An optimal transport approach to Value-at-Risk bounds with partial dependence information

Thibaut Lux [a,1,*], Ludger Rüschendorf [b,2,*]

ABSTRACT

We establish a method to compute Value-at-Risk estimates for aggregations of multiple risks in the presence of model ambiguity. We consider the situation of dependence uncertainty, where the univariate marginal distributions of individual risk factors are known but at most partial information about the dependence structure between them is available. The partial dependence information is introduced in the form of pointwise upper and lower bounds on the joint distribution function of the risks. This extends the well-known improved standard bounds on Value-at-Risk obtained by Embrechts and Puccetti [8] and Embrechts et al. [9] so as to account for two-sided bounds on the distribution. Using the Monge-Kantorovich duality theory, we first derive sharp dual bounds on the Value-at-Risk of aggregations over the class of distributions which comply with the available information. Typically, these dual bounds are both, numerically and analytically intractable. Therefore we develop a computable optimization scheme based on the dual characterization, which allows us to derive estimates on Value-at-Risk using partial dependence information. Moreover, we show that the scheme yields asymptotically sharp risk bounds in the certainty limit. In general however, the scheme produces Value-at-Risk bounds that are not sharp over the constrained class of distributions but numerical applications demonstrate that the estimates may be significantly narrower than the improved standard bounds.

KEYWORDS: Value-at-Risk bounds, dependence uncertainty, copulas, Fréchet-Hoeffding bounds, optimal transport, duality

AUTHORS INFO

[a] *Institute of Mathematics, TU Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany*
[b] *Institute of Mathematics, University of Freiburg, Eckerstraße 1, 79104 Freiburg, Germany*
[1] thibaut.lux@mail.tu-berlin.de
[2] ruschen@stochastik.uni-freiburg.de

* *We thank Peter Bank and Ruodu Wang, for useful discussions during the work on these topics. TL gratefully acknowledges the financial support from the DFG Research Training Group 1845 "Stochastic Analysis with Applications in Biology, Finance and Physics".*

PAPER INFO

AMS CLASSIFICATION: *91B30, 62E17, 62H05*

1. Introduction

In order to increase the resilience of the financial sector to periods of heightened stress, new regulatory provisions require the computation of robust risk estimates as a key ingredient in the determination of capital reserves. This in turn calls for new methods to compute risk estimates with partial information about the distribution of the underlying risk drivers. While the computation of portfolio risk estimates from a given model (distribution) for the risk factors poses primarily computational difficulties, fundamentally different challenges arise once we scrutinize the assumption of completely specified model for the risks and consider the framework of model ambiguity. Ambiguity in the Knightian sense occurs whenever uncertainty about the joint law of the risks is introduced. In practice, such uncertainty may stem e.g. from a shortage of historical data to estimate the possibly high-dimensional distribution of the risk factors, or from risk dynamics that are too volatile as to be described adequately by a static model. In this situation, practitioners face the challenge to quantify the portfolio risk in the absence a completely specified model for the underlying factors. This typically involves the computation of worst-case estimates that correspond to the maximal risk over all possible models that are compatible with reliable information or certain views about the risk factors. These estimates are considered robust with respect to the class of admissible models. In this paper, we present a novel approach to compute robust Value-at-Risk (VaR) estimates for portfolios in the presence of model uncertainty. Specifically, we assume only partial information about the distribution of the risk factors that is available e.g. in the form of reliable estimates or expert views. In applications, we show that our approach provides ways to compute model risk estimates that comply with the five fundamental criteria for robust scenario aggregation presented in Cambou and Filipović [6], namely (1) no penalty for conservative internal models (2) focus on tail loss (3) control over distance from internal model (4) robustness of capital requirements and (5) tractability.

In our framework, we consider an $\mathbb{R}^d$-valued risk vector $\mathbf{X} = (X_1, \dots, X_d)$ with partially specified distribution and an aggregation function $\psi\colon \mathbb{R}^d \to \mathbb{R}$. Specifically, we focus on the situation of *dependence uncertainty*, where the marginal distributions of the components $X_i \sim F_i$ are known or can be inferred from data, while the dependence structure of $\mathbf{X}$ is only partially known.

A significant part of the literature on dependence uncertainty focuses on the marginals-only case, where merely the marginal distributions of $\mathbf{X}$ are known and no information at all about the dependence structure between its constituents is available. In this case, sharp VaR bounds can be obtained by the Rearrangement Algorithm (RA) introduced in Puccetti and Rüschendorf [18] and Embrechts, Puccetti, and Rüschendorf [10]. For a presentation of general results in the marginals-only case see also [10]. The complete absence of information on the dependence structure however leads typically to very wide risk bounds that are not sufficiently informative for practical applications.

This observation has led to a series of papers discussing VaR bounds with additional dependence information in the form of a one-sided, upper or lower bound on the joint distribution function of the risks. The associated VaR bounds are in the literature referred to as *improved standard bounds*. For the ample literature on this we refer to Williamson and Downs [31], Denuit, Genest, and Marceau [7], Puccetti and Rüschendorf [19], Bignozzi, Puccetti, and Rüschendorf [4], Bernard and Vanduffel [3], Puccetti, Rüschendorf, and Manko [21] and Lux and Papapantoleon [16]. A survey on this and improved VaR bounds using partial dependence information is provided in Rüschendorf [25]. As a result it has been found that this kind of one-sided information leads to reasonably narrow risk estimates when the distributional bounds inscribe strong (positive or negative) dependence information or when the dimension d is relatively small. In higher dimensions the bounds remain typically too wide as to be of practical relevance since they yield no substantial improvement over the marginals-only case.

To address this problem, we consider in this paper the situation of dependence uncertainty with additional two-sided constraints on the joint distribution function F of $\mathbf{X}$, i.e. we assume that a lower and an upper bound on F respectively the copula C of $\mathbf{X}$ are given. We then make use of the Kantorovich Duality Theorem in order to derive a sharp dual representation of the VaR bounds over the constrained class of distributions. This extends the results in Embrechts and Puccetti [8] in two ways: (i) we show that the dual bounds are sharp under mild conditions on the aggregation function ψ and (ii) we include two-sided information on the copula of the risk vector. The dual bounds however are both analytically and numerically intractable. Nevertheless, the dual formulation allows us to obtain a reduced and tractable optimization scheme for the computation of VaR bounds with two-sided dependence information. Our scheme corresponds to an optimization over a tractable subset of admissible functions for the duals, hence the resulting VaR bounds are no longer sharp in general. Nevertheless, we show that our reduced scheme yields asymptotically sharp risk bounds in the certainty limit. Moreover we illustrate in numerical examples that our method produces reasonable results in higher dimensions. In particular, our VaR estimates are significantly tighter than the improved standard bounds.

2. Bounds on Value-at-Risk using copula information

In this paper we consider an $\mathbb{R}^d$-valued random vector of risks $\mathbf{X} = (X_1, \dots, X_d)$ and an aggregation function $\psi\colon \mathbb{R}^d \to \mathbb{R}$. We want to compute the VaR of the aggregation $\psi(\mathbf{X})$. The VaR of $\psi(\mathbf{X})$ relates to the quantile function in the following way: when $\psi(\mathbf{X}) \sim F$ then the VaR of $\psi(\mathbf{X})$ for a certain confidence level $\alpha \in (0,1)$ is given by the quantity

$$\mathrm{VaR}_\alpha(\psi(\mathbf{X})) = F^{-1}(\alpha) = \inf\{x \in \mathbb{R}\colon F^-(x) > \alpha\}.$$

Here we define VaR in terms of the left-continuous quantile function $F^-_{\psi(\mathbf{X})}$, while in the literature it is sometimes defined in terms of $F_{\psi(\mathbf{X})}$. Typical levels of α are close to 1 in practice, assuming that risks (or losses) correspond to the right tail of the distribution.The most commonly considered aggregation function ψ is the sum of the individual risks $X_1 + \cdots + X_d$, but also the maximum and minimum of the risks, $\max\{X_1, \dots, X_d\}$ and $\min\{X_1, \dots, X_d\}$ are of interest in applications.

We are concerned with the situation of model ambiguity and assume that only partial information about the distribution of $\mathbf{X}$ is available. In practice, the univariate distributions of the constituents $X_1, ..., X_d$ are often assumed to be known or can be estimated rather accurately while the dependence structure between the individual components is at best partially known. This form of model ambiguity is referred to as *dependence uncertainty*. Specifically, it is assumed that the unknown joint distribution F of $\mathbf{X}$ is in the Fréchet class $\mathcal{F}(F_1, ..., F_d)$ of d-dimensional distribution functions with marginals $F_1, ..., F_d$. Then it follows from Sklar's Theorem that F can be expressed as

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad \text{for all } (x_1, ..., x_d) \in \mathbb{R}^d, \tag{2.1}$$

for a copula C. This implies that dependence uncertainty is in fact uncertainty about the copula of $\mathbf{X}$. For an introduction to the theory of copulas see e.g. Nelsen [17] and for related applications in risk management c.f. Bouyé, Durrleman, Nikeghbali, Riboulet, and Roncalli [5].

When no information about the risk vector besides its marginal distributions is provided, then every copula yields a possible joint distribution for $\mathbf{X}$ via (2.1). In this situation, VaR bounds for aggregations $\psi(\mathbf{X})$ over the set of all copulas are typically too wide to be relevant in practice. In addition, it is rarely the case that no information at all about the dependence structure of the risk vector is available, since partial information such as correlations between the risk factors can be estimated or inferred with sufficient accuracy. Such additional information about $\psi(\mathbf{X})$ can be translated into a lower and an upper bound on the copula of $\mathbf{X}$. The derivation of bounds on copulas that improve the generic Fréchet–Hoeffding bounds when partial dependence information is available has recently attracted attention and a series of papers establishing improved Fréchet–Hoeffding bounds for many kinds of partial information has emerged; see e.g. Nelsen [17] and Tankov [27] for $d = 2$ or Lux and Papapantoleon [15, 16] and Puccetti et al. [21] for $d > 2$. Our approach will allow us to translate these improved Fréchet–Hoeffding bounds into bounds on VaR. In fact, we consider first a much more general problem. Instead of bounding VaR we derive bounds on the expectation of a more general functional $\varphi(\mathbf{X})$ when bounds on the copula of $\mathbf{X}$ are provided.

Let us denote the expectation operator for a measurable $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ and an $\mathbb{R}^d$-valued random vector $\mathbf{X}$ with marginal distributions $F_1, .., F_d$ and copula C, by

$$\mathbb{E}_C[\varphi] = \int_{\mathbb{R}^d} \varphi(x_1, ..., x_d)\, \mathrm{d}C(F_1(x_1), ..., F_d(x_d)).$$

Using $C \leq C'$ to refer to the pointwise inequality between d-variate functions, we formulate the following primal problems:

$$\underline{P}_\varphi := \inf \left\{\mathbb{E}_C[\varphi] \colon C \in \mathcal{C}^d, \underline{Q} \leq C \leq \overline{Q}\right\}, \tag{2.2}$$
$$\overline{P}_\varphi := \sup \left\{\mathbb{E}_C[\varphi] \colon C \in \mathcal{C}^d, \underline{Q} \leq C \leq \overline{Q}\right\}, \tag{2.3}$$

where $\mathcal{C}^d$ is the set of all d-copulas and $\overline{Q}, \underline{Q}$ are quasi-copulas with $\underline{Q} \leq \overline{Q}$. The notion of quasi-copulas generalizes the copula concept as follows:

Definition 2.1. A function $Q\colon [0, 1]^d \to [0, 1]$ is a *d-quasi-copula* if the following properties hold:

($\mathbf{QC1}$) Q satisfies, for all $i \in \{1, \ldots, d\}$, the boundary conditions

$$Q(u_1, \ldots, u_i = 0, \ldots, u_d) = 0 \quad \text{and} \quad Q(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i.$$

($\mathbf{QC2}$) Q is increasing in each argument.

($\mathbf{QC3}$) Q is Lipschitz continuous, i.e. for all $\mathbf{u}, \mathbf{v} \in [0, 1]^d$

$$|Q(u_1, \ldots, u_d) - Q(v_1, \ldots, v_d)| \leq \sum_{i=1}^{d} |u_i - v_i|.$$

In particular, by allowing quasi-copulas as bounds in the formulation of the primal problems, we are able to include the lower Fréchet–Hoeffding bound or improved Fréchet–Hoeffding bounds which are quasi-copulas but often fail to be proper copulas; c.f. Lux and Papapantoleon [15]. Allowing quasi-copulas as bounds, there might however not exist a copula that complies with the constraints and so $\mathcal{C}_b := \{C \in \mathcal{C}^d : \underline{Q} \le C \le \overline{Q}\} = \emptyset$. Consider e.g. $\overline{Q} = \underline{Q} = W_d$, then $\mathcal{C}_b$ is empty whenever $d > 2$. In this case, we set $\overline{P}_\varphi = \infty$ and $\underline{P}_\varphi = -\infty$.

Remark 2.2. When $\underline{Q}$ and $\overline{Q}$ are equal to the lower and upper Fréchet–Hoeffding bound respectively, i.e.

$$\begin{aligned} \underline{Q}(u_1, ..., u_d) &= \max\left\{0, \sum_{i=1}^{d} u_i - d + 1\right\} =: W_d(\mathbf{u}) \quad \text{and} \\ \overline{Q}(u_1, ..., u_d) &= \min(u_1, ..., u_d) =: M_d \end{aligned} \tag{2.4}$$

then the optimization corresponds to a standard Fréchet problem where only the marginals are known and no information about the dependence structure at all is prescribed. ♦

In order to compute or approximate the bounds $\overline{P}_\varphi$ and $\underline{P}_\varphi$ we proceed as follows: First, we derive a dual characterization of the the primal problems yielding sharp bounds on the expectation of $\varphi(\mathbf{X})$ under rather general assumptions on the function φ. Based on the dual characterization we then develop a tractable optimization scheme to compute bounds on the expectation for specific functions φ. In particular, the scheme allows us to determine robust VaR estimates of aggregations using copula bounds.

3. Dual Characterization of Value-at-Risk Bounds with additional Information

In this section we establish a dual characterization of the primal problems $\overline{P}_\varphi$ and $\underline{P}_\varphi$ and prove strong duality between the two formulations. To this end, let us introduce the class

$$\mathcal{R} := \left\{ h = \sum_{n=1}^{k} \alpha_n \Lambda_{\mathbf{u}^n} : k \in \mathbb{N}, \alpha_1, ..., \alpha_k \ge 0, \mathbf{u}^1, ..., \mathbf{u}^k \in \overline{\mathbb{R}}^d \right\},$$

where the functions $\Lambda_{\mathbf{u}}$ are of the form

$$\Lambda_{\mathbf{u}} : \mathbb{R}^d \ni (x_1, ..., x_d) \mapsto \mathbb{1}_{x_1 \le u_1, ..., x_d \le u_d}.$$

The elements in $\mathcal{R}$ are hence positive, linear combinations of indicator functions of rectangles of the form $(-\infty, u_1] \times \cdots \times (-\infty, u_d]$. Analogously, we denote the lower semicontinuous version of $\Lambda_{\mathbf{u}}$ by

$$\Lambda^-_{\mathbf{u}} : (x_1, ..., x_d) \mapsto \mathbb{1}_{x_1 < u_1, ..., x_d < u_d},$$

and for $h = \sum_{n=1}^{k} \alpha_n \Lambda_{\mathbf{u}^n} \in \mathcal{R}$ we define $h^- := \sum_{n=1}^{k} \alpha_n \Lambda^-_{\mathbf{u}^n}$.

