$  at time  $n$  is defined by  $f_n^{\bar{\mu}}(x) := \kappa_n^{\bar{\mu}} x$ , then  $\pi^{\bar{\mu}} := (f_n^{\bar{\mu}})_{n=0}^{N-1}$  forms an optimal trading strategy for  $P^{\bar{\mu}}$ .*

**Proof** (i) We intend to apply Theorem C.1. Let  $M_n^{P^{\bar{\mu}}} := M'$  and  $F_n^{P^{\bar{\mu}}} := F'$  for any  $n = 0, \dots, N-1$ , where  $M' := \{v_\theta : \theta \in \mathbb{R}_{++}\}$  and  $F' := \{f_\kappa : \kappa \in K\}$  with

$v_\theta(x) := \theta u_\alpha(x)$  and  $f_\kappa(x) = \kappa x$ ,  $x \in \mathbb{R}_+$ . It can be inferred from Lemma 6.5 that  $\mathbb{M}_n^{P^{\bar{\mu}}} := \mathbb{M}' \subseteq \mathbb{M}_n^{P^{\bar{\mu}}}(\mathbb{R}_+)$  for any  $n = 0, \dots, N-1$ , where  $\mathbb{M}_n^P(\mathbb{R}_+)$  is defined as in Sect. C.1. Moreover, we clearly have  $F_n^{P^{\bar{\mu}}} := F' \subseteq F =: F_n$  for any  $n = 0, \dots, N-1$ .

Below, we show that conditions (a)–(c) of Theorem C.1 are met. Thus the Bellman iteration scheme in part (i) of Theorem C.1 holds true, and so

$$\begin{aligned}
 V_{N-1}^{P^{\bar{\mu}}}(x_{N-1}) &= T_{N-1}^{P^{\bar{\mu}}} V_N^{P^{\bar{\mu}}}(x_{N-1}) = T_{N-1}^{P^{\bar{\mu}}} r_N(x_{N-1}) = T_{N-1}^{P^{\bar{\mu}}} u_\alpha(x_{N-1}) \\
 &= \sup_{f_{N-1} \in F_{N-1}} T_{N-1, f_{N-1}}^{P^{\bar{\mu}}} u_\alpha(x_{N-1}) \\
 &= \sup_{f_{N-1} \in F_{N-1}} \int_{\mathbb{R}_+} u_\alpha(x_N) P_{N-1} \left( (x_{N-1}, f_{N-1}(x_{N-1})) \right) dx_N \\
 &= \sup_{f_n \in F_{N-1}} \int_{\mathbb{R}_+} (Z_N^0 x_{N-1} + \langle f_{N-1}(x_{N-1}), z - Z_N^0 \mathbf{1} \rangle)^\alpha \mu_N(dz) \\
 &= \sup_{\kappa \in K} \int_{\mathbb{R}_+} (Z_N^0 x_{N-1} + \langle \kappa x_{N-1}, z - Z_N^0 \mathbf{1} \rangle)^\alpha \mu_N(dz) \\
 &= u_\alpha(x_{N-1}) \sup_{\kappa \in K} \int_{\mathbb{R}_+} (Z_N^0 + \langle \kappa, z - Z_N^0 \mathbf{1} \rangle)^\alpha \mu_N(dz) \\
 &= u_\alpha(x_{N-1}) \sup_{\kappa \in K} \gamma_{N-1}^{\bar{\mu}, \kappa} \\
 &= \bar{\gamma}_{N-1}^{\bar{\mu}} u_\alpha(x_{N-1}) = \phi_{N-1}^{\bar{\mu}} u_\alpha(x_{N-1})
 \end{aligned} \tag{C.1}$$

for any  $x_{N-1} \in \mathbb{R}_+$ . If we continue this way, we successively get  $V_n^{P^{\bar{\mu}}}(\cdot) = \phi_n^{\bar{\mu}} u_\alpha(\cdot)$ ,  $n = N-2, \dots, 0$ .

