Ordering risk bounds in factor models

Jonathan Ansari, Ludger Rüschendorf

October 1, 2018

Abstract

Conditionally comonotonic risk vectors have been proved in Bernard et al. (2017) to yield worst case dependence structures maximizing the risk of the portfolio sum in partially specified risk factor models. In this paper we investigate the question how risk bounds depend on the specification of the pairwise copulas of the risk components X_i with the systemic risk factor. As basic tool we introduce a new ordering based on sign changes of the derivatives of copulas. This together with discretization by n-grids and the theory of supermodular transfers allows us to derive concrete ordering criteria for the maximal risks.

Keywords products of copulas, supermodular ordering, risk bounds, conditionally comonotonic distributions, mass transfer theory, elliptical distributions, Archimedean copulas

1 Introduction

In recent years a lot of effort has been undertaken to base the evaluation of risk bounds for the joint portfolio $S = \sum_{i=1}^d X_i$ of a risk vector $X = (X_1, \dots, X_d)$ on reliable information on the marginals F_i of X_i and on the joint dependence structure of X. Considering law-invariant convex risk measures Ψ it is well-known that Ψ is consistent with respect to the convex order, i.e.

$$S_1 \leq_{cx} S_2 \quad \Longrightarrow \quad \Psi(S_1) \leq \Psi(S_2) \tag{1}$$

assuming generally that $S_i \in L^1(P)$ are integrable and defined on a non-atomic measure space $(\Omega, \mathcal{A}, P)$. Thus it is sufficient to determine (sharp) upper bounds w.r.t. $\leq_{cx}$ in order to determine (sharp) upper risk bounds for $\sum_{i=1}^d X_i$.

In the case that there is only marginal information but no further dependence information on the risk vector X available, an upper bound for the joint portfolio $S = \sum_{i=1}^d X_i$ in convex order is given by the comonotonic sum $S^c = \sum_{i=1}^d X_i^c = \sum_{i=1}^d F_i^{-1}(U)$ with $U \sim U(0,1)$ uniformly distributed on $(0,1)$,

$$S \leq_{cx} S^c . \tag{2}$$

For many applications, the comonotonic upper bound $\Psi(S^c)$ of the risk $\Psi(S)$ is too wide to be useful. Therefore, in recent years various approaches have been investigated to introduce additional dependence information and structural information in order to tighten the risk bounds.

A promising approach in this direction, the *partially specified risk factor models*, have been introduced in Bernard et al. (2017). It is assumed in this approach that the risk vector X is described by a factor model

$$X_i = f_i(Z, \varepsilon_i)\,, \quad 1 \leq i \leq d\,,$$

for functions f_i , where Z is a systemic risk factor and ε_i are individual risk factors. It is assumed that the joint distributions H_i of (X_i, Z) , $1 \leq i \leq d$, are known. The joint distributions of (ε_i) and Z however are not specified, in contrast to the usual independence assumption in factor models. This means that both the copulas $C_{X_i,Z}$ of (X_i, Z) and the marginal distributions of $X_i \sim F_i$ and $Z \sim G$ are known, but the dependence structure of $(X_1, \dots, X_d)|Z = z$ is not specified.

The common systemic risk factor Z however can be used to reduce the dependence uncertainty (DU). It has been shown in Bernard et al. (2017, Proposition 3.2) that in the partially specified risk factor model a sharp upper bound in convex order is given by the conditionally comonotonic sum, i.e. for $U \sim U(0,1)$ independent of Z holds

$$S \leq_{cx} S_Z^c := \sum_{i=1}^{d} F_{X_i|Z}^{-1}(U)\,. \tag{3}$$

Furthermore, S_Z^c is an improvement of the comonotonic sum S^c , i.e.

$$S_Z^c \leq_{cx} S^c\,. \tag{4}$$

In this paper, we assume that Z is a real-valued random variable. Then, the upper bound S_Z^c depends only on the specified marginals F_i and G and on the bivariate copulas $C^i = C_{X_i,Z} \in \mathcal{C}_2$, where $\mathcal{C}_d$ denotes the set of d-copulas. The conditionally comonotonic sum S_Z^c thus solves the optimization problem

$$S_Z^c = \max\left\{\sum_{i=1}^{d} X_i\,,\ X_i \sim F_i\,,\ Z \sim G\,,\ C_{X_i,Z} = C^i\right\}, \tag{5}$$

where the max is w.r.t. convex order $\leq_{cx}$.