Note that for a copula C it holds that

$$\mathbb{E}_C[\Lambda_{\mathbf{u}}] = \int_{\mathbb{R}^d} \Lambda_{\mathbf{u}}(x_1, ..., x_d)\, \mathrm{d}C\big(F_1(x_1), ..., F_d(x_d)\big) = C\big(F_d(x_1), ..., F_d(x_d)\big), \tag{3.1}$$

and analogously we obtain that $\mathbb{E}_C[\Lambda_{\mathbf{u}}^-] = C(F_d^-(x_1), ..., F_d^-(x_d))$, where F_i^- is the left-continuous version of F_i for $i = 1, ..., n$.

Moreover, let us define, for a quasi-copula Q and a function $h = \sum_{n=1}^k \alpha_n \Lambda_{\mathbf{u}^n} \in \mathcal{R}$,

$$Q(h) := \sum_{n=1}^k \alpha_n Q\big(F_1(u_1^n), ..., F_d(u_d^n)\big); \quad Q(h^-) := \sum_{n=1}^k \alpha_n Q\big(F_1^-(u_1^n), ..., F_d^-(u_d^n)\big).$$

If $Q = C$ for a copula C, we have that $Q(h) = \mathbb{E}_Q[h]$ as well as $Q(h^-) = \mathbb{E}_Q[h^-]$.

This leads us to the dual problem corresponding to $\underline{P}_\varphi$:

$$\begin{aligned}\underline{D}_\varphi = \sup \Big\{ \underline{Q}(h) - \overline{Q}(g^-) + \sum_{i=1}^d \mathbb{E}_i[\nu_i] \colon \nu_i \in \mathcal{L}(F_i), i = 1, ..., d;& \\ h, g \in \mathcal{R} \text{ s.t. } h - g^- + \sum_{i=1}^d \nu_i \le \varphi \Big\},&\end{aligned} \tag{3.2}$$

where $\mathbb{E}_i[\nu_i] = \int \nu_i \, \mathrm{d}F_i$ and $\mathcal{L}(F_i)$ is the class of F_i-integrable functions ν_i, i.e. $\mathbb{E}_i[\nu_i] < \infty$, for $i = 1, ..., d$. Analogously, for the upper bound $\overline{P}_\varphi$, the corresponding dual is given by

$$\begin{aligned}\overline{D}_\varphi = \inf \Big\{ \overline{Q}(h^-) - \underline{Q}(g) + \sum_{i=1}^d \mathbb{E}_i[\nu_i] \colon \nu_i \in \mathcal{L}(F_i), i = 1, ..., d;& \\ h, g \in \mathcal{R} \text{ s.t. } h^- - g + \sum_{i=1}^d \nu_i \ge \varphi \Big\}.&\end{aligned} \tag{3.3}$$

Note, that the roles of $\underline{Q}$ and $\overline{Q}$ are reversed in $\underline{D}_\varphi$, i.e. we subtract the sum w.r.t. $\overline{Q}$ from the sum w.r.t. $\underline{Q}$ in the formulation of $\underline{D}_\varphi$ and vice versa for $\overline{D}_\varphi$. In the remainder of this section, we show that strong duality between the primal and the dual problem holds under mild assumptions on the function φ, so that:

$$\underline{P}_\varphi = \underline{D}_\varphi \quad \text{and} \quad \overline{P}_\varphi = \overline{D}_\varphi.$$

Several approaches to proving duality results of this type have been established in the literature. For instance, Rüschendorf [23] and Rüschendorf and Gaffke [22] establish duality results for functionals of multivariate random variables with given marginals using a Hahn-Banach separation argument. More recently, a duality result for the martingale optimal transport problem was established by Beiglböck, Henry-Labordère, and Penkner [2] using a minmax argument, and Bartl, Cheredito, Kupper, and Tangpi [1] derive a general duality result for convex functionals with countably many marginal constraints using the Daniell-Stone Theorem. An account of the history of the Monge-Kantorovich duality theory and associated references can be found in the survey by Rüschendorf [24] or in the book by Villani [29].

The proof of our Duality Theorem 3.4 below is based on the following auxiliary results.

Lemma 3.1 (Kantorovich Duality for copulas). *Let $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ be lower semicontinuous and such that*

$$\sum_{i=1}^{d} \varrho_i(x_i) \geq |\varphi(x_1, ..., x_d)| \quad \text{for all } (x_1, ..., x_d) \in \mathbb{R}^d \tag{3.4}$$

for some continuous functions $\varrho_i \in \mathcal{L}(F_i)$, $i = 1, ..., d$. Then the following duality holds

$$\inf\left\{\mathbb{E}_C[\varphi]\colon C \in \mathcal{C}^d\right\} = \sup\left\{\sum_{i=1}^{d} \mathbb{E}_i[\nu_i]\colon \nu_i \in \mathcal{L}(F_i),\ i = 1, ..., d,\ \sum_{i=1}^{d} \nu_i \leq \varphi\right\} \tag{3.5}$$

$$\tag{3.6}$$

Proof. The statement follows immediately by an application of Sklar's Theorem. In particular, we have that

$$\inf\left\{\mathbb{E}_C[\varphi]\colon C \in \mathcal{C}^d\right\} = \inf\left\{\int_{\mathbb{R}^d} \varphi(\mathbf{x})\,\mathrm{d}F(\mathbf{x})\colon F \in \mathcal{F}(F_1, ..., F_d)\right\}.$$

Then, due to (3.4) we can apply the classical Kantorovich Duality Theorem (see e.g. Theorem 5.10 in [29]) to the right-hand side of the equation and obtain (3.5). □

The following Lemma links the uniform convergence of copulas to the weak convergence of the associated random vectors. The proof is an immediate consequence of Theorem 2.1 in Lindner and Szimayer [13] and therefore omitted.

Lemma 3.2. *Let $(X_1^n, ..., X_d^n)_{n\in\mathbb{N}}$ be a sequence of $\mathbb{R}^d$-valued random vectors with marginals $X_i^n \sim F_i$ for $i = 1, ..., d$ and all $n \in \mathbb{N}$. Moreover, denote by C_n the copula of $(X_1^n, ..., X_d^n)$ for $n \in \mathbb{N}$. If the sequence of copulas C_n converges uniformly to a copula C then $(X_1^n, ..., X_d^n)$ converges weakly to a random vector $(X_1, ..., X_d)$ with marginals $F_1, ..., F_d$ and copula C.*

The following Minmax Theorem is presented as Corollary 2 in Terkelsen [28].

Lemma 3.3 (Minmax Theorem). *Let B_1 be a compact convex subset of a topological vector space V_1 and B_2 be a convex subset of a vector space V_2. If $f\colon B_1 \times B_2 \to \mathbb{R}$ is such that*

1. *$f(\cdot, b_2)$ is lower semicontinuous and convex on B_1 for all $b_2 \in B_2$,*
2. *$f(b_1, \cdot)$ is concave on B_2 for all $b_1 \in B_1$,*

then

$$\inf_{b_1\in B_1} \sup_{b_2\in B_2} f(b_1, b_2) = \sup_{b_2\in B_2} \inf_{b_1\in B_1} f(b_1, b_2).$$

With these result, we are now in the position to establish our main Duality Theorem.

Theorem 3.4. *Let* $\varphi\colon \mathbb{R}^d \to \mathbb{R}$ *be such that*

$$\sum_{i=1}^{d} \varrho_i(x_i) \geq |\varphi(x_1, ..., x_d)| \quad \textit{for all } (x_1, ..., x_d) \in \mathbb{R}^d, \tag{3.7}$$

for some continuous functions $\varrho_i \in \mathcal{L}(F_i)$, $i = 1, ..., d$. *Moreover, assume that there exists a copula* $C \in \mathcal{C}^d$ *with* $\underline{Q} \leq C \leq \overline{Q}$. *Then if* φ *is lower semicontinuous the following duality holds:*

$$\underline{P}_\varphi = \underline{D}_\varphi.$$

Conversely, when φ *is upper semicontinuous the following duality holds:*

$$\overline{P}_\varphi = \overline{D}_\varphi.$$

Moreover, the primal values are attained, i.e. there exist copulas $\underline{C}, \overline{C}$ *such that* $\mathbb{E}_{\underline{C}}[\varphi] = \underline{P}_\varphi$ *and* $\mathbb{E}_{\overline{C}}[\varphi] = \overline{P}_\varphi$.

Proof. We show that the statement holds for the lower bound, i.e. $\underline{D}_\varphi = \underline{P}_\varphi$. The proof for the upper bound can be derived by applying analogous arguments to the function $-\varphi$.

First, assume that φ is bounded and continuous. By ν_i we refer to functions in $\mathcal{L}(F_i)$. It follows that

$$\overline{D}_\varphi = \sup_{h,g \in \mathcal{R}} \sup_{\substack{\nu_1,...,\nu_d \\ h - g^- + \sum_{i=1}^d \nu_i \leq \varphi}} \left\{ \underline{Q}(h) - \overline{Q}(g^-) + \sum_{i=1}^{d} \mathbb{E}_i[\nu_i] \right\} \tag{3.8}$$

$$= \sup_{h,g \in \mathcal{R}} \sup_{\substack{\nu_1,...,\nu_d \\ \sum_{i=1}^d \nu_i \leq \varphi - h + g^-}} \left\{ \underline{Q}(h) - \overline{Q}(g^-) + \sum_{i=1}^{d} \mathbb{E}_i[\nu_i] \right\} \tag{3.9}$$

$$= \sup_{h,g \in \mathcal{R}} \inf_{C \in \mathcal{C}^d} \left\{ \underline{Q}(h) - \overline{Q}(g^-) + \mathbb{E}_C[\varphi - h + g^-] \right\} \tag{3.10}$$

$$= \sup_{h,g \in \mathcal{R}} \inf_{C \in \mathcal{C}^d} \left\{ \big(\underline{Q}(h) - C(h)\big) - \big(\overline{Q}(g^-) - C(g^-)\big) + \mathbb{E}_C[\varphi] \right\} \tag{3.11}$$

$$= \inf_{C \in \mathcal{C}^d} \left\{ \sup_{h,g \in \mathcal{R}} \left\{ \big(\underline{Q}(h) - C(h)\big) - \big(\overline{Q}(g^-) - C(g^-)\big) \right\} + \mathbb{E}_C[\varphi] \right\} \tag{3.12}$$

$$= \inf_{\substack{C \in \mathcal{C}^d \\ \underline{Q}(h) \leq C(h) \leq \overline{Q}(h),\ \forall h \in \mathcal{R}}} \mathbb{E}_C[\varphi] \tag{3.13}$$

$$= \inf_{\underline{Q} \leq C \leq \overline{Q}} \mathbb{E}_C[\varphi] = \underline{P}_\varphi. \tag{3.14}$$

Equation (3.10) follows from an application of Lemma 3.1 to the function $\varphi' := \varphi - h + g^-$. Note, that the application of Lemma 3.1 is justified since φ' is lower semicontinuous being the sum of the lower semicontinuous functions φ, $-h$ and g^-. Moreover, since h and g are of the form

$$h = \sum_{n=1}^{k} \alpha_n \Lambda_{\mathbf{u}^n}, \quad g = \sum_{n=1}^{m} \beta_n \Lambda_{\mathbf{v}^n}$$

for $\alpha_1, ..., \alpha_k, \beta_1, ..., \beta_m \in \mathbb{R}_+$, we obtain

$$|(\varphi + h - g^-)(x_1, ..., x_d)| \leq \sum_{i=1}^{d} \varrho_i(x_i) + \sum_{n=1}^{k} \alpha_n + \sum_{n=1}^{m} \beta_n.$$

Equation (3.11) then follows by rearranging the terms, using the linearity of the expectation and the definition of the operator $C(h)$ for $h \in \mathcal{R}$. Now, applying the Minmax Theorem 3.3 to the function

$$f : \mathcal{C}^d \times \mathcal{R}^2 \ni \big(C, (h, g)\big) \mapsto \big(\underline{Q}(h) - C(h)\big) - \big(\overline{Q}(g^-) - C(g^-)\big) + \mathbb{E}_C[\varphi]$$

yields equation (3.12). Note, that the requirements of Theorem 3.3 are satisfied, since

$$\mathcal{C}_b = \big\{C \in \mathcal{C}^d : \underline{Q} \leq C \leq \overline{Q}\big\}$$

is a closed, bounded and equicontinuous subset of the topological space of all continuous functions on $[0, 1]^d$, equipped with the uniform metric. Hence, it follows from the Arzelà-Ascoli Theorem that $\mathcal{C}_b$ is compact. Moreover, $\mathcal{C}_b$ and $\mathcal{R}^2$ are convex sets. On the other hand, it follows from Lemma 3.2 that for all $h, g \in \mathcal{R}$ the map $f\big(\cdot\,, (h, g)\big)$ is continuous w.r.t. the uniform convergence of copulas since we assume φ to be bounded and continuous. Furthermore, we have that $f\big(\cdot\,, (h, g)\big)$ is convex on $\mathcal{C}^d$. Also, for all $C \in \mathcal{C}^d$ it holds that $f(C, \cdot)$ is linear on $\mathcal{R}^2$. To verify (3.13), assume that $\underline{Q}(h) \leq C(h) \leq \overline{Q}(h)$ does not hold for one $h \in \mathcal{R}$, i.e. let w.l.o.g. $C(h) < \underline{Q}(h)$, then for each $\alpha > 0$ it follows that

$$\big(\underline{Q}(\alpha h) - C(\alpha h)\big) = \alpha\big(\underline{Q}(h) - C(h)\big) > 0$$

and thus, by scaling α, the supremum is ∞ and C can be disregarded in the infimum in (3.12). Hence, it holds that

$$\underline{Q}(h) \leq C(h) \leq \overline{Q}(h), \text{for all } h \in \mathcal{R}.$$

This entails that $\underline{Q}(h) - C(h) \leq 0$ and $-(\overline{Q}(g) - C(g)) \leq 0$ for all $(g, h) \in \mathcal{R}^2$ and thus the supremum is attained for $h, g \equiv 0$. Finally, (3.14) holds due to the fact that $\underline{Q}(h) \leq C(h) \leq \overline{Q}(h)$ for all $h \in \mathcal{R}$ implies

$$\underline{Q}(F_1(x_1), ..., F_d(x_d)) \leq C(F_1(x_1), ..., F_d(x_d)) \leq \overline{Q}(F_1(x_1), ..., F_d(x_d))$$

for all $(x_1, ..., x_d) \in \mathbb{R}^d$ and $\mathbb{E}_C[\varphi] = \mathbb{E}_{C'}[\varphi]$ for all copulas C and C' with $C(F_1(x_1), ..., F_d(x_d)) = C'(F_1(x_1), ..., F_d(x_d))$.