It remains to verify that conditions (a)–(c) of Theorem C.1 are met. Condition (a) is trivially satisfied, and (b) can be shown by proceeding analogously to (C.1). Furthermore, similarly to (C.1), we obtain for any  $n = 0, \dots, N-1$  and  $\theta \in \mathbb{R}_{++}$  that  $\sup_{f_n \in F_n} T_{n, f_n}^{P^{\bar{\mu}}} v_\theta(\cdot) = \theta u_\alpha(\cdot) \sup_{\kappa \in K} \gamma_n^{\bar{\mu}, \kappa}$  and  $\theta u_\alpha(\cdot) \gamma_n^{\bar{\mu}, \kappa} = T_{n, f_\kappa}^{P^{\bar{\mu}}} v_\theta(\cdot)$  for all  $\kappa \in K$ . By Lemma C.2, we know that there exists a maximum point  $\kappa_n^{\bar{\mu}} \in K$  of the map  $\kappa \mapsto \gamma_n^{\bar{\mu}, \kappa}$ , and therefore for any  $n = 0, \dots, N-1$ , the decision rule  $f_{\kappa_n^{\bar{\mu}}}^{\bar{\mu}}$  is a maximiser of  $v_\theta$ . This shows that condition (c) is satisfied, too.

(ii) In the proof of (i), we have seen that the assumptions of Theorem C.1 are fulfilled. Thus Theorem C.1 (i) gives  $V_{n+1}^{P^{\bar{\mu}}} \in \mathbb{M}' =: \mathbb{M}_n^{P^{\bar{\mu}}}$  for any  $n = 0, \dots, N-1$ . In particular, the above elaborations under (c) show that for any  $n = 0, \dots, N-1$ , the decision rule  $f_{\kappa_n^{\bar{\mu}}}^{\bar{\mu}} \in F' =: F_n^{P^{\bar{\mu}}}$  provides a maximiser of  $V_{n+1}^{P^{\bar{\mu}}}$ . Hence Theorem C.1 (iii) ensures that the strategy  $\pi^{\bar{\mu}} := (f_{\kappa_n^{\bar{\mu}}}^{\bar{\mu}})_{n=0}^{N-1} \in \Pi_{\text{lin}}$  forms an optimal trading strategy for  $P^{\bar{\mu}}$ .  $\square$



The following lemma specifies the value functions for fixed strategies. Recall that the values  $\gamma_n^{\bar{\mu};\kappa}$ ,  $n = 0, \dots, N-1$ ,  $\kappa \in K$ , were defined before Lemma C.2, and note that for any  $\bar{\kappa} = (\kappa_n)_{n=0}^{N-1} \in K^N$ , one defines a trading strategy  $\pi_{\bar{\kappa}} := (f_n^{\bar{\kappa}})_{n=0}^{N-1} \in \Pi_{\text{lin}}$  by setting  $f_n^{\bar{\kappa}}(x) := \kappa_n x$  for any  $x \in \mathbb{R}_+$  and  $n = 0, \dots, N-1$ .

**Lemma C.4** *Let  $\bar{\mu} \in \mathcal{M}_1^\alpha(\mathbb{R}_{++})^N$  and  $\bar{\kappa} = (\kappa_n)_{n=0}^{N-1} \in K^N$ . For  $n = 0, \dots, N-1$ , the time- $n$  value function  $V_n^{P^{\bar{\mu}}; \pi_{\bar{\kappa}}} : \mathbb{R}_+ \rightarrow \mathbb{R}$  associated with strategy  $\pi_{\bar{\kappa}}$  then admits the representation  $V_n^{P^{\bar{\mu}}; \pi_{\bar{\kappa}}}(\cdot) = \phi_n^{\bar{\mu}; \pi_{\bar{\kappa}}} u_\alpha(\cdot)$ , where  $\phi_n^{\bar{\mu}; \pi_{\bar{\kappa}}} := \prod_{j=n}^{N-1} \gamma_j^{\bar{\mu}; \kappa_j}$ .*