In the following, we investigate how the solution in (5) varies in dependence on the constraints C^i . More generally, we aim to determine criteria for copula classes $\mathcal{S}^i \subset \mathcal{C}_2$ of bivariate copulas and classes $\mathcal{F}_i$ of univariate distribution functions such that a solution of the maximization problem

$$\max\left\{\sum_{i=1}^{d} X_i\,,\ X_i \sim F_i\,,\ Z \sim G\,,\ C_{X_i,Z} \in \mathcal{S}^i\right\} \quad \text{w.r.t. } \leq_{cx} \tag{6}$$

exists and can be determined for all $F_i \in \mathcal{F}_i$ and for all continuous distribution functions G. Equivalently, maximization problem 6 can be formulated as

$$\max\left\{\sum_{i=1}^d f_i(U_i)\,,\ U_i\sim U(0,1)\,,\ Z\sim G\,,\ C_{U_i,Z}\in\mathcal{S}^i\right\}\quad \text{w.r.t. } \leq_{cx} \tag{7}$$

for classes $\mathcal{S}^i$ of copulas, transformation functions $f_i\in\mathcal{G}_i=\{F_i^{-1}|F_i\in\mathcal{F}_i\}$ and continuous distribution functions G.

After the problem formulation and motivation we introduce in Section 2 the upper product of bivariate copulas which describes the dependence structure of conditionally comonotonic random vectors. We develop several tools for approximation of these products. In particular, we deal with the approximation by n-grid copulas. In Section 3 we reduce ordering properties of the portfolio sums in partially specified factor models by approximation to ordering properties on n-grid models. As a basic new tool, we introduce the ordering $\leq_{\partial\Delta}$ of sign changes of the derivatives of the copulas. In our main result, Theorem 3.10, we show that the $\leq_{\partial\Delta}$-ordering is sufficient for ordering upper products and thus for ordering risk bounds in factor models. For the ordering of the n-grid copulas we make essential use of the ordering results by mass transfer theory as developed in Müller (2013). The partially quite technical proofs are deferred to the appendix. In Section 4, we give an application to financial data. We improve the standard DU interval for the Average Value-at-Risk of a portfolio of European options on different assets by up to 30%.

2 The upper product of bivariate copulas

A d-copula is a distribution function on the d-dimensional unit cube $[0,1]^d$ with uniform univariate margins. Due to Sklar's Theorem, the distribution of a random vector can be separated in its univariate margins and a copula which completely describes the dependence structure of the random vector. Some specific copulas that we need in the following are the copulas $M^d\in\mathcal{C}_d$, $\Pi^d\in\mathcal{C}_d$ and $W^2\in\mathcal{C}_2$ which model comonotonicity, independence and countermonotonicity, respectively. For an introduction to copulas, we refer to Nelsen (2006).
Since the univariate margins are fixed, a solution of (5) and (6) only depends on the copula of $(X_1,\dots,X_d)$. Varying the solution in dependence on the constraints $C^i\in\mathcal{C}_2$ motivates to introduce upper products of bivariate copulas which are the copulas of conditionally comonotonic distributions.