We proceed by relaxing the condition of φ being bounded and continuous. So let φ merely be lower semicontinuous. We can w.l.o.g. assume that $\varphi \geq 0$ as otherwise there exist, due to condition (3.7), functions $\varrho_1, ...\varrho_d$ with $\varphi + \sum_{i=1}^d \varrho_i \geq 0$ and

$$\overline{P}_\varphi = \overline{P}_{\varphi+\sum_{i=1}^d \varrho_i} - \sum_{i=1}^d \mathbb{E}_i[\varrho_i].$$

Now, since φ is lower semicontinuous there exists a sequence of positive, bounded, continuous functions $\varphi_1 \leq \varphi_2 \leq \cdots$ with $\varphi = \lim_n \varphi_n$ pointwise and $\underline{P}_{\varphi_n} \leq \underline{P}_\varphi$. Furthermore, due to the compactness of $\mathcal{C}_b$ there exist optimizers $C_1, C_2, ...$ of $\underline{P}_{\varphi_1}, \underline{P}_{\varphi_2}, ...$ and we can, by passing to a subsequence, assume that $C_1, C_2, ...$ converges to some $C^* \in \mathcal{C}_b$. Then it follows by monotone convergence that

$$\underline{P}_\varphi \leq \mathbb{E}_{C^*}[\varphi] = \lim_n \mathbb{E}_{C^*}[\varphi_n] = \lim_n \lim_j \mathbb{E}_{C_j}[\varphi_n] \leq \lim_j \mathbb{E}_{C_j}[\varphi_j] = \lim_j \underline{P}_{\varphi_j} = \lim_j \underline{D}_{\varphi_j}$$

Lastly, we note that the optimizers of the primal problems are attained due to the compactness of $\mathcal{C}_b$ which completes the proof. □

Remark 3.5. Assuming the existence of a copula $C \in \mathcal{C}^d$ with $\underline{Q} \leq C \leq \overline{Q}$ in Theorem 3.4 rules out the degenerate situation where no probabilistic model exists which is compatible with the prescribed information. Verifying this assumption however is a delicate task in general. The existence of a copula C with $\underline{Q} \leq C$ follows immediately from the fact that the upper Fréchet–Hoeffding bound M_d is a copula and hence $\underline{Q} \leq M_d$. The difficulty thus lies in verifying $C \leq \overline{Q}$ which fails e.g. when $\overline{Q} = W_d$ and $d > 2$, where W_d is the lower Fréchet–Hoeffding bound given in (2.4). Nevertheless, when $\underline{Q}$ and $\overline{Q}$ are improved Fréchet–Hoeffding bounds it is often straight-forward to verify that $\{C \in \mathcal{C}^d : \underline{Q} \leq C \leq \overline{Q}\}$ is not empty. ♦

The following counter-example shows that the dual optimizers are not attained in general.

Example 3.6. Consider the case $d = 2$ and let F_1 and F_2 be uniform marginal distribution on $[0, 1]$. Moreover, let $\underline{Q}(u_1, u_2) = \overline{Q}(u_1, u_2) = \Pi(u_1, u_2) = u_1 u_2$ for all $(u_1, u_2) \in [0, 1]^2$ and consider $\varphi : \mathbb{R}^2 \to \mathbb{R} : (u_1, u_2) \mapsto \mathbb{1}_{\psi(u_1,u_2)<1}$ where $\psi(u_1, u_2) = \sqrt{u_1^2 + u_2^2}$, i.e. φ is the characteristic function of the circular segment of the unit circle on $[0, 1]^2$. It then follows from $\underline{Q} = \overline{Q} = \Pi$, that

$$\underline{P}_\varphi = \overline{P}_\varphi = \int_{[0,1]^2} \mathbb{1}_{\sqrt{u_1^2+u_2^2}<1} \, \mathrm{d}u_1 \mathrm{d}u_2 = \frac{\pi}{4}.$$

Now, assume the dual optimizer for $\underline{D}_\varphi$ is attained. Then it is of the form

$$f^* := h - g + \nu_1 + \nu_2 = \sum_{n=1}^k \alpha_n \Lambda_{\mathbf{u}^n} + \sum_{n=1}^m -\beta_n \Lambda_{\mathbf{v}^n} + \nu_1 + \nu_2$$

and since $f^*(u_1, u_2) \leq \mathbb{1}_{\psi(u_1,u_2)<1}$ for all $(u_1, u_2) \in [0, 1]^2$ and $\mathbb{E}_\Pi[f^*] = \frac{\pi}{4}$, we have that

$$f^*(u_1, u_2) = (h - g + \nu_1 + \nu_2)(u_1, u_2) = \mathbb{1}_{\psi(u_1,u_2)<1} \quad \lambda\text{-a.s.} \tag{3.15}$$

Moreover, we can assume w.l.o.g. that $\nu_1 \equiv \nu_2 \equiv 0$ λ-a.s. since it follows from equation (3.15) that

$$(h-g)(u_1,1) = \mathbb{1}_{\psi(u_1,1)<1} - \nu_1(u_1) - \nu_2(1) = -\nu_1(u_1) - c \quad \lambda\text{-a.s.}$$

where the last inequality is due to $\mathbb{1}_{\psi(u_1,1)<1} = 0$ λ-a.s. and $\nu_2(1) =: c$. Now, by the same argument it follows that

$$(h-g)(1,u_2) = -\nu_2(u_2) - c' \quad \lambda\text{-a.s.},$$

and thus we obtain

$$(h-g)(u_1,u_d) = \Big(\sum_{n=1}^{k} \alpha_n \Lambda_{\mathbf{u}^n} + \sum_{n=1}^{m} -\beta_n \Lambda_{\mathbf{v}^n}\Big)(u_1,u_2) = \mathbb{1}_{\psi(u_1,u_2)<1} \quad \lambda\text{-a.s.}$$

This however, corresponds to a construction of the characteristic function of the circular segment by a finite number of rectangular characteristic functions which contradicts the impossibility of the squaring of the circle. $\Diamond$

4. A reduction scheme to compute bounds on Value-at-Risk

The dual characterizations of $\underline{P}_\varphi$ and $\overline{P}_\varphi$ in Section 3 lend themselves to the development of a scheme to compute VaR estimates, accounting for an upper and a lower bound on the copula of the risks. In general, the dual problems are intractable and closed form solutions have only been obtained in the situation where $\underline{Q} = W_d$, $\overline{Q} = M_d$ with homogeneous marginals $F_1 = \cdots = F_d$ fulfilling additional constraints; c.f. Puccetti and Rüschendorf [20] and Wang and Wang [30]. We therefore develop in this section a scheme that corresponds to an optimization over a tractable subset of admissible functions for the duals $\underline{D}_\varphi$ and $\overline{D}_\varphi$ and produces narrow VaR bounds. Furthermore, we show that scheme produces asymptotically sharp bounds in the certainty limit, i.e. when $\underline{Q}$ and $\overline{Q}$ converge to some copula C.

4.1. A reduction scheme for $\underline{D}_\varphi$

Consider the function $\varphi(x_1,...,x_d) = \mathbb{1}_{\psi(x_1,...,x_d)<s}$ for componentwise increasing $\psi\colon \mathbb{R}^d \to \mathbb{R}$ and recall from (2.2) that our primal problem of interest reads

$$\underline{P}_\varphi := \inf\left\{\mathbb{E}_C[\varphi] \colon C \in \mathcal{C}^d, \underline{Q} \le C \le \overline{Q}\right\},$$

for quasi-copulas $\underline{Q}$ and $\overline{Q}$, whereas the corresponding dual problem is given in (3.2) by

$$\underline{D}_\varphi = \sup\Big\{\underline{Q}(h) - \overline{Q}(g^-) + \sum_{i=1}^{d} \mathbb{E}_i[\nu_i] \colon \nu_i \in \mathcal{L}(F_i),\ i = 1,...,d,$$
$$h, g \in \mathcal{R} \text{ s.t. } h - g^- + \sum_{i=1}^{d} \nu_i \le \varphi\Big\}.$$

In the following, we identify admissible functions for the dual $\underline{D}_\varphi$ by $(d+2)$-tuples in the class

$$\underline{\mathcal{A}} := \left\{(h, g, \nu_1, ..., \nu_d) \colon \nu_i \in \mathcal{L}(F_i),\ i = 1, ..., d,\ h, g \in \mathcal{R} \text{ s.t. } h - g^- + \sum_{i=1}^{d} \nu_i \le \varphi\right\}$$

and for each admissible tuple the corresponding value of the objective function amounts to

$$\underline{Q}(h) - \overline{Q}(g^-) + \sum_{i=1}^{d} \mathbb{E}_i[\nu_i].$$

Regarding the Improved Standard Bounds in [8, 9], we note that when the copula C of $\mathbf{X}$ is bounded from below by $\underline{Q}$, i.e. $\underline{Q} \le C$, then the lower Improved Standard Bound is given by

$$\mathbb{E}_C[\mathbb{1}_{\psi(\mathbf{X})<s}] \ge \sup_{u_1,\dots,u_{d-1}\in\mathbb{R}} \underline{Q}\big(F_1(u_1), \dots, F_{d-1}(u_{d-1}), F_d^-(\psi^*_{u_{-d}}(s))\big) = \underline{m}_{\underline{Q},\psi}(s),$$

which corresponds, in the case of continuous marginals, to the maximization of $\underline{Q}(h)$ over functions $h = \Lambda_{\mathbf{u}} \in \mathcal{R}$ with $\mathbf{u} \in \big\{\big(u_1, ..., u_{d-1}, \psi^*_{u_{-d}}(s)\big) \colon (u_1, ...u_{d-1}) \in \mathbb{R}^{d-1}\big\}$. Hence, $\underline{m}_{\underline{Q},\psi}(s)$ can be viewed as an optimization over a – rather small – subset of admissible elements in $\underline{\mathcal{A}}$, i.e. tuples of the form $(h, 0, ..., 0) \in \underline{\mathcal{A}}$.

Leveraging this observation, we develop an optimization scheme over a larger subset of admissible functions. To this end, let us first consider admissible $(h, g, \nu_1, ..., \nu_d)$ with

- $(\mathbf{A1})$ $\nu_1, ..., \nu_d \equiv 0$,
- $(\mathbf{A2})$ $h, g \in \mathcal{R}^\square$, where

$$\mathcal{R}^\square := \left\{\sum_{n=1}^{k} \Lambda_{\mathbf{u}^n} \colon k \in \mathbb{N}, \mathbf{u}_1, ..., \mathbf{u}_k \in \mathcal{U}_\psi(s)\right\},$$

and $\mathcal{U}_\psi(s) = \{(x_1, ..., x_d) \in \mathbb{R}^d \colon \psi(x_1, ..., x_d) < s\}$. We thus obtain a subset of projections of admissible functions in $\underline{\mathcal{A}}$ given by

$$\underline{\mathcal{A}}^\square := \big\{(h, g) \colon h, g \in \mathcal{R}^\square \text{ s.t. } h - g^- \le \varphi\big\}.$$

The optimization over the subset $\underline{\mathcal{A}}^\square$ remains however intractable due to the constraint $h - g^- \le \varphi$. Moreover, optimizing over $\underline{\mathcal{A}}^\square$ requires a truncation of k and an appropriate choice for such a truncation is not obvious. We therefore proceed with the development of an unconstrained optimization scheme over a finite number of elements in $\mathcal{U}_\psi(s)$. An informal description and illustration of the scheme and the idea of the proofs is provided in Section 5. For notational ease, let us introduce the notion of multisets (c.f. Definition 2 in Syropoulos [26]).

Definition 4.1. Let $\mathcal{B}$ be some set. A multiset over $\mathcal{B}$ is a pair $\langle\mathcal{B}, f\rangle$ where $f \colon \mathcal{B} \to \mathbb{N}$ and f is called *multiplicity function*.

Multisets generalize the notion of a set so as to allow for finite but multiple occurrences of elements. The following example illustrates this feature.

Example 4.2. By the conventional notion of a set we have that $\mathcal{B} := \{1, 1, 2\} = \{1, 2\}$. Using the notion multisets we refer to $\{1, 1, 2\}$ as $\langle \mathcal{B}, f \rangle$ with $f(1) = 2$ and $f(2) = 1$. The multiplicity function f hence counts the number of occurences of each element of $\mathcal{B}$. ◊

Our scheme is based on the following Inclusion-Exclusion Principle for multisets.

Lemma 4.3 (Multiset Inclusion-Exclusion Principle). *Let $B_1, ..., B_k \subset \mathbb{R}^d$ and define for $m = 1, ..., k$ the multisets*

$$\begin{aligned} &\langle \mathcal{B}^o, f^o \rangle, \quad \mathcal{B}^o := \{B_{i_1} \cap \cdots \cap B_{i_m} : 0 \le i_1 \le \cdots \le i_m \le k,\ m \textit{ odd}\} \\ &\langle \mathcal{B}^e, f^e \rangle, \quad \mathcal{B}^e := \{B_{i_1} \cap \cdots \cap B_{i_m} : 0 \le i_1 \le \cdots \le i_m \le k,\ m \textit{ even}\} \end{aligned} \tag{4.1}$$

where

$$f^o(B) = |\{(i_1, ..., i_m) \colon 0 \le i_1 \le \cdots \le i_m \le k,\ m \textit{ odd},\ B = B_{i_1} \cap \cdots \cap B_{i_m}\}|,$$

for $B \in \mathcal{B}^o$ and f^e is defined analogously. Then

$$\mathbb{1}_{B_1 \cup \cdots \cup B_k} = \sum_{B \in \mathcal{B}^o} \big(f^o(B) - f^e(B)\big)^+ \mathbb{1}_B - \sum_{B \in \mathcal{B}^e} \big(f^e(B) - f^o(B)\big)^+ \mathbb{1}_B.$$

Proof. Applying the classical Inclusion-Exclusion Principle (see e.g. Loera, Hemmecke, and Köppe [14, Lemma 6.1.2]) to $\mathbb{1}_{B_1 \cup \cdots \cup B_k}$ yields

$$\mathbb{1}_{B_1 \cup \cdots \cup B_k} = \sum_{B \in \mathcal{B}^o} f^o(B)\, \mathbb{1}_B - \sum_{B \in \mathcal{B}^e} f^e(B)\, \mathbb{1}_B.$$

Then, by rearranging the terms and using the fact that $f^e(B) = 0$ when $B \in \mathcal{B}^o \setminus \mathcal{B}^e$ and $f^o(B) = 0$ when $B \in \mathcal{B}^e \setminus \mathcal{B}^o$ we obtain

$$\begin{aligned} \sum_{B \in \mathcal{B}^o} f^o(B)\, \mathbb{1}_B - \sum_{B \in \mathcal{B}^e} f^e(B) \mathbb{1}_B = &\sum_{B \in \mathcal{B}^o \setminus \mathcal{B}^e} f^o(B)\, \mathbb{1}_B - \sum_{B \in \mathcal{B}^e \setminus \mathcal{B}^o} f^e(B)\, \mathbb{1}_B \\ &+ \sum_{B \in \mathcal{B}^o \cap \mathcal{B}^e} \big(f^o(B) - f^e(B)\big)\, \mathbb{1}_B \end{aligned}$$

The statement then follows from

$$\big(f^o(B) - f^e(B)\big) = \big(f^o(B) - f^e(B)\big)^+ - \big(f^e(B) - f^o(B)\big)^+.$$

which completes the proof. □

Remark 4.4. Lemma 4.3 establishes a non-redundant version of the classical Inclusion-Exclusion Principle. To illustrate this, consider $B_1, B_2, B_3 \in \mathbb{R}^d$ such that $B_1 \cap B_2 \cap B_3 = B_1 \cap B_2$ and $B_1 \neq B_2$. Then applying the classical Inclusion-Exclusion Principle to $B_1 \cup B_2 \cup B_3$ yields

$$\mathbb{1}_{B_1 \cup B_2 \cup B_3} = \mathbb{1}_{B_1} + \mathbb{1}_{B_2} + \mathbb{1}_{B_3} - \mathbb{1}_{B_1 \cap B_2} - \mathbb{1}_{B_1 \cap B_3} - \mathbb{1}_{B_2 \cap B_3} + \mathbb{1}_{B_1 \cap B_2 \cap B_3},$$

where the terms $-\mathbb{1}_{B_1 \cap B_2}$ and $+\mathbb{1}_{B_1 \cap B_2 \cap B_3}$ cancel each other out. This superfluous subtraction and addition of terms is avoided using the multisets $\langle \mathcal{B}^o, f^o \rangle$ and $\langle \mathcal{B}^e, f^e \rangle$ as in Lemma 4.3. Due to $f^o(B_1 \cap B_2) = f^e(B_1 \cap B_2 \cap B_3)$ we then obtain

$$\mathbb{1}_{B_1 \cup B_2 \cup B_3} = \mathbb{1}_{B_1} + \mathbb{1}_{B_2} + \mathbb{1}_{B_3} - \mathbb{1}_{B_1 \cap B_3} - \mathbb{1}_{B_2 \cap B_3},$$

and thus a more parsimonious representation of $\mathbb{1}_{B_1 \cup B_2 \cup B_3}$. ♦

In the following we denote the componentwise minimum of vectors $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathbb{R}^d$ by

$$\min(\mathbf{u}^1, ..., \mathbf{u}^k) = \Big(\min_{n=1,...,k} \{u_1^n\}, ..., \min_{n=1,...,k} \{u_d^n\} \Big).$$

Moreover, let us define the sets

$$\begin{aligned} \mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k) &:= \{\min(\mathbf{u}^{i_1}, ..., \mathbf{u}^{i_m}) \colon 0 \le i_1 \le \cdots \le i_m \le k \;\; m \text{ odd}\} \\ \mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k) &:= \{\min(\mathbf{u}^{i_1}, ..., \mathbf{u}^{i_m}) \colon 0 \le i_1 \le \cdots \le i_m \le k \;\; m \text{ even}\}. \end{aligned} \tag{4.2}$$

We refer to the multiplicity function

$$l^o(\mathbf{u}) := |\{(i_1, ..., i_m) \colon 0 \le i_1 \le \cdots \le i_m \le k, \; m \text{ odd}, \; \mathbf{u} = \min(\mathbf{u}^{i_1}, ..., \mathbf{u}^{i_m})\}|$$

for $\mathbf{u} \in \mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ as the *multiplicity function associated to* $\mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ and define the multiplicity function associated to $\mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$, denoted by l^e, analogously. The normalization of a vector $\mathbf{u} \in \mathbb{R}^d$ by the marginals $F_1, ..., F_d$ is denoted by $F(\mathbf{u}) := \big(F_1(u_1), ..., F_d(u_d)\big)$ as well as the left-continuous version $F^-(\mathbf{u}) := \big(F_1^-(u_1), ..., F_d^-(u_d)\big)$. Finally, for $\boldsymbol{\varepsilon} := (\varepsilon, ..., \varepsilon) \in \mathbb{R}^d$ and $\mathbf{u} \in \mathbb{R}^d$ we denote $\mathbf{u} + \boldsymbol{\varepsilon} = (u_1 + \varepsilon, ..., u_d + \varepsilon)$.