**Proof** For any  $n = 0, \dots, N-1$  and  $x_n \in \mathbb{R}_+$ , we have

$$\begin{aligned} V_n^{P^{\bar{\mu}}; \pi_{\bar{\kappa}}}(x) &= \mathbb{E}^{x_0, P^{\bar{\mu}}; \pi_{\bar{\kappa}}}[r_N(X_n) | X_n = x_n] = \mathbb{E}^{x_0, P^{\bar{\mu}}; \pi_{\bar{\kappa}}}[u_\alpha(X_n) | X_n = x_n] \\ &= \int_{\mathbb{R}_+} \cdots \int_{\mathbb{R}_+} \int_{\mathbb{R}_+} u_\alpha(x_N) P^{\bar{\mu}; \pi_{\bar{\kappa}}}(x_{N-1}, dx_N) \\ &\quad P^{\bar{\mu}; \pi_{\bar{\kappa}}}(x_{N-2}, dx_{N-1}) \cdots P^{\bar{\mu}; \pi_{\bar{\kappa}}}(x_n, dx_{n+1}) \\ &= \int_{\mathbb{R}_+} \cdots \int_{\mathbb{R}_+} \int_{\mathbb{R}_+} (Z_N^0 x_{N-1} + \langle \kappa_{N-1} x_{N-1}, z - Z_N^0 \mathbf{1} \rangle)^\alpha \mu_N(dz) \\ &\quad P^{\bar{\mu}; \pi_{\bar{\kappa}}}(x_{N-2}, dx_{N-1}) \cdots P^{\bar{\mu}; \pi_{\bar{\kappa}}}(x_n, dx_{n+1}) \\ &= \int_{\mathbb{R}_+} \cdots \int_{\mathbb{R}_+} x_{N-1}^\alpha \int_{\mathbb{R}_+} (Z_N^0 + \langle \kappa_{N-1}, z - Z_N^0 \mathbf{1} \rangle)^\alpha \mu_N(dz) \\ &\quad P^{\bar{\mu}; \pi_{\bar{\kappa}}}(x_{N-2}, dx_{N-1}) \cdots P^{\bar{\mu}; \pi_{\bar{\kappa}}}(x_n, dx_{n+1}) \\ &= \gamma_N^{\bar{\mu}; \kappa_N} \int_{\mathbb{R}_+} \cdots \int_{\mathbb{R}_+} u_\alpha(x_{N-1}) P^{\bar{\mu}; \pi_{\bar{\kappa}}}(x_{N-2}, dx_{N-1}) \cdots P^{\bar{\mu}; \pi_{\bar{\kappa}}}(x_n, dx_{n+1}). \end{aligned}$$

If we continue successively in this way, we end up with

$$V_n^{P^{\bar{\mu}}; \pi_{\bar{\kappa}}}(x) = \prod_{j=n+1}^{N-1} \gamma_j^{\bar{\mu}; \kappa_j} \int_{\mathbb{R}_+} u_\alpha(x_{n+1}) P^{\bar{\mu}; \pi_{\bar{\kappa}}}(x_n, dx_{n+1}) = \prod_{j=n}^{N-1} \gamma_j^{\bar{\mu}; \kappa_j} u_\alpha(x_n).$$

This proves the assertion of the lemma, since  $\prod_{j=n}^{N-1} \gamma_j^{\bar{\mu}; \kappa_j} = \phi_n^{\bar{\mu}; \pi_{\bar{\kappa}}}$ . □

**Acknowledgements** The author would like to thank an Associate Editor and two referees for their constructive comments and suggestions that led to an improved version of the article.