For a family $\mathbf{C}=\{C_t\}_{t\in[0,1]}\subset\mathcal{C}_2$ of bivariate copulas and $A,B\in\mathcal{C}_2$, Durante et al. (2007) define the $\mathbf{C}$-product $A*_{\mathbf{C}}B\colon[0,1]^2\to[0,1]$ of A and B through

$$(A*_{\mathbf{C}}B)(u_1,u_2):=\int_0^1 C_t\left(\frac{\partial}{\partial t}A(u_1,t),\frac{\partial}{\partial t}B(t,u_2)\right)\mathsf{d}t \tag{8}$$

which again is a bivariate copula (see Durante et al. (2007, Proposition 3.1).

In the case that $C_t = \Pi^2$ for all t, where Π^2 denotes the bivariate independence copula, there is a correspondence of the $\mathbf{C}$-product with Markov processes (see Darsow et al. (1992, Theorem 3.2 and Theorem 3.3). For our purposes, we are interested in a d-dimensional extension of the case that $C_t = M^2$ for all t.

An extension of (8) to the case of d-fold products as needed in the partially specified factor model is given as follows.

Proposition 2.1 *Let $\mathbf{C} = \{C_t\}_{t\in[0,1]} \subset \mathcal{C}_d$ be a family of d-copulas. Then, for $C^1, \ldots, C^d \in \mathcal{C}_2$, the $\mathbf{C}$-product $*_{\mathbf{C}}(C^1, \ldots, C^d)$ given through*

$$*_{\mathbf{C}}(C^1, \ldots, C^d)(u_1, \ldots, u_d) := \int_0^1 C_t\left(\partial_2 C^1(u_1, t), \ldots, \partial_2 C^d(u_d, t)\right) \mathsf{d}t$$

for $(u_1, \ldots, u_d) \in [0,1]^d$ is a d-copula, where ∂_2 denotes the first partial derivative with respect to the second argument.

Proof: Let $U_i, Z \sim U(0,1)$ with $C_{U_i,Z} = A^i$. Then,

$$\int_0^v \partial_2 A^i(u_i, t)\, \mathsf{d}t = A^i(u_i, v) = P(U_i \leq u_i, Z \leq v) = \int_0^v F_{U_i|Z=t}(u_i)\, \mathsf{d}t \tag{9}$$

Since $F_{U_i|Z=t}$ can be considered as a distribution function for all t, Lebesgue's differential theorem shows that $\partial_2 A^i(u_i, t) = F_{U_i|Z=t}(u_i)$, and $\partial_2 A^i(\cdot\,, t)$ can also be considered as a distribution function for almost all t. From Sklar's Theorem it follows that $C_t(\partial_2 C^1(\cdot\,, t), \ldots, \partial_2 C^d(\cdot\,, t))$ defines a distribution function for almost all t. Thus also the mixture $*_{\mathbf{C}}(C^1, \ldots, C^d)$ defines a distribution function. Since $\partial_2 C^i(1, t) = 1$ for all t, $*_{\mathbf{C}}(C^1, \ldots, C^d)$ has uniform margins and thus is a d-copula. ■

Note that copulas are almost surely partially differentiable (see Nelsen (2006, Theorem 2.2.7)) and the integral is defined as a Lebesgue-integral.

Remark 2.2 *Let $(X_1, \ldots, X_d, Z)$ be a $(d+1)$-dimensional random vector such that F_Z is continuous. Then, from Sklar's Theorem, the transformation formula and Proposition 2.1 it follows that*

$$(X_1, \ldots, X_d) \sim *_{\mathbf{C}}(C^1, \ldots, C^d)(F_{X_1}, \ldots, F_{X_d}),$$

where $C^i = C_{X_i,Z}$ and $\mathbf{C} = \{C_t\}_{t\in(0,1)}$ for $C_t = C_{X_1,\ldots,X_d|Z=F_Z^{-1}(t)}$ being the copula of the conditional vector $(X_1, \ldots, X_d)|Z = F_Z^{-1}(t)$.