Lemma 4.5. *Let $k \in \mathbb{N}$ and $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi(s)$. Define the functions*

$$h := \sum_{\mathbf{u} \in \mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \Lambda_{\mathbf{u}}; \quad g_{\boldsymbol{\varepsilon}} := \sum_{\mathbf{u} \in \mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \Lambda_{\mathbf{u}+\boldsymbol{\varepsilon}},$$

for l^o and l^e being the multiplicity functions associated to $\mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ and $\mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$ respectively. Then for every $\boldsymbol{\varepsilon} > \mathbf{0}$ it holds that $(h, g_{\boldsymbol{\varepsilon}})$ is admissible for the dual problem $\underline{D}_\varphi$, i.e. $(h, g_{\boldsymbol{\varepsilon}}) \in \underline{\mathcal{A}}^\square$, and the value of the objective function is given by

$$\sum_{\mathbf{u} \in \mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \underline{Q}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u} \in \mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \overline{Q}\big(F^-(\mathbf{u} + \boldsymbol{\varepsilon})\big).$$

Proof. It suffices to show that $h - g_{\boldsymbol{\varepsilon}}^- \le \varphi$ for any $\boldsymbol{\varepsilon} > \mathbf{0}$. Recalling the notion of the sublevel set

$$\mathcal{U}_\psi(s) = \big\{(u_1, ...u_d) \in \mathbb{R}^d \colon \psi(u_1, ..., u_d) < s\big\},$$

we have for every $(u_1, ..., u_d) \in \mathcal{U}_\psi(s)$ that

$$\big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon x_1 \le u_1, ..., x_d \le u_d\big\} \subset \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \psi(x_1, \cdots, x_d) < s\big\},$$

due to the fact that ψ is increasing in each coordinate.

Hence, for $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi(s)$ and $B_n := \{(x_1, ..., x_d) \in \mathbb{R}^d \colon x_1 \le u_1^n, ..., x_d \le u_d^n\}$ for $n = 1, ..., k$ we have that

$$\bigcup_{n=1}^{k} B_n \subset \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \psi(x_1, \cdots, x_d) < s\big\}.$$

Now applying the Inclusion-Exclusion Principle for multisets (Lemma 4.3) to $\bigcup_{n=1}^k B_n$ we obtain

$$\mathbb{1}_{B_1 \cup \cdots \cup B_k} = \sum_{B \in \mathcal{B}^o} \big(f^o(B) - f^e(B)\big)^+ \mathbb{1}_B - \sum_{B \in \mathcal{B}^e} \big(f^e(B) - f^o(B)\big)^+ \mathbb{1}_B,$$

where $\mathcal{B}^o$ and $\mathcal{B}^e$ are as in (4.1) and f^o, f^e are the respective multiplicity functions. Moreover, we have for $\bigcap_{l=1}^m B_{n_l} \in \mathcal{B}^o \cup \mathcal{B}^e$ that

$$\begin{aligned}
\bigcap_{l=1}^{m} B_{n_l} &= \bigcap_{l=1}^{m} \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon x_1 \le u_1^{n_l}, ..., x_d \le u_d^{n_l}\big\} \\
&= \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon x_1 \le \min(u_1^{n_1}, ..., u_1^{n_m}), ..., x_d \le \min(u_d^{n_1}, ..., u_d^{n_m})\big\} \\
&= \{\mathbf{x} \in \mathbb{R}^d \colon \mathbf{x} \le \min(\mathbf{u}^{n_1}, ..., \mathbf{u}^{n_m})\},
\end{aligned}$$

and thus $\mathbb{1}_{B_{n_1} \cap \cdots \cap B_{n_m}} = \Lambda_{\min(\mathbf{u}^{n_1}, ..., \mathbf{u}^{n_m})}$ for $\min(\mathbf{u}^{n_1}, ..., \mathbf{u}^{n_m}) \in \mathcal{M}^o \cup \mathcal{M}^e$. Also, if $B = \bigcap_{l=1}^m B_{n_l} \in \mathcal{B}^o$ we have that

$$f^o(B) = l^o(\min(\mathbf{u}^{n_1}, ..., \mathbf{u}^{n_m}))$$

and $f^e(B) = l^e(\min(\mathbf{u}^{n_1}, ..., \mathbf{u}^{n_m}))$ for $B = \bigcap_{l=1}^m B_{n_l} \in \mathcal{B}^e$. In particular, it follows for any $\boldsymbol{\varepsilon} > \mathbf{0}$ that

$$\begin{aligned}
h(\mathbf{x}) - g_{\boldsymbol{\varepsilon}}^-(\mathbf{x}) &\le h(\mathbf{x}) - g_{\mathbf{0}}(\mathbf{x}) \\
&\le \sum_{\mathbf{u} \in \mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \Lambda_{\mathbf{u}}(\mathbf{x}) - \sum_{\mathbf{u} \in \mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \Lambda_{\mathbf{u}}(\mathbf{x}) \\
&= \sum_{B \in \mathcal{B}^o} \big(f^o(B) - f^e(B)\big)^+ \mathbb{1}_B(\mathbf{x}) - \sum_{B \in \mathcal{B}^e} \big(f^e(B) - f^o(B)\big)^+ \mathbb{1}_B(\mathbf{x}) \\
&= \mathbb{1}_{B_1 \cup \cdots \cup B_k}(\mathbf{x}) \le \mathbb{1}_{\psi(\mathbf{x}) < s} = \varphi(\mathbf{x})
\end{aligned}$$

for all $\mathbf{x} \in \mathbb{R}^d$ and so $(h, g_{\boldsymbol{\varepsilon}}) \in \underline{\mathcal{A}}^\square$ which completes the proof. $\square$

We are now in the position to state the reduced optimization problem for $\underline{D}_\varphi$.

Corollary 4.6. *Let* $\varphi(x_1, ..., x_d) = \mathbb{1}_{\psi(x_1, ..., x_d) < s}$ *for* $\psi \colon \mathbb{R}^d \to \mathbb{R}$ *increasing in each coordinate and let*

$$\begin{aligned}
\underline{D}_\varphi^\square(k) := \sup\Bigg\{ &\sum_{\mathbf{u} \in \mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \underline{Q}\big(F(\mathbf{u})\big) \\
&- \sum_{\mathbf{u} \in \mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \overline{Q}\big(F(\mathbf{u})\big) \colon \mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi(s) \Bigg\},
\end{aligned} \tag{4.3}$$

where l^o and l^e are the canonical multiplicity functions associated to $\mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ and $\mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$ respectively. Then

$$\underline{D}^{\square}_{\varphi}(k) \leq \underline{D}^{\square}_{\varphi}(k+1) \leq ... \leq \underline{D}_{\varphi}.$$

Proof. We first show that $\underline{D}^{\square}_{\varphi}(k) \leq \underline{D}^{\square}_{\varphi}(k+1)$ for $k \in \mathbb{N}$. Therefore note that when $\mathbf{u}^k = \mathbf{u}^{k+1} \in \mathcal{U}_{\psi}(s)$ it follows that $\mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^{k+1}) = \mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ and $\mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^{k+1}) = \mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$. Moreover, by straight-forward calculations we obtain

$$\begin{aligned}
&\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^{k+1})} \big(l^o_{k+1}(\mathbf{u}) - l^e_{k+1}(\mathbf{u})\big)^+ \underline{Q}\big(F(\mathbf{u})\big) \\
&\qquad\qquad - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^{k+1})} \big(l^e_{k+1}(\mathbf{u}) - l^o_{k+1}(\mathbf{u})\big)^+ \overline{Q}\big(F(\mathbf{u})\big) \\
&= \sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^{k})} \big(l^o_{k}(\mathbf{u}) - l^e_{k}(\mathbf{u})\big)^+ \underline{Q}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^{k})} \big(l^e_{k}(\mathbf{u}) - l^o_{k}(\mathbf{u})\big)^+ \overline{Q}\big(F(\mathbf{u})\big),
\end{aligned}$$

where l^o_j, l^e_j are the canonical multiplicity functions associated to $\mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^j)$ and $\mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^j)$ respectively for $j = k, k+1$. This implies in particular that $\underline{D}^{\square}_{\varphi}(k) \leq \underline{D}^{\square}_{\varphi}(k+1)$ for all $k \in \mathbb{N}$.

Furthermore, the inequality $\underline{D}^{\square}_{\varphi}(k) \leq \underline{D}_{\varphi}$ for all $k \in \mathbb{N}$ follows by an application of Lemma 4.5 to $(h, g_{\boldsymbol{\varepsilon}})$ where $\boldsymbol{\varepsilon} > \mathbf{0}$ and

$$h = \sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \Lambda_{\mathbf{u}}; \quad g_{\boldsymbol{\varepsilon}} = \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \Lambda_{\mathbf{u}+\boldsymbol{\varepsilon}}.$$

This yields that $(h, g_{\boldsymbol{\varepsilon}}) \in \underline{\mathcal{A}}^{\square}$ and the respective value of the objective function is given by

$$\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \underline{Q}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \overline{Q}\big(F^-(\mathbf{u}+\boldsymbol{\varepsilon})\big).$$

In particular, we have for all $k \in \mathbb{N}$ that

$$\begin{aligned}
\sup\Bigg\{ &\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \underline{Q}\big(F(\mathbf{u})\big) \\
- &\sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \overline{Q}\big(F^-(\mathbf{u}+\boldsymbol{\varepsilon})\big) : \boldsymbol{\varepsilon} > \mathbf{0};\ \mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_{\psi}(s)\Bigg\} =: c^* \leq \underline{D}_{\varphi}.
\end{aligned} \tag{4.4}$$

Moreover, it holds for all $\boldsymbol{\varepsilon} > \mathbf{0}$ that

$$\begin{aligned}
&\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \underline{Q}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \overline{Q}\big(F^-(\mathbf{u}+\boldsymbol{\varepsilon})\big) \\
&\leq \sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \underline{Q}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \overline{Q}\big(F(\mathbf{u})\big)
\end{aligned}$$

as well as $\lim_{\boldsymbol{\varepsilon}\to\mathbf{0}} \overline{Q}\big(F^{-}(\mathbf{u}+\boldsymbol{\varepsilon})\big) = \overline{Q}\big(F(\mathbf{u})\big)$ due to the Lipschitz continuity of $\overline{Q}$. Hence using (4.4) it follows that

$$c^* = \sup\Bigg\{ \sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,\dots,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \, \underline{Q}\big(F(\mathbf{u})\big) \\ - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,\dots,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \, \overline{Q}\big(F(\mathbf{u})\big) \colon \mathbf{u}^1,\dots,\mathbf{u}^k \in \mathcal{U}_\psi(s)\Bigg\}$$

which completes the proof. □

Corollary 4.6 establishes a tractable optimization problem to compute a lower bound on $\underline{D}_\varphi$ and thus also on $\underline{P}_\varphi$. The optimization takes place over vectors in the sublevel set $\mathcal{U}_\psi(s)$ and the trade-off between the computational effort and the quality of the bound is moderated by the variable k. For fixed k, $\underline{D}^{\square}_\varphi(k)$ amounts to a $k \cdot d$ dimensional optimization that can be solved with standard optimization packages. Note, that most mathematical programming environments also provide efficient built-in procedures to compute the multiplicity functions l^o and l^k.

4.2. A reduction scheme for $\overline{D}_\varphi$

We proceed with the development of a similar reduction scheme based on the dual $\overline{D}_\varphi$. Recall from (3.3) that

$$\overline{D}_\varphi = \inf\Bigg\{\overline{Q}(h^-) - \underline{Q}(g) + \sum_{i=1}^d \mathbb{E}_i[\nu_i] \colon \nu_i \in \mathcal{L}(F_i),\ i = 1,\dots,d, \\ h, g \in \mathcal{R} \text{ s.t. } h^- - g + \sum_{i=1}^d \nu_i \geq \varphi\Bigg\}.$$

We refer to the class of admissible functions for $\overline{D}_\varphi$ by

$$\overline{\mathcal{A}} := \Bigg\{(h, g, \nu_1, \dots, \nu_d) \colon \nu_i \in \mathcal{L}(F_i),\ i = 1,\dots,d,\ h, g \in \mathcal{R} \text{ s.t. } h^- - g + \sum_{i=1}^d \nu_i \geq \varphi\Bigg\}$$

and for each admissible function the corresponding value of the objective function is given by

$$\overline{Q}(h^-) - \underline{Q}(g) + \sum_{i=1}^d \mathbb{E}_i[\nu_i].$$

Again, for our reduction scheme, we consider a subclass of admissible pairs (h, g) such that $h, g \in \mathcal{R}^{\square}$, i.e.

$$\overline{\mathcal{A}}^{\square} := \big\{(h, g) \colon h, g \in \mathcal{R}^{\square} \text{ s.t. } h^- - g \geq \varphi\big\}$$

We now turn to the formal construction of admissible functions in $\overline{\mathcal{A}}^{\square}$ with an auxiliary version of the multiset Inclusion-Exclusion Principle for intersections.