**Funding Note** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article

are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Bäuerle, N., Rieder, U.: Markov Decision Processes with Applications to Finance. Springer, Berlin (2011)
2. Bauer, H.: Measure and Integration Theory. de Gruyter, Berlin (2001)
3. Bellini, F., Klar, B., Müller, A., Rosazza Gianin, E.: Generalized quantiles as risk measures. *Insur. Math. Econ.* **54**, 41–48 (2014)
4. Berge, C.: Topological Spaces. Macmillan, New York (1963)
5. Bickel, P.J., Freedman, D.A.: Some asymptotic theory for the bootstrap. *Ann. Stat.* **9**, 1196–1217 (1981)
6. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer, New York (2000)
7. Charpentier, A., Segers, J.: Convergence of Archimedean copulas. *Stat. Probab. Lett.* **78**, 412–419 (2008)
8. Cheridito, P., Li, T.: Risk measures on Orlicz hearts. *Math. Finance* **19**, 189–214 (2009)
9. Claus, M., Krättschmer, V., Schultz, R.: Weak continuity of risk functionals with applications to stochastic programming. *SIAM J. Optim.* **27**, 91–109 (2017)
10. Deheuvels, P.: Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes. *Publ. Inst. Stat. Univ. Paris* **23**, 1–36 (1978)
11. Demarta, S., McNeil, A.J.: The  $t$  copula and related copulas. *Int. Stat. Rev.* **73**, 111–129 (2005)
12. Durante, F., Sempi, C.: Principles of Copula Theory. Taylor & Francis, London (2015)
13. Embrechts, P., Liu, H., Wang, R.: Quantile-based risk sharing. *Oper. Res.* **66**, 936–949 (2018)
14. Embrechts, P., Puccetti, G.: Bounds for functions of dependent risks. *Finance Stoch.* **10**, 341–352 (2006)
15. Embrechts, P., Puccetti, G., Rüschendorf, L.: Model uncertainty and VaR aggregation. *J. Bank. Finance* **37**, 2750–2764 (2013)
16. Embrechts, P., Schied, A., Wang, R.: Robustness in the optimization of risk measures. *Oper. Res.* **70**, 95–110 (2021)
17. Embrechts, P., Wang, B., Wang, R.: Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance Stoch.* **19**, 763–790 (2015)
18. Emmer, S., Kratz, M., Tasche, D.: What is the best risk measure in practice? A comparison of standard measures. *J. Risk* **18**, 31–60 (2015)
19. Fernández Sánchez, J., Trutschnig, W.: Conditioning-based metrics on the space of multivariate copulas and their interrelation with uniform and levelwise convergence and iterated function systems. *J. Theor. Probab.* **28**, 1311–1336 (2015)
20. Filipović, D., Kupper, M.: Optimal capital and risk transfers for group diversification. *Math. Finance* **18**, 55–76 (2007)
21. Filipović, D., Svindland, G.: Optimal capital and risk allocations for law- and cash-invariant convex functions. *Finance Stoch.* **12**, 423–439 (2008)
22. Föllmer, H., Schied, A.: Robust preferences and convex measures of risk. In: Sandmann, K., Schönbucher, P. (eds.) *Advances in Finance and Stochastics. Essays in Honour of Dieter Sondermann*, pp. 39–56. Springer, Berlin (2002)
23. Föllmer, H., Schied, A.: *Stochastic Finance. An Introduction in Discrete Time*, 3rd revised and extended edn. de Gruyter, Berlin (2011)
24. Kasper, T., Fuchs, S., Trutschnig, W.: On weak conditional convergence of bivariate Archimedean and extreme value copulas, and consequences to nonparametric estimation. *Bernoulli* **27**, 2217–2240 (2021)
25. Kern, P.: Sensitivity and Statistical Inference in Markov Decision Models and Collective Risk Models. PhD thesis, Saarland University, Saarbrücken, (2020). Available online at <https://doi.org/10.22028/D291-32385>