2.1 Definition of the upper product and elementary properties

For application to risk bounds in partially specified risk factor models we consider the special case of the $*_{\mathbf{C}}$-product with $\mathbf{C} = \mathbf{M^d} := \{M^d\}_{0\leq t\leq 1}$ leading to the notion of the *upper product.* Due to Proposition 2.1 the operator $\bigvee$ in the following definition is well-defined.

Definition 2.3 (Upper product) *The* upper product $\bigvee$ *of bivariate copulas* $C^1, \ldots, C^d$ *is defined through* $\bigvee_{i=1}^d C^i := C^1 \vee \cdots \vee C^d := *_{\mathbf{M}^d}(C_1, \ldots, C^d)$ *, i.e.*

$$\bigvee_{i=1}^{d} C^i(u_1, \ldots, u_d) := \int_0^1 \min_{1 \le i \le d} \left\{ \partial_2 C^i(u_i, t) \right\} \mathsf{d}t$$

for all $(u_1, \ldots, u_d) \in [0,1]^d$ *.*

The following proposition gives some elementary properties of the upper product. Point (i) explains the choice of the name "upper" product. Point (ii) explains that the upper product describes the case of conditionally comonotonic copulas and thus gives the connection to risk bounds in partially specified factor models (see also Remark 2.5 (a)).

Proposition 2.4 *For* $\mathbf{C} = \{C_t\}_{t \in (0,1)} \subset \mathcal{C}_d$ *, for* $A^1, \ldots, A^d, D \in \mathcal{C}_2$ *and for a random vector* $(U_1, \ldots, U_d)$ *on* $(\Omega, \mathcal{A}, P)$ *holds:*

(i) $*_{\boldsymbol{C}}(A^1, \ldots, A^d) \le_{sm} \bigvee_{i=1}^d A^i$ *.*

(ii) $U = (U_1, \ldots, U_d) \sim \bigvee_{i=1}^d A^i \iff \exists Z \sim U(0,1)$ *and* $V = (V_1, \ldots, V_d)$ *such that* $V \stackrel{\mathrm{d}}{=} U$ *,* $C_{V_i, Z} = A^i$ *and* $V | Z = z$ *is comonotonic for all* z *.*

(iii) In general, the upper product is neither commutative nor associative.

(iv) Marginalization property: For $J \subset (1, \ldots, d)$ *, the* J*-margin of* $\bigvee_{i=1}^d A_i$ *is given by* $\bigvee_{i \in J} A^i$ *.*

(v) $\bigvee_{i=1}^d A^i = M^d$ *if and only if* $A^i = A^j$ *for all* $i \ne j$ *.*

(vi) $D \vee M^2 = D$ *and* $M^2 \vee D = D^*$ *, where* $D^*(u, v) = D(v, u)$ *.*

(vii) $D \vee W^2(u, v) = D(u, 1 - v)$ *and* $W^2 \vee D(u, v) = D^*(1 - u, v)$ *.*

(viii) $\overline{A^1 \vee \cdots \vee A^d}(u) = 1 - \int \max_i \{\partial_2 A^i(u_i, t)\} \, \mathsf{d}t$ *, where* $\overline{F}$ *denotes the survival function of a distribution function* F *.*

Proof: (i) follows from $C_t \le_{sm} M^d$ (see Tchen (1980, Theorem 5) or Rüschendorf (1983, Corollary 3a)) and the closure of the supermodular ordering under mixtures (see Shaked and Shanthikumar (2007, Theorem 9.A.9.(d))).
(ii): Assume that $(U_1, \ldots, U_d) \sim \bigvee_i A^i$. Consider bivariate random vectors $(W_i, Z) \sim A^i$. Define $V_i := F^{-1}_{W_i|Z}(\xi)$ for $\xi \sim U(0,1)$ independent of Z . Then, $V|Z = z$ is comonotonic for all z by construction. Further, it holds that $V_i | Z = z \sim F_{W_i | Z = z}$. But this means that

$$\begin{aligned} P(V_i \le u_i, Z \le z) &= \int_0^z P(V_i \le u_i | Z = t) \, \mathsf{d}t \\ &= \int_0^z F_{W_i | Z = t}(u_i) \, \mathsf{d}t = A^i(u_i, z) \, . \end{aligned}$$