Lemma 4.7. *Let* $B_1^n, ..., B_d^n \subset \mathbb{R}^d$ *for* $n = 1, ..., k$ *and* $k \in \mathbb{N}$ *and define*

$$G_{(i_1,...,i_k)} := (B_{i_1}^1 \cap \cdots \cap B_{i_k}^k) \quad \textit{for } (i_1, ..., i_k) \in \{1, ..., d\}^k, \quad \textit{and} \quad \mathcal{B} = \bigcap_{n=1}^{k} \bigcup_{l=1}^{d} B_l^n.$$

Moreover, for an enumeration $\{\mathbf{i}^1, ..., \mathbf{i}^{dk}\}$ *of the set* $\{1, ..., d\}^k$ *define the multisets*

$$\langle \mathcal{G}^o, f^o \rangle, \quad \mathcal{G}^o := \{G_{\mathbf{i}^{n_1}} \cap \cdots \cap G_{\mathbf{i}^{n_m}} : 0 \le n_1 \le \cdots \le n_m \le k, \ m \textit{ odd}\} \tag{4.5}$$

$$\langle \mathcal{G}^e, f^e \rangle, \quad \mathcal{G}^e := \{G_{\mathbf{i}^{n_1}} \cap \cdots \cap G_{\mathbf{i}^{n_m}} : 0 \le n_1 \le \cdots \le n_m \le k, \ m \textit{ even}\} \tag{4.6}$$

where

$$f^o(G) = |\{(\mathbf{i}^{n_1}, ..., \mathbf{i}^{n_m}) \colon 0 \le n_1 \le \cdots \le n_m \le k, \ m \textit{ odd}, \ G = G_{\mathbf{i}^{n_1}} \cap \cdots \cap G_{\mathbf{i}^{n_m}}\}|$$

and f^e *is defined analogously. Then it holds that*

$$\mathbb{1}_{\mathcal{B}} = \sum_{G \in \mathcal{G}^o} \big(f^o(G) - f^e(G)\big)^+ \mathbb{1}_G - \sum_{G \in \mathcal{G}^e} \big(f^e(G) - f^o(G)\big)^+ \mathbb{1}_G.$$

Proof. Since the union and the intersection of sets commute we have that $\mathcal{B} = \mathcal{G}_{\mathbf{i}^1} \cup \cdots \cup \mathcal{G}_{\mathbf{i}^{kd}}$ and hence the statement follows by a straight-forward application of Lemma 4.3. □

We are now ready to establish an explicit construction of admissible pairs $(h, g) \in \overline{\mathcal{A}}^{\square}$. To this end, let us denote for $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathbb{R}^d$ and an enumeration $\{\mathbf{i}^1, ..., \mathbf{i}^{dk}\}$ of $\{1, ..., d\}^k$

$$\mathrm{U}_{\mathbf{i}^n} := \min(\mathrm{pr}_{i_1}(\mathbf{u}^1), ..., \mathrm{pr}_{i_k}(\mathbf{u}^k)) \quad \text{for } (i_1, ..., i_k) = \mathbf{i}^n, n = 1, ..., dk,$$

where $\mathrm{pr}_i(\mathbf{u}) := (\infty, ..., \infty, u_i, \infty, ..., \infty)$ for $i \in \{1, ..., d\}$. Moreover, we define

$$\begin{aligned} \mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k) &:= \{\min(\mathrm{U}_{\mathbf{i}^{n_1}}, ..., \mathrm{U}_{\mathbf{i}^{n_m}}) \colon 0 \le n_1 \le \cdots \le n_m \le k \ \ m \text{ odd}\} \\ &= \mathcal{M}^o(\mathrm{U}_{\mathbf{i}^1}, ..., \mathrm{U}_{\mathbf{i}^{dk}}) \\ \mathcal{W}^e(\mathbf{u}^1, ..., \mathbf{u}^k) &:= \{\min(\mathrm{U}_{\mathbf{i}^{n_1}}, ..., \mathrm{U}_{\mathbf{i}^{n_m}}) \colon 0 \le n_1 \le \cdots \le n_m \le k \ \ m \text{ even}\} \\ &= \mathcal{M}^e(\mathrm{U}_{\mathbf{i}^1}, ..., \mathrm{U}_{\mathbf{i}^{dk}}). \end{aligned} \tag{4.7}$$

Finally, we write $\mathbf{u} < \mathbf{v}$ for vectors $\mathbf{u}, \mathbf{v} \in \overline{\mathbb{R}}^d$ such that $u_i < v_i$ for $i = 1, ..., d$.

Lemma 4.8. *Let* $k \in \mathbb{N}$ *and* $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_{\psi}^c(s)$. *Define the functions*

$$h_{\boldsymbol{\varepsilon}} := \sum_{\mathbf{u} \in \mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \Lambda_{\mathbf{u} - \boldsymbol{\varepsilon}}; \quad g := \sum_{\mathbf{u} \in \mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \Lambda_{\mathbf{u}},$$

for l^o *and* l^e *being the multiplicity functions associated to* $\mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ *and* $\mathcal{W}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$ *respectively. Then for every* $\boldsymbol{\varepsilon} > \mathbf{0}$ *it holds that* $(h_{\boldsymbol{\varepsilon}}, g)$ *is admissible for the dual problem* $\overline{D}_{\varphi}$, *i.e.* $(h_{\boldsymbol{\varepsilon}}, g) \in \overline{\mathcal{A}}^{\square}$, *and the value of the objective function is given by*

$$\sum_{\mathbf{u} \in \mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \, \overline{Q}\big(F^-(\mathbf{u} + \boldsymbol{\varepsilon})\big) \ - \sum_{\mathbf{u} \in \mathcal{W}^e(\mathbf{u}^1, ..., \mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \, \underline{Q}\big(F(\mathbf{u})\big).$$

Proof. We need to show that $h_{\varepsilon}^{-} - g \geq \varphi$. Note, that for every $(u_1, ..., u_d) \in \mathcal{U}_{\psi}^{c}(s)$ we have that

$$\big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon x_1 \geq u_1, ..., x_d \geq u_d\big\}^c \supset \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \psi(x_1, \cdots, x_d) < s\big\},$$

due to the fact that ψ is increasing in each coordinate. Hence, for $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_{\psi}^{c}(s)$ and $B_n := \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon x_1 \geq u_1^n, ..., x_d \geq u_d^n\big\}$ for $n = 1, ..., k$ it follows that

$$\bigcap_{n=1}^{k} B_n^c \supset \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \psi(x_1, \cdots, x_d) < s\big\}. \tag{4.8}$$

Moreover, for $n = 1, ..., k$ we have that

$$B_n^c = \bigcup_{i=1}^{d} \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \mathbf{x} < \mathrm{pr}(u^n)_i\big\},$$

which follows from

$$\begin{aligned}
\big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon x_1 \geq u_1, ..., x_d \geq u_d\big\}^c &= \big([u_1, \infty) \times \cdots \times [u_d, \infty)\big)^c \\
&= \bigcup_{i=1}^{d} \mathbb{R} \times \cdots \times \mathbb{R} \times (-\infty, u_i) \times \mathbb{R} \times \cdots \times \mathbb{R} \\
&= \bigcup_{i=1}^{d} \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \mathbf{x} < \mathrm{pr}(u^n)_i\big\}.
\end{aligned}$$

Hence, we obtain

$$\bigcap_{n=1}^{k} B_n^c = \bigcap_{n=1}^{k} \bigcup_{i=1}^{d} \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \mathbf{x} < \mathrm{pr}(u^n)_i\big\}. \tag{4.9}$$

Now, defining

$$\begin{aligned}
&H_i^n := \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \mathbf{x} < \mathrm{pr}(u^n)_i)\big\}, \text{ for } i = 1, ..., d, n = 1, ..., k \text{ and} \\
&\mathcal{G}_{(i_1, ..., i_k)} := H_{i_1}^1 \cap \cdots \cap H_{i_k}^k,\ (i_1, ..., i_k) \in \{1, ..., d\}^k,
\end{aligned}$$

and applying Lemma 4.7 to equation (4.9) we arrive at

$$\mathbb{1}_{B_1^c \cap \cdots \cap B_k^c} = \sum_{G \in \mathcal{G}^o} \big(f^o(G) - f^e(G)\big)^+ \mathbb{1}_G - \sum_{G \in \mathcal{G}^e} \big(f^e(G) - f^o(G)\big)^+ \mathbb{1}_G,$$

where $\mathcal{G}^o, \mathcal{G}^e$ and f^o, f^e are defined in 4.5. Finally, note that for $1 \leq n_1 \leq \cdots \leq n_m \leq dk$

$$\mathcal{G}_{(i_1, ..., i_k)} = \big\{(x_1, ..., x_d) \in \mathbb{R}^d \colon \mathbf{x} < \min\big(\mathrm{pr}_{i_1}(\mathbf{u}^1), ..., \mathrm{pr}_{i_k}(\mathbf{u}^k)\big)\big\},$$

so that with the definition of $\mathrm{U}_{\mathbf{i}}$ for $\mathbf{i} \in \{1, ..., d\}^k$ it follows for every $1 \leq n_1, ..., n_m \leq dk$ that

$$
\begin{aligned}
\bigcap_{l=1}^{m} \mathcal{G}_{\mathbf{i}^{n_l}} &= \bigcap_{l=1}^{m} \left\{ (x_1, ..., x_d) \in \mathbb{R}^d \colon \mathbf{x} < \mathrm{U}_{\mathbf{i}^{n_l}} \right\} \\
&= \left\{ (x_1, ..., x_d) \in \mathbb{R}^d \colon \mathbf{x} < \min(\mathrm{U}_{\mathbf{i}^{n_1}}, ..., \mathrm{U}_{\mathbf{i}^{n_m}}) \right\},
\end{aligned}
$$

and thus $\mathbb{1}_{\mathcal{G}_{\mathbf{i}^{n_1}} \cap \cdots \cap \mathcal{G}_{\mathbf{i}^{n_m}}} = \Lambda^-_{\min(\mathrm{U}_{\mathbf{i}^{n_1}}, ..., \mathrm{U}_{\mathbf{i}^{n_m}})}$ for $\min(\mathrm{U}_{\mathbf{i}^{n_1}}, ..., \mathrm{U}_{\mathbf{i}^{n_m}}) \in \mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k) \cup \mathcal{W}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$. In particular, using equation (4.8) and the fact that

$$
l^o(\min(\mathrm{U}_{\mathbf{i}^{n_1}}, ..., \mathrm{U}_{\mathbf{i}^{n_m}})) = f^o(G)
$$

for $G = \bigcap_{l=1}^{m} G_{\mathbf{i}^{n_l}} \in \mathcal{G}^o$ and vice versa for l^e, we conclude that

$$
h^-_{\boldsymbol{\varepsilon}}(\mathbf{x}) - g(\mathbf{x}) \geq h^-_{\mathbf{0}}(\mathbf{x}) - g^-(\mathbf{x}) = \mathbb{1}_{B_1^c \cap \cdots \cap B_k^c}(\mathbf{x}) \geq \mathbb{1}_{\psi(\mathbf{x}) < s} = \varphi(\mathbf{x})
$$

for all $\mathbf{x} \in \mathbb{R}^d$, which completes the proof. □

We are now in the position to establish our reduction scheme based on $\overline{D}_\varphi$. The proof of the following corollary is analogous to the proof of Corollary 4.6 and therefore omitted.

Corollary 4.9. *Let* $\varphi(x_1, ..., x_d) = \mathbb{1}_{\psi(x_1, ..., x_d) < s}$ *for* $\psi \colon \mathbb{R}^d \to \mathbb{R}$ *increasing in each coordinate and let*

$$
\begin{aligned}
\overline{D}^{\square}_{\varphi}(k) := \inf \Bigg\{ & \sum_{\mathbf{u} \in \mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \left(l^o(\mathbf{u}) - l^e(\mathbf{u}) \right)^+ \overline{Q}\big(F(\mathbf{u})\big) \\
& - \sum_{\mathbf{u} \in \mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)} \left(l^e(\mathbf{u}) - l^o(\mathbf{u}) \right)^+ \underline{Q}\big(F(\mathbf{u})\big) \colon \mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}^c_\psi(s) \Bigg\},
\end{aligned} \tag{4.10}
$$

where l^o *and* l^e *are the canonical multiplicity functions associated to* $\mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ *and* $\mathcal{W}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$ *respectively. Then*

$$
\overline{D}^{\square}_{\varphi}(k) \geq \overline{D}^{\square}_{\varphi}(k+1) \geq ... \geq \overline{D}_{\varphi}.
$$

4.3. Sharp asymptotic bounds in the certainty limit

In general, the schemes $\underline{D}^{\square}_{\varphi}(k)$ and $\overline{D}^{\square}_{\varphi}(k)$ do not approximate the dual bounds $\underline{D}_\varphi$ and $\overline{D}_\varphi$ respectively for $k \to \infty$. In the homogeneous, complete dependence uncertainty case, i.e. $F_1 = \cdots = F_d = F$ and $\underline{Q} = W_d$ and $\overline{Q} = M_d$, Puccetti and Rüschendorf [20] derived an explicit solution to the dual $\underline{D}_\varphi$ under additional requirements on the marginals. They showed, that the optimizer is of the form $d \cdot \nu$ for a piecewise linear function $\nu \in \mathcal{L}(F)$ which cannot be represented by the linear combinations in $\mathcal{R}$.

The counterpart to the situation of complete dependence uncertainty is the case of certainty, i.e. the limit when $\underline{Q}$ and $\overline{Q}$ converge from below and above respectively to a copula C. A natural feature of any bound on the expectation of φ using the information from $\underline{Q}$ and $\overline{Q}$ should be that it converges to $\mathbb{E}_C[\varphi]$ as $\underline{Q}, \overline{Q} \to C$. The following theorem shows that for $k \to \infty$ the reduced bounds $\underline{D}_\varphi^\square(k)$ and $\overline{D}_\varphi^\square(k)$ indeed converge to the desired object in the certainty limit.

Theorem 4.10. *Let $\varphi(x_1, ..., x_d) = \mathbb{1}_{\psi(x_1,...,x_d)<s}$ for $\psi\colon \mathbb{R}^d \to \mathbb{R}$ increasing in each coordinate. Moreover, define for $k \in \mathbb{N}$*

$$\left[\underline{D}_\varphi^\square(k)\right](\underline{Q}, \overline{Q}) := \sup\Bigg\{ \sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \left(l^o(\mathbf{u}) - l^e(\mathbf{u})\right)^+ \underline{Q}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \left(l^e(\mathbf{u}) - l^o(\mathbf{u})\right)^+ \overline{Q}\big(F(\mathbf{u})\big) \colon \mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi(s)\Bigg\}$$

where $\mathcal{M}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ and $\mathcal{M}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$ are defined in (4.2) and l^o, l^e are the associated multiplicity functions. Analogously, let

$$\left[\overline{D}_\varphi^\square(k)\right](\underline{Q}, \overline{Q}) := \inf\Bigg\{ \sum_{\mathbf{u}\in\mathcal{W}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \left(f^o(\mathbf{u}) - f^e(\mathbf{u})\right)^+ \underline{Q}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u}\in\mathcal{W}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \left(f^e(\mathbf{u}) - f^o(\mathbf{u})\right)^+ \overline{Q}\big(F(\mathbf{u})\big) \colon \mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi^c(s)\Bigg\}.$$

where $\mathcal{W}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ and $\mathcal{W}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$ as in (4.7) with associated multiplicity functions f^o, f^e. Then it holds for any copula C and sequences of quasi-copulas $\left(\underline{Q}^j\right)_{j=1,2...}$ and $\left(\overline{Q}^j\right)_{j=1,2...}$ with $\underline{Q}^j \le C \le \overline{Q}^j$ for all $j \in \mathbb{N}$ and $\underline{Q}^j, \overline{Q}^j \to_j C$ pointwise, that

$$\lim_j \inf_k \left[\overline{D}_\varphi^\square(k)\right](\underline{Q}^j, \overline{Q}^j) = \lim_j \sup_k \left[\underline{D}_\varphi^\square(k)\right](\underline{Q}^j, \overline{Q}^j) = \mathbb{P}_C(\psi(X_1, ..., X_d) < s)$$

Proof. We show that

$$\lim_j \sup_k \left[\underline{D}_\varphi^\square(k)\right](\underline{Q}^j, \overline{Q}^j) = \mathbb{P}_C(\psi(X_1, ..., X_d) < s).$$

The proof for $\overline{D}_\varphi^\square(k)$ follows along similar lines.