26. Kern, P., Simroth, A., Zähle, H.: First-order sensitivity of the optimal value in a Markov decision model with respect to deviations in the transition probability function. *Math. Methods Oper. Res.* **92**, 165–197 (2020)
27. Krätschmer, V., Schied, A., Zähle, H.: Comparative and qualitative robustness for law-invariant risk measures. *Finance Stoch.* **18**, 271–295 (2014)
28. Krätschmer, V., Schied, A., Zähle, H.: Domains of weak continuity of statistical functionals with a view toward robust statistics. *J. Multivar. Anal.* **158**, 1–19 (2017)
29. Krätschmer, V., Zähle, H.: Statistical inference for expectile-based risk measures. *Scand. J. Stat.* **44**, 425–454 (2017)
30. Li, X., Mikusiński, P., Taylor, M.D.: Strong approximation of copulas. *J. Math. Anal. Appl.* **225**, 608–623 (1998)
31. Lindner, A., Szimayer, A.: A limit theorem for copulas. Discussion paper 433, Sonderforschungsbereich 386, Ludwig-Maximilian-Universität München (2005). Available online at <https://pub.uni-muenchen.de/1802/>
32. Markowitz, H.M.: Portfolio selection. *J. Finance* **7**, 77–91 (1952)
33. McNeil, A.J., Frey, R., Embrechts, P.: *Quantitative Risk Management*, 1st edn. Princeton University Press, Princeton (2005)
34. Müller, A.: How does the value function of a Markov decision process depend on the transition probabilities? *Math. Oper. Res.* **22**, 872–885 (1997)
35. Müller, A.: Integral probability metrics and their generating classes of functions. *Adv. Appl. Probab.* **29**, 429–443 (1997)
36. Nelsen, R.B.: *An Introduction to Copulas*, 2nd edn. Springer, New York (2006)
37. Ogryczak, W., Ruszczyński, A.: From stochastic dominance to mean–risk models: semideviations as risk measures. *Eur. J. Oper. Res.* **116**, 33–50 (1999)
38. Ogryczak, W., Ruszczyński, A.: Dual stochastic dominance and related mean–risk models. *SIAM J. Optim.* **13**, 60–78 (2002)
39. Puccetti, G.: Sharp bounds on the expected shortfall for a sum of dependent random variables. *Stat. Probab. Lett.* **83**, 1227–1232 (2013)
40. Rachasingho, J., Tasena, S.: A metric space of subcopulas – an approach via Hausdorff distance. *Fuzzy Sets Syst.* **378**, 144–156 (2020)
41. Rachev, S.T.: *Probability Metrics and the Stability of Stochastic Models*. Wiley, Chichester (1991)
42. Rudin, W.: *Functional Analysis*, 2nd edn. McGraw-Hill, New York (1991)
43. Rüschendorf, L.: Random variables with maximum sums. *Adv. Appl. Probab.* **14**, 623–632 (1982)
44. Rüschendorf, L.: *Mathematical Risk Analysis*. Springer, Berlin (2013)
45. Saida, A.B., Prigent, J.L.: On the robustness of portfolio allocation under copula misspecification. *Ann. Oper. Res.* **262**, 631–652 (2018)
46. Schultz, R., Tiedemann, S.: Conditional value-at-risk in stochastic programs with mixed-integer recourse. *Math. Program.* **105**, 365–386 (2006)
47. Sempi, C.: Convergence of copulas: a critical remark. *Rad. Mat.* **12**, 241–249 (2004)
48. Shiryaev, A.N.: *Probability*, 2nd edn. Springer, New York (1996)
49. Sklar, A.: Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris* **8**, 229–231 (1959)
50. Trutschnig, W.: On a strong metric on the space of copulas and its induced dependence measure. *J. Math. Anal. Appl.* **384**, 690–705 (2011)
51. Trutschnig, W.: Some results on the convergence of (quasi-)copulas. *Fuzzy Sets Syst.* **191**, 113–121 (2012)
52. van der Vaart, A.W.: *Asymptotic Statistics*. Cambridge University Press, Cambridge (1998)
53. Wang, S.S., Dhaene, J.: Comonotonicity, correlation order and premium principles. *Insur. Math. Econ.* **22**, 235–242 (1998)
54. Zolotarev, V.M.: Probability metrics. *Theory Probab. Appl.* **28**, 278–302 (1983)

© The Author(s) 2022. This work is published under

<http://creativecommons.org/licenses/by/4.0/>

(the “License”). Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License.