Further, we obtain

$$\begin{aligned} P(V_i \leq u_i \ \forall i) &= \int_0^1 P(F^{-1}_{W_i|Z}(U) \leq u_i \ \forall i \ |Z = t) \, \mathsf{d}t \\ &= \int_0^1 P(U \leq F_{W_i|Z}(u_i) \ \forall i \ |Z = t) \, \mathsf{d}t \\ &= \int_0^1 \min_i \left\{ F_{W_i|Z=t}(u_i) \right\} \, \mathsf{d}t \qquad (10) \\ &= \int_0^1 \min_i \left\{ \partial_2 A^i(u_i, t) \right\} \, \mathsf{d}t = \bigvee_{i=1}^d A^i(u_1, \ldots, u_d) \, , \end{aligned}$$

where the fourth equality holds with an argument as in (9). Hence, it holds $V \stackrel{\mathrm{d}}{=} U$. The reverse direction follows from the equations in (10).
(iii): From (vi) follows that $\bigvee$ is not commutative if D is not symmetric.
(iv): Let $u \in [0,1]^d$ with $u_i = 1$ for all $i \notin J$. Then,

$$\begin{aligned} \bigvee_{i=1}^d A^i \ (u) &= \int_0^1 \min_{1 \leq i \leq d} \left\{ \partial_2 A^i(u_i, t) \right\} \, \mathsf{d}t \\ &= \int_0^1 \min_{i \in J} \left\{ \partial_2 A^i(u_i, t) \right\} \, \mathsf{d}t = \bigvee_{i \in J} A^i(u_J) \end{aligned}$$

where $u_J := (u_{i_1}, \ldots, u_{i_k})$ for $J = (i_1, \ldots, i_k)$.
(v): If $A^i = A^j$ for all $i \neq j$, then

$$\begin{aligned} \bigvee_i A^i(u_1, \ldots, u_d) &= \int_0^1 \min_i \left\{ \partial_2 A^1(u_i, t) \right\} \, \mathsf{d}t \\ &= \int_0^1 \partial_2 A^1(\min_i \{u_i\}, t) \, \mathsf{d}t = \min_i \{u_i\} \end{aligned}$$

for all $(u_1, \ldots, u_d) \in [0,1]^d$.
Assume without loss of generality that $A^1 \neq A^2$. Due to the continuity of copulas there exist $(v_1, v_2) \in (0,1)^2$ and $\varepsilon > 0$ such that $\partial_2 A^1(u,t) > \partial_2 A_2(u,t)$ for all $(u,t) \in B_\varepsilon((v_1, v_2)) \subset (0,1)^2$. This yields for $u_1 = u_2 = u$ that

$$\begin{aligned} M^2(u_1, u_2) = u_1 &= \int_0^1 \partial_2 A^1(u_1, t) \, \mathsf{d}t \\ &> \int_0^1 \min \left\{ \partial_2 A^1(u_1, t), \partial_2 A^2(u_2, t) \right\} \, \mathsf{d}t \\ &= A^1 \vee A^2(u_1, u_2) \, . \end{aligned}$$

Then, the assertion follows from (iv).

(vi) and (vii): For all $(u_1,u_2)\in[0,1]^2$ holds

$$\begin{aligned} D\vee M^2(u_1,u_2) &= \int_0^1 \min\left\{\partial_2 D(u_1,t), \mathbb{1}_{\{u_2\geq t\}}\right\} \mathrm{d}t \\ &= \int_0^{u_2} \partial_2 D(u_1,t)\,\mathrm{d}t = D(u_1,u_2)\,. \end{aligned}$$

The other cases follow similarly.
(viii): Due to (ii) assume that $(U_1,\ldots