First, note that there exists a sequence $\mathbf{u}^1, \mathbf{u}^2, ... \in \mathcal{U}_\psi(s)$ such that for

$$h^k := \sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \left(l^o(\mathbf{u}) - l^e(\mathbf{u})\right)^+ \Lambda_\mathbf{u}; \quad g^k := \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \left(l^e(\mathbf{u}) - l^o(\mathbf{u})\right)^+ \Lambda_\mathbf{u},$$

we have that $h^k - g^k \to_k \varphi$ pointwise. To verify the existence of such a sequence one can choose as $(\mathbf{u}^n)_{n=1,...,k}$, any discretization of the set $\mathcal{U}_\psi(s)$, whose mesh converges to zero for $k \to \infty$.

From the fact that $\mathbf{u}^1, \mathbf{u}^2, ... \in \mathcal{U}_\psi(s)$ and the proof of Lemma 4.5 it follows that $h^k - g^k \le \varphi$ for all $k \in \mathbb{N}$, and the corresponding value of the objective function is given by

$$\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \underline{Q}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \overline{Q}\big(F(\mathbf{u})\big) \le \big[\underline{D}_\varphi^\square(k)\big](\underline{Q}^j, \overline{Q}^j). \tag{4.11}$$

Using the fact that C is a copula and applying the Dominated Convergence Theorem yields

$$\mathbb{E}_C\big[h^k - g^k\big] \to_k \mathbb{E}_C[\varphi] = \mathbb{P}_C(\psi(X_1, ..., X_d) < s). \tag{4.12}$$

Therefore we can w.l.o.g. assume that for fixed j, there exists an $N \in \mathbb{N}$ such that

$$\mathbb{E}_C[h^k - g^k] \ge \big[\underline{D}_\varphi^\square(k)\big](\underline{Q}^j, \overline{Q}^j) \quad \text{for all } k \ge N,$$

since otherwise

$$\big[\underline{D}_\varphi^\square(k_m)\big](\underline{Q}^j, \overline{Q}^j) \ge \mathbb{E}_C[h^{k_m} - g^{k_m}] \to_m \mathbb{E}_C[\varphi],$$

along a subsequence $(k_m)_m$ and we are done.

We proceed by showing that the convergence

$$\lim_j \sup_k \big[\underline{D}_\varphi^\square(k)\big](\underline{Q}^j, \overline{Q}^j) = \mathbb{P}_C(\psi(X_1, ..., X_d) < s)$$

holds. To this end, fix an arbitrary $\varepsilon > 0$. Due to (4.12) we can choose $k \ge N$ such that

$$\big|\mathbb{P}_C(\psi(X_1, ..., X_d) < s) - \mathbb{E}_C\big[h^k - g^k\big]\big| < \frac{\varepsilon}{2}. \tag{4.13}$$

Moreover, the fact that quasi-copulas are Lipschitz continuous yields, by an application of the Arzelà-Ascoli Theorem, that $\underline{Q}^j \to_j C$ and $\overline{Q}^j \to_j C$ uniformly. Thus, for

$$p := \sum_{\mathbf{u}\in\mathcal{M}^o\mathbf{u}^1,...,\mathbf{u}^k)} l^o(\mathbf{u}) + \sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)} l^e(\mathbf{u})$$

we can choose an $j \in \mathbb{N}$ so that

$$\|C - \underline{Q}^j\|_\infty + \|C - \overline{Q}^j\|_\infty < \frac{\varepsilon}{2p}. \tag{4.14}$$

With this choice of k and j we arrive at

$$
\begin{aligned}
&\left|\mathbb{P}_C(\psi(X_1,...,X_d)<s)-\left[\underline{D}_\varphi^\square(k)\right](\underline{Q}^j,\overline{Q}^j)\right| \\
&\leq \left|\mathbb{P}_C(\psi(X_1,...,X_d)<s)-\mathbb{E}_C\left[h^k-g^k\right]+\mathbb{E}_C\left[h^k-g^k\right]-\left[\underline{D}_\varphi^\square(k)\right](\underline{Q}^j,\overline{Q}^j)\right| \\
&\leq \left|\mathbb{P}_C(\psi(X_1,...,X_d)<s)-\mathbb{E}_C\left[h^k-g^k\right]\right| \\
&+\Bigg|\mathbb{E}_C\left[h^k-g^k\right]-\Bigg(\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)}\left(l^o(\mathbf{u})-l^e(\mathbf{u})\right)^+\underline{Q}^j(F(\mathbf{u})) \\
&\qquad\qquad -\sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)}\left(l^e(\mathbf{u})-l^o(\mathbf{u})\right)^+\overline{Q}^j(F(\mathbf{u}))\Bigg)\Bigg| \\
&<\frac{\varepsilon}{2}+\Bigg|\sum_{\mathbf{u}\in\mathcal{M}^o(\mathbf{u}^1,...,\mathbf{u}^k)}\left(l^o(\mathbf{u})-l^e(\mathbf{u})\right)^+\Big(C(F(\mathbf{u}))-\underline{Q}^j(F(\mathbf{u}))\Big) \\
&\qquad\qquad -\sum_{\mathbf{u}\in\mathcal{M}^e(\mathbf{u}^1,...,\mathbf{u}^k)}\left(l^o(\mathbf{u})-l^e(\mathbf{u})\right)^+\Big(C(F(\mathbf{u}))-\overline{Q}^j(F(\mathbf{u}))\Big)\Bigg| \\
&<\frac{\varepsilon}{2}+p\big(\|C-\underline{Q}^j\|_\infty+\|C-\overline{Q}^j\|_\infty\big)<\frac{\varepsilon}{2}+p\frac{\varepsilon}{2p}=\varepsilon.
\end{aligned}
$$

The second inequality is a consequence of equation (4.11) and the fact that $\mathbb{E}_C[h^k-g^k]\geq\left[\underline{D}_\varphi^\square(k)\right](\underline{Q}^j,\overline{Q}^j)$. The third inequality follows from equation (4.13) and last inequality holds due to equation (4.14).

Finally, since ε was arbitrary we have shown that

$$\lim_j\sup_k\left[\underline{D}_\varphi^\square(k)\right](\underline{Q}^j,\overline{Q}^j)=\mathbb{P}_C(\psi(X_1,...,X_d)<s)$$

holds and hence the proof is complete. □

5. Illustrations and numerical examples

In this section we provide an informal description of the reduction schemes in section 4.1 and 4.2 in order to illustrate the underlying idea. Furthermore, we provide several numerical examples comparing the performance of our reduction scheme to the improved standard bounds from [8, 9].

A graphical illustration of $\underline{D}_\varphi^\square(k)$

For a graphical illustration of the scheme $\underline{D}_\varphi^\square(k)$ let us consider let $\psi(x_1,x_2)=x_1+x_2$ and F_1,F_2 uniform distributions on $[0,1]$. Due to assumption **(A1)** and **(A2)**, we consider admissible functions which are sums of characteristic functions of rectangular regions in $\mathcal{U}_\psi(s)$, as in

$$h-g=\sum_{n=1}^k\Lambda_{\mathbf{u}^k}-\sum_{n=1}^m\Lambda_{\mathbf{v}^n}\leq\mathbb{1}_{x_1+x_2<s}$$

for $\mathbf{u}^1, ..., \mathbf{u}^k, \mathbf{v}^1, ..., \mathbf{v}^m \in \mathcal{U}_\psi(s)$ and $k, m \in \mathbb{N}$. The corresponding value of the objective function to be maximized is given by $\underline{Q}(h^-) - \overline{Q}(g)$ for each pair $h, g \in \mathcal{R}^\square$. Figure 5 illustrates the structure of admissible functions of this type that we shall consider.

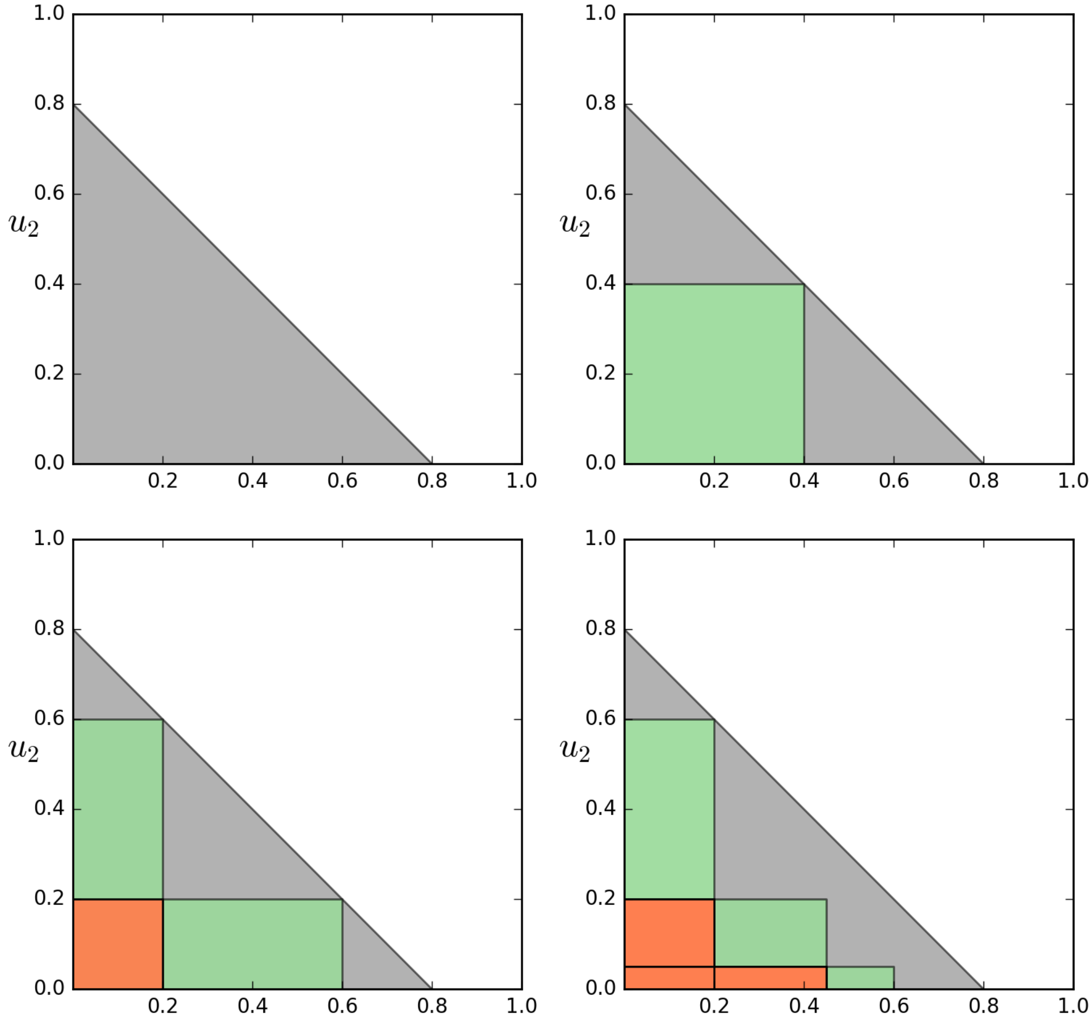

Figure 1: Constraint set of admissible functions

The gray triangular region in figure 5, corresponds to the area

$$\mathcal{U}_\psi(s) = \{(u_1, u_2) \in [0, 1]^2 \colon u_1 + u_2 < s\}$$

for $s = 0.8$. The upper LHS simply depicts the region $\mathcal{U}_\psi(s)$. The green area in the upper RHS figure corresponds to the rectangle $[0, 0.4]^2 \subset \mathcal{U}_\psi(s)$ as induced by the function $h = \Lambda_{(0.4,0.4)}$, i.e. $\Lambda_{(0.4,0.4)}(u_1, u_2) = \varphi(u_1, u_2) = 1$ for all $(u_1, u_2) \in [0, 0.4]^2$. The value of the objective function for h is given by

$$\underline{Q}(\Lambda^-_{(0.4,0.4)}) = \underline{Q}(F_1(0.4), F_2(0.4)) = \underline{Q}(0.4, 0.4).$$

Similarly, the lower LHS represents the rectangles $[0,0.2]\times[0,0.6]$ and $[0,0.6]\times[0,0.2]$ induced by $\Lambda_{(0.2,0.6)}$ and $\Lambda_{(0.6,0.2)}$. The red area corresponds to $[0,0.2]\times[0,0.2]$ where an overlap occurs due to

$$h(\mathbf{u}) = \Lambda_{(0.2,0.6)}(\mathbf{u}) + \Lambda_{(0.6,0.2)}(\mathbf{u}) = 2 > \varphi(\mathbf{u}) \quad \text{for all } \mathbf{u} \in [0,0.2]\times[0,0.2].$$

This overlap is then compensated by applying the Inclusion-Exclusion Principle and subtracting $g = \Lambda_{(0.2,0.2)}$, yielding the admissible function

$$h - g = \Lambda_{(0.2,0.6)} + \Lambda_{(0.6,0.2)} - \Lambda_{(0.2,0.2)}.$$

The respective value of the objective function is equal to $\underline{Q}(0.2,0.6)+\underline{Q}(0.6,0.2)-\overline{Q}(0.2,0.2)$. Finally, the lower RHS represents the function constructed by

$$h = \Lambda_{(0.2,0.6)} + \Lambda_{(0.45,0.2)} + \Lambda_{(0.6,0.05)}$$

and an appropriate compensation of the overlap by $g = \Lambda_{(0.2,0.2)} + \Lambda_{(0.45,0.05)}$ so that $(h,g) \in \underline{\mathcal{A}}$ and the corresponding value of the objective function is equal to

$$\underline{Q}(0.2,0.6) + \underline{Q}(0.45,0.2) + \underline{Q}(0.6,0.05) - \overline{Q}(0.2,0.2) - \overline{Q}(0.45,0.05).$$

Note, that the construction of (h,g) depends entirely on the choice $\mathbf{u}_1,\ldots,\mathbf{u}_k \in \mathbb{R}^d$. In particular, maximizing over all (h,g) that are constructed in this way amounts to an optimization over $\mathbf{u}_1,\ldots,\mathbf{u}_k \in \mathcal{U}_\psi(s)$.

Lemma 4.5 formalizes this construction of admissible (h,g) whereas Corollary 4.6 establishes the optimization over such (h,g) as to obtain bounds on over which we then maximize in order to obtain bounds on $\underline{D}_\varphi$.

A graphical illustration of $\overline{D}_\varphi^\square(k)$

Using again $\psi(x_1,x_2) = x_1 + x_2$ and F_1, F_2 uniform distributions on $[0,1]$ let us illustrate the idea of the scheme $\overline{D}_\varphi^\square(k)$. This time, (h,g) with $h,g \in \mathcal{R}^\square$ are admissible when

$$h - g = \sum_{n=1}^{k} \Lambda_{\mathbf{u}^k} - \sum_{n=1}^{m} \Lambda_{\mathbf{v}^n} \geq \mathbb{1}_{x_1+x_2<s}.$$

Figure **??** illustrates two possible constructions of admissible pairs (h,g). Again the green area corresponds to $\{\mathbf{x} \in [0,1]^2 \colon h(\mathbf{x}) = 1\}$ whereas the red shaded area marks overlaps $(h(\mathbf{x}) > 1)$ which we compensate using the inclusion exclusion principle. The LHS corresponds to

$$h - g = \Lambda_{(0.8,0.4)} + \Lambda_{(0.4,0.8)} - \Lambda_{(0.4,0.4)} \in \overline{\mathcal{A}}$$

and the respective value of the objective function amounts to $\overline{Q}(0.8,0.4)+\overline{Q}(0.4,0.8)-\underline{Q}(0.4,0.4)$. The RHS represents the admissible function given by $h-g$ for

$$h = \Lambda_{(0.8,0.2)} + \Lambda_{(0.2,0.8)} + \Lambda_{(0.6,0.6)}; \quad g = \Lambda_{(0.2,0.6)} + \Lambda_{(0.6,0.2)}$$

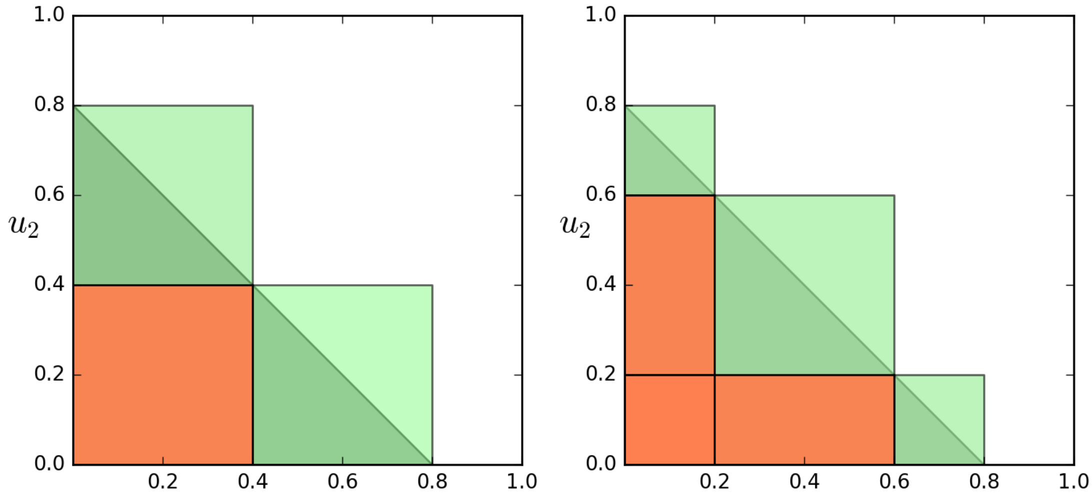

Figure 2: Constraint set of admissible functions

with corresponding value of the objective function

$$\overline{Q}(0.8, 0.2) + \overline{Q}(0.2, 0.8) - \overline{Q}(0.6, 0.6) - \underline{Q}(0.2, 0.6) - \underline{Q}(0.6, 0.2).$$

Note, that in contrast to $\underline{D}_{\varphi}^{\square}(k)$ it does not suffice to consider $h = \sum_{n=1}^{k} \mathbf{u}^k$ for $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_{\psi}(s)$. A construction of admissible functions in the spirit of section 4.1 is however possible if we formulate it in terms of characteristic functions of upper level sets of the form

$$\{(x_1, x_2) \colon x_1 \geq u_1, \ x_2 \geq u_2\},$$

for $\mathbf{u}$ in the complement $\mathcal{U}_{\psi}^c(s)$. Returning to the LHS of figure **??** we then note that the region

$$\{\mathbf{x} \in \mathbb{R}^2 \colon \Lambda_{(0.8,0.4)}(\mathbf{x}) + \Lambda_{(0.4,0.8)}(\mathbf{x}) - \Lambda_{(0.4,0.4)}(\mathbf{x}) = 1\}$$

can be expressed in terms of complements of upper level sets via

$$\big([0.4, 1] \times [0.4, 1]\big)^c \cap \big([0, 1] \times [0.8, 1]\big)^c \cap \big([0.8, 1] \times [0, 1]\big)^c,$$

and a similar representation holds for the RHS. Moreover, we can represent the complements via

$$\big([u_1, 1] \times [u_2, 1]\big)^c = ([u_1, 1] \times [0, 1]) \cup ([0, 1] \times [u_2, 1]),$$

where the right-hand side of the equation is the union of sets that can be evaluated by means of the quasi-copulas $\underline{Q}, \overline{Q}$. This construction is made rigorous in Lemma 4.8 and the resulting optimization is provided in Corollary 4.9.

The following numerical examples show, how the reduction schemes can be applied in order to account for copula information in the computation of VaR estimates. Our results are compared to the Improved Standard Bounds using the same information.

Example 5.1. Consider an $\mathbb{R}^2$-valued risk vector (X_1, X_2) with copula C and Pareto$_2$-marginals. We assume that the copula C lies in the vicinity of a reference copula C^* as measured by the Kolmogorov–Smirnov distance, i.e.

$$\mathcal{D}_{\text{KS}}(C, C^*) \leq \delta$$

for some $\delta > 0$. Hereby, C^* is assumed to be a Gaussian copula with correlation ρ. We then compute VaR estimates on the sum $X_1 + X_2$, using the copula information via the reduction schemes presented in Section 4.1 and 4.2. Applying Theroem 5.4 in Lux and Papapantoleon [16] in conjunction with the explicit representation of the improved Fréchet–Hoeffding bounds given in [16, Lemma 5.7], we obtain

$$\begin{aligned} \underline{Q}^{\mathcal{D}_{\text{KS}},\delta}(\mathbf{u}) = \max\{C^*(\mathbf{u}) - \delta, W_d(\mathbf{u})\} &\leq C(\mathbf{u}) \\ &\leq \min\{C^*(\mathbf{u}) + \delta, M_2(\mathbf{u})\} = \overline{Q}^{\mathcal{D}_{\text{KS}},\delta}(\mathbf{u}), \end{aligned} \tag{5.1}$$

for all $\mathbf{u} \in \mathbb{I}^d$. Note that because of $d = 2$ the bounds apply analogously to the survival function of C, i.e.

$$\widehat{\underline{Q}}^{\mathcal{D}_{\text{KS}},\delta} \leq \widehat{C} \leq \widehat{\overline{Q}}^{\mathcal{D}_{\text{KS}},\delta}. \tag{5.2}$$

Our reduction schemes now allow us to translate these improved Fréchet–Hoeffding bounds into VaR estimates. As a benchmark to demonstrate the quality of our estimates, we compare them to the Improved Standard Bounds, which are given by the inverses of the following bounds on the quantile function of $X_1 + X_2$:

$$\begin{aligned} \mathbb{P}(X_1 + X_2 < s) &\geq \sup_{x \in \mathbb{R}} \underline{Q}^{\mathcal{D}_{\text{KS}},\delta}\big(F_1(x), F_2^-(s-x)\big) \\ \mathbb{P}(X_1 + X_2 < s) &\leq \inf_{x \in \mathbb{R}} 1 - \widehat{\underline{Q}}^{\mathcal{D}_{\text{KS}},\delta}\big(F_1(x), F_2(s-x)\big). \end{aligned}$$

The following tables show the values of the Improved Standard Bounds on the VaR of the sum $X_1 + X_2$ for different confidence levels α. These are compared to the VaR bounds obtained by inverting $\underline{D}_{\varphi}^{\square}(3)$ and $\overline{D}_{\varphi}^{\square}(3)$, for $\varphi(x_1, x_2) = \mathbb{1}_{x_1+x_2<s}$, along the variable s. Analogously, we compute VaR estimates by inverting $\widehat{\underline{D}}_{\varphi}^{\square}(3)$ and $\widehat{\overline{D}}_{\varphi}^{\square}(3)$, using the bounds on the survival copula in (5.2). We thus obtain two lower and two upper VaR estimates for each α, of which the largest lower bound and the lowest upper bound respectively are reported in each of the tables. For $k \geq 4$ no further improvement of the bounds was obtained. For the sake of legibility, the results are rounded to one decimal place. Table 1 shows the VaR estimates for different levels of the correlation of the reference copula and $\delta = 0.0001$, while for Table 2 we assume that $\delta = 0.0005$.

The improvement obtained by including two-sided information via the reduction scheme ranges from 54% in the case of positive correlation and $\delta = 0.0005$, up to a considerable 93% in the case of negative correlation and $\delta = 0.0001$. Overall, the improvement is more pronounced when negative correlation is prescribed. Moreover, the improvement is particularly strong for high levels of the confidence threshold α, except for the case of positive correlation and $\delta = 0.0005$. $\diamondsuit$

	$\rho = -0.5$			$\rho = 0$			$\rho = 0.5$		
α	i. standard (low : up)	scheme (low : up)	impr. %	i. standard (low : up)	scheme (low : up)	impr. %	i. standard (low : up)	scheme (low : up)	impr. %
0.95	(1.5 : 10.7)	(5.5 : 6.7)	86	(3.5 : 10.7)	(5.5 : 7.1)	77	(3.5 : 10.1)	(5.6 : 8.2)	61
0.99	(2.3 : 27)	(13.1 : 16.2)	87	(4.3 : 27)	(13.5 : 16.8)	85	(7.7 : 26.5)	(14.1 : 20.5)	66
0.995	(2.8 : 38)	(18.3 : 20.7)	93	(5.5 : 38)	(19.5 : 23.8)	87	(10.4 : 38)	(19.5 : 29.7)	63

Table 1: Improved Standard Bounds on VaR of $X_1 + X_2$ and VaR estimates via reduction schemes using $\underline{Q}^{\mathcal{D}_{\mathrm{KS}},\delta}$ and $\underline{Q}^{\mathcal{D}_{\mathrm{KS}},\delta}$ for $\delta = 0.0001$.

	$\rho = -0.5$			$\rho = 0$			$\rho = 0.5$		
α	i. standard (low : up)	scheme (low : up)	impr. %	i. standard (low : up)	scheme (low : up)	impr. %	i. standard (low : up)	scheme (low : up)	impr. %
0.95	(1.5 : 10.7)	(5.5 : 7.1)	82	(2.2 : 10.7)	(5.5 : 7.9)	72	(3.4 : 10.2)	(5.5 : 8.6)	54
0.99	(2.3 : 27)	(13.1 : 16.6)	85	(4.3 : 27)	(13.1 : 18.8)	74	(7.3 : 27)	(14 : 22)	60
0.995	(2.8 : 38)	(18.3 : 23.4)	85	(5.3 : 38)	(19.2 : 27)	76	(9.9 : 38)	(19.5 : 33.3)	50

Table 2: Improved Standard Bounds on VaR of $X_1 + X_2$ and VaR estimates via reduction schemes using $\underline{Q}^{\mathcal{D}_{\mathrm{KS}},\delta}$ and $\underline{Q}^{\mathcal{D}_{\mathrm{KS}},\delta}$ for $\delta = 0.0005$.

Example 5.2. We now consider an $\mathbb{R}^4$-valued risk vector (X_1, X_2, X_3, X_4) with copula C and Pareto$_2$-marginals. Moreover, we assume that

$$C^{\underline{\Sigma}} \le C \le C^{\overline{\Sigma}}$$

where $C^{\underline{\Sigma}}$ and $C^{\overline{\Sigma}}$ denote 4-dimensional Gaussian copulas with correlation matrices $\underline{\Sigma} = (\underline{\rho}_{ij})_{i,j=1,\dots,4}$ and $\overline{\Sigma} = (\overline{\rho}_{ij})_{i,j=1,\dots,4}$ respectively. Also, we assume that $\underline{\rho}_{ij} \le \overline{\rho}_{ij}$ for $i, j = 1, \dots, 4$, which by Slepian's Lemma guarantees non-degeneracy in the sense that $C^{\underline{\Sigma}} \le C^{\overline{\Sigma}}$; c.f. also Gupta, Eaton, Olkin, Perlman, Savage, and Sobel [12, Theorem 5.1].

This corresponds to a situation of correlation uncertainty which occurs naturally in many practical applications. Whenever correlation is estimated from data one obtains, rather than an exact estimate, a confidence interval for the pairwise correlations $(\underline{\rho}_{ij}, \overline{\rho}_{ij}) \subset [-1, 1]$, in which the parameters lie with a certain probability. Due to the radial symmetry of the Gaussian copula, the bounds $C^{\underline{\Sigma}}$ and $C^{\overline{\Sigma}}$ apply similarly to the survival copula $\widehat{C}$ (c.f. [11, Property 1]), i.e.

$$C^{\underline{\Sigma}}(\mathbf{1} - \cdot) \le \widehat{C} \le C^{\overline{\Sigma}}(\mathbf{1} - \cdot).$$

We then relate the bounds on C and $\widehat{C}$ respectively to the VaR of $X_1 + \cdots + X_4$, using our reduction schemes and again we compare the results to the Improved Standard Bounds obtained from $C^{\underline{\Sigma}}$ and $C^{\underline{\Sigma}}(\mathbf{1} - \cdot)$. Table 3 shows the results for different confidence levels α, assuming that $\underline{\Sigma}$ and $\overline{\Sigma}$ are equicorrelation matrices with correlation parameters $\underline{\rho}$ and $\overline{\rho}$ respectively. The VaR estimates were obtained by inverting $\underline{D}_{\varphi}^{\square}(5)$ and $\overline{D}_{\varphi}^{\square}(5)$ as well as $\widehat{\underline{D}}_{\varphi}^{\square}(5)$ and $\widehat{\overline{D}}_{\varphi}^{\square}(5)$ for

$\varphi(x_1, ..., x_4) = \mathbb{1}_{x_1+\cdots+x_4<s}$ along the variable s. Thus, we obtain two upper and two lower VaR estimates of which the largest lower bound and the lowest upper bound are reported. No further improvement of the bounds was obtained for $k > 5$. For the sake of legibility the results are rounded to full integers.

	$\underline{\rho} = -0.1, \overline{\rho} = 0.2$			$\underline{\rho} = 0.3, \overline{\rho} = 0.5$		
α	i. standard (low : up)	scheme (low : up)	impr. %	i. standard (low : up)	scheme (low : up)	impr. %
0.95	(3 : 32)	(8 : 26)	38	(1 : 30)	(7 : 29)	24
0.99	(9 : 74)	(20 : 52)	51	(2 : 74)	(18 : 63)	37
0.995	(13 : 104)	(26 : 70)	52	(3 : 104)	(25 : 86)	40

Table 3: Improved Standard Bounds on VaR of $X_1 + \cdots + X_4$ and VaR estimates computed via reduction schemes using $C^{\underline{\Sigma}}$ and $C^{\overline{\Sigma}}$.

The improvement of the spread reaches from 24% in the case of moderate positive correlation up to 52% in the case of low correlation. Moreover, the improvement is particularly pronounced for high levels of the confidence threshold α. $\Diamond$

A. Using information on the survival copula

In this section we show that the reduction schemes in Section 4.1 and 4.2 can be applied similarly when information on the survival copula is provided. Specifically, we assume that the copula of $\mathbf{X}$ is such that $\underline{\widehat{Q}} \leq \widehat{C} \leq \overline{\widehat{Q}}$ where $\widehat{C}$ is the survival-function of C and $\underline{\widehat{Q}}, \overline{\widehat{Q}}$ are quasi-survival functions. We hence consider the primal problems

$$\underline{\widehat{P}}_\varphi := \inf \left\{ \mathbb{E}_C[\varphi] \colon C \in \mathcal{C}^d, \underline{\widehat{Q}} \leq \widehat{C} \leq \overline{\widehat{Q}} \right\}, \tag{A.1}$$

$$\overline{\widehat{P}}_\varphi := \sup \left\{ \mathbb{E}_C[\varphi] \colon C \in \mathcal{C}^d, \underline{\widehat{Q}} \leq \widehat{C} \leq \overline{\widehat{Q}} \right\}. \tag{A.2}$$

Note that due to

$$\widehat{C}(F_1(u_1), ..., F_d(u_d)) = \mathbb{P}(X_1 > u_1, ..., X_d > u_d) = \mathbb{P}(-X_1 < -u_1, ..., -X_d < -u_d),$$

for all $\mathbf{u} \in \mathbb{I}^d$, the condition $\underline{\widehat{Q}} \leq \widehat{C} \leq \overline{\widehat{Q}}$ is equivalent to $\underline{Q} \leq C_{-\mathbf{X}} \leq \overline{Q}$, where $\underline{Q}(\mathbf{u}) := \underline{\widehat{Q}}(\mathbf{1} - \mathbf{u})$, $\overline{Q}(\mathbf{u}) := \overline{\widehat{Q}}(\mathbf{1} - \mathbf{u})$ and $C_{-\mathbf{X}}$ is the copula of $-\mathbf{X}$. In particular, since $\underline{Q}$ and $\overline{Q}$ are quasi-copulas it follows from our Duality Theorem 3.4 and a transformation of variables that the sharp dual bound corresponding to $\underline{\widehat{P}}_\varphi$ is given by

$$\underline{\widehat{P}}_\varphi = \underline{\widehat{D}}_\varphi = \sup \left\{ \underline{\widehat{Q}}(h) - \overline{\widehat{Q}}(g^-) + \sum_{i=1}^{d} \mathbb{E}_i[\nu_i] \colon (h, g, \nu_1, ..., \nu_d) \in \underline{\widehat{\mathcal{A}}} \right\}, \tag{A.3}$$

where

$$\underline{\widehat{\mathcal{A}}} := \left\{ (h, g, \nu_1, ..., \nu_d) \colon \nu_i \in L(F_i),\ i = 1, ..., d,\ h, g \in \widehat{\mathcal{R}} \text{ s.t. } h - g^- + \sum_{i=1}^{d} \nu_i \leq \varphi \right\}.$$

and

$$\widehat{\mathcal{R}} := \left\{ h = \sum_{n=1}^{k} \alpha_n \widehat{\Lambda}_{\mathbf{u}^n} \colon k \in \mathbb{N}, \alpha_1, ..., \alpha_k \geq 0, \mathbf{u}^1, ..., \mathbf{u}^k \in \overline{\mathbb{R}}^d \right\},$$

for $\widehat{\Lambda}_{\mathbf{u}}$ of the form

$$\widehat{\Lambda}_{\mathbf{u}} \colon \mathbb{R}^d \ni (x_1, ..., x_d) \mapsto \mathbb{1}_{x_1 \geq u_1, ..., x_d \geq u_d}.$$

Moreover, we denote $\widehat{\Lambda}_{\mathbf{u}}^-(x_1, ..., x_d) := \mathbb{1}_{x_1 > u_1, ..., x_d > u_d}$ and h^- for $h \in \widehat{\mathcal{R}}$ is defined accordingly. Finally, for quasi-survival functions $\widehat{Q}$ and $h = \sum_{n=1}^{k} \alpha_n \widehat{\Lambda}_{\mathbf{u}^n} \in \mathcal{R} \in \widehat{\mathcal{R}}$ we define,

$$Q(h) := \sum_{n=1}^{k} \alpha_n \widehat{Q}\big(F_1(u_1^n), ..., F_d(u_d^n)\big); \quad Q(h^-) := \sum_{n=1}^{k} \alpha_n \widehat{Q}\big(F_1^-(u_1^n), ..., F_d^-(u_d^n)\big).$$

Analogously, the sharp dual bound associated to $\overline{\widehat{P}}_\varphi$ is equal to

$$\overline{\widehat{P}}_\varphi = \overline{\widehat{D}}_\varphi = \inf \left\{ \overline{\widehat{Q}}(h^-) - \underline{\widehat{Q}}(g) + \sum_{i=1}^{d} \mathbb{E}_i[\nu_i] \colon (h, g, \nu_1, ..., \nu_d) \in \overline{\widehat{\mathcal{A}}} \right\}, \tag{A.4}$$

where

$$\widehat{\widehat{\mathcal{A}}} := \left\{(h, g, \nu_1, ..., \nu_d) \colon \nu_i \in L(F_i),\ i = 1, ..., d,\ h, g \in \widehat{\mathcal{R}} \text{ s.t. } h^- - g + \sum_{i=1}^{d} \nu_i \geq \varphi\right\}.$$

Based on these dual characterizations the following corollaries establish the corresponding reduction schemes. Using the fact that

$$\mathbb{P}(\psi(X_1, ..., X_d) < s) = 1 - \mathbb{P}(\psi(X_1, ..., X_d) \geq s),$$

the proofs involve similar arguments as the proofs of Corollary 4.6 and 4.9 and therefore they are omitted. We denote the componentwise maximum of vectors $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathbb{R}^d$ by

$$\max(\mathbf{u}^1, ..., \mathbf{u}^k) = \Big(\max_{n=1,...,k}\{u_1^n\}, ..., \max_{n=1,...,k}\{u_d^n\}\Big).$$

Corollary A.1. *Let $\varphi(x_1, ..., x_d) = \mathbb{1}_{\psi(x_1,...,x_d)<s}$ for $\psi \colon \mathbb{R}^d \to \mathbb{R}$ increasing in each coordinate and let*

$$\begin{aligned}\widehat{\widehat{D}}{}^{\square}_{\varphi}(k) := \inf\Bigg\{1 - &\sum_{\mathbf{u}\in\widehat{\mathcal{M}}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \,\underline{\widehat{Q}}\big(F(\mathbf{u})\big) \\ &- \sum_{\mathbf{u}\in\widehat{\mathcal{M}}^e(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \,\widehat{\overline{Q}}\big(F(\mathbf{u})\big) \colon \mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi^c(s)\Bigg\},\end{aligned} \tag{A.5}$$

where

$$\begin{aligned}\widehat{\mathcal{M}}^o(\mathbf{u}^1, ..., \mathbf{u}^k) &:= \{\max(\mathbf{u}^{i_1}, ..., \mathbf{u}^{i_m}) \colon 0 \leq i_1 \leq \cdots \leq i_m \leq k \;\; m \textit{ odd}\},\\ \widehat{\mathcal{M}}^e(\mathbf{u}^1, ..., \mathbf{u}^k) &:= \{\max(\mathbf{u}^{i_1}, ..., \mathbf{u}^{i_m}) \colon 0 \leq i_1 \leq \cdots \leq i_m \leq k \;\; m \textit{ even}\}.\end{aligned}$$

and

$$\begin{aligned}l^o(\mathbf{u}) &:= |\{(i_1, ..., i_m) \colon 0 \leq i_1 \leq \cdots \leq i_m \leq k,\ m \textit{ odd}, \mathbf{u} = \max(\mathbf{u}^{i_1}, ..., \mathbf{u}^{i_m})\}|,\\ l^e(\mathbf{u}) &:= |\{(i_1, ..., i_m) \colon 0 \leq i_1 \leq \cdots \leq i_m \leq k,\ m \textit{ odd}, \mathbf{u} = \max(\mathbf{u}^{i_1}, ..., \mathbf{u}^{i_m})\}|,\end{aligned} \tag{A.6}$$

for $\mathbf{u} \in \widehat{\mathcal{M}}^o(\mathbf{u}^1, ..., \mathbf{u}^k)$ and $\mathbf{u} \in \widehat{\mathcal{M}}^e(\mathbf{u}^1, ..., \mathbf{u}^k)$ respectively. Then it holds that

$$\widehat{\widehat{D}}{}^{\square}_{\varphi}(k) \geq \widehat{\widehat{D}}{}^{\square}_{\varphi}(k+1) \geq ... \geq \widehat{\widehat{D}}_{\varphi}.$$

Remark A.2. Corollary A.1 extends the upper Improved Standard Bound from Embrechts et al. [9] presented in Theorem **??** in the sense that

$$\inf_{h\in\widehat{\mathcal{R}}} 1 - \underline{\widehat{Q}}(h) = \inf_{\mathbf{u}\in\mathcal{U}_\psi^c(s)} 1 - \underline{\widehat{Q}}(F(\mathbf{u})) = \overline{M}_{\underline{\widehat{Q}},\psi}(s).$$

The following corollary establishes a similar reduction scheme for $\underline{\widehat{D}}$. To this end, let us denote for $\mathbf{u}^1, ..., \mathbf{u}^k \in \mathbb{R}^d$ and an enumeration $\{\mathbf{i}^1, ..., \mathbf{i}^{dk}\}$ of $\{1, ..., d\}^k$

$$\widehat{\mathrm{U}}_{\mathbf{i}^n} := \max(\widehat{\mathrm{pr}}_{i_1}(\mathbf{u}^1), ..., \widehat{\mathrm{pr}}_{i_k}(\mathbf{u}^k)) \quad \text{for } (i_1, ..., i_k) = \mathbf{i}^n, n = 1, ..., dk,$$

where $\widehat{\mathrm{pr}}_i(\mathbf{u}) := (-\infty, ..., -\infty, u_i, -\infty, ..., -\infty)$ for $i \in \{1, ..., d\}$.

Corollary A.3. *Let* $\varphi(x_1, ..., x_d) = \mathbb{1}_{\psi(x_1,...,x_d)<s}$ *for* $\psi\colon \mathbb{R}^d \to \mathbb{R}$ *increasing in each coordinate and let*

$$\underline{\widehat{D}}^{\square}_{\varphi}(k) := \sup\Bigg\{1 - \sum_{\mathbf{u}\in\widehat{\mathcal{W}}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^o(\mathbf{u}) - l^e(\mathbf{u})\big)^+ \,\widehat{\overline{Q}}\big(F(\mathbf{u})\big) - \sum_{\mathbf{u}\in\widehat{\mathcal{W}}^o(\mathbf{u}^1,...,\mathbf{u}^k)} \big(l^e(\mathbf{u}) - l^o(\mathbf{u})\big)^+ \,\underline{\widehat{Q}}\big(F(\mathbf{u})\big) \colon \mathbf{u}^1, ..., \mathbf{u}^k \in \mathcal{U}_\psi(s)\Bigg\}, \tag{A.7}$$

where

$$\begin{aligned} \widehat{\mathcal{W}}^o(\mathbf{u}^1, ..., \mathbf{u}^k) &:= \widehat{\mathcal{M}}^o(\widehat{\mathrm{U}}_{\mathbf{i}^1}, ..., \widehat{\mathrm{U}}_{\mathbf{i}^{dk}}), \\ \widehat{\mathcal{W}}^e(\mathbf{u}^1, ..., \mathbf{u}^k) &:= \widehat{\mathcal{M}}^e(\widehat{\mathrm{U}}_{\mathbf{i}^1}, ..., \widehat{\mathrm{U}}_{\mathbf{i}^{dk}}), \end{aligned} \tag{A.8}$$

for an enumeration $\{\mathbf{i}^1, ..., \mathbf{i}^{dk}\}$ *of* $\{1, ..., d\}^k$ *and* l^o, l^e *are given in* (A.6)*. Then it holds that*

$$\underline{\widehat{D}}^{\square}_{\varphi}(k) \geq \underline{\widehat{D}}^{\square}_{\varphi}(k+1) \geq ... \geq \underline{\widehat{D}}_{\varphi}.$$

References

[1] D. Bartl, P. Cheredito, M. Kupper, and L. Tangpi. Duality for increasing convex functionals with countably many marginal constraints. Preprint, arXiv:1509.08988v2, 2015.

[2] M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices: A mass transport approach. *Finance Stoch.*, 17:477–501, 2013.

[3] C. Bernard and S. Vanduffel. A new approach to assessing model risk in high dimensions. *J. Banking Finance*, 58:166–178, 2015.

[4] V. Bignozzi, G. Puccetti, and L. Rüschendorf. Reducing model risk using positive and negative dependence assumptions. *Insurance Math. Econ.*, 61:17–26, 2015.

[5] E. Bouyé, V. Durrleman, A. Nikeghbali, G. Riboulet, and T. Roncalli. Copulas: an open field for risk management. working paper, ssrn:1032557, 2001.

[6] M. Cambou and Filipović. Model uncertainty and scenario aggregation. *Math. Finance*, 2015. forthcoming, doi:10.1111/mafi.12097.

[7] M. Denuit, C. Genest, and É. Marceau. Stochastic bounds on sums of dependent risks. *Insurance Math. Econom.*, 25:85–104, 1999.

[8] P. Embrechts and G. Puccetti. Bounds for functions of dependent risks. *Finance Stoch.*, 10:341–352, 2006.

[9] P. Embrechts, A. Höing, and A. Juri. Using copulae to bound the value-at-risk for functions of dependent risks. *Finance Stoch.*, 7:145–167, 2003.

[10] P. Embrechts, G. Puccetti, and L. Rüschendorf. Model uncertainty and VaR aggregation. *J. Banking Finance*, 37:2750–2764, 2013.

[11] P. Georges, A.-G. Lamy, E. Nicolas, G. Quibel, and T. Roncalli. Multivariate survival modelling: a unified approach with copulas. Preprint, ssrn:1032559, 2001.

[12] S. D. Gupta, M. L. Eaton, I. Olkin, M. Perlman, L. J. Savage, and M. Sobel. Inequalitites on the probability content of convex regions for elliptically contoured distributions. In *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, pages 241–265. Univ. of Calif. Press, 1972.

[13] A. Lindner and A. Szimayer. A limit theorem for copulas. technical report, Technical University Munich, 2005.

[14] J. A. D. Loera, R. Hemmecke, and M. Köppe. *Algebraic and Geometric Ideas in the Theory of Discrete Optimization*. MOS-SIAM, 2013.

[15] T. Lux and A. Papapantoleon. Improved Fréchet–Hoeffding bounds on d-copulas and applications in model-free finance. Preprint, arXiv:1602.08894v3, 2015.

[16] T. Lux and A. Papapantoleon. Model-free bounds on value-at-risk using partial dependence information. Preprint, arXiv:1610.09734, 2016.

[17] R. B. Nelsen. *An Introduction to Copulas*. Springer, 2nd edition, 2006.

[18] G. Puccetti and L. Rüschendorf. Computation of sharp bounds on the distribution of a function of dependent risks. *J. Comput. Appl. Math.*, 236:1833–1840, 2012.

[19] G. Puccetti and L. Rüschendorf. Bounds for joint portfolios of dependent risks. *Statistics and Risk Modeling*, 29:107–132, 2012.

[20] G. Puccetti and L. Rüschendorf. Sharp bounds for sums of dependent risks. *J. Appl. Probab.*, 50: 42–53, 2013.

[21] G. Puccetti, L. Rüschendorf, and D. Manko. VaR bounds for joint portfolios with dependence constraints. Preprint, ssrn:2794873, 2016.

[22] L. Rüschendorf and N. Gaffke. On a class of extremal problems in statistics. *Math. Operationsforschung Statist.*, 12:123–135, 1981.

[23] L. Rüschendorf. Sharpness of Fréchet-bounds. *Z. Wahrsch. Verw. Gebiete*, 57:293–302, 1981.

[24] L. Rüschendorf. Monge-Kantorovich transportation problem and optimal couplings. *Jahresbericht der DMV*, 109:113–137, 2007.

[25] L. Rüschendorf. Risk bounds and partial dependence information. Preprint, Univ. Freiburg, 2016.

[26] A. Syropoulos. Mathematics of multisets. *Lecture Notes in Computer Science*, 2235:347–358, 2001.

[27] P. Tankov. Improved Fréchet bounds and model-free pricing of multi-asset options. *J. Appl. Probab.*, 48:389–403, 2011.

[28] F. Terkelsen. Some minmax theorems. *Math. Scand.*, 31:405–413, 1972.

[29] C. Villani. *Optimal Transport. Old and New*. Springer, Heidelberg, 2009.

[30] B. Wang and R. Wang. The complete mixability and convex minimization problems with monotone marginal densities. *J. Multivariate Anal.*, 102:1344–1360, 2011.

[31] R. C. Williamson and T. Downs. Numerical methods for calculating convolutions and dependency bounds. *Internat. J. Approx. Reason.*, 4:89–158, 1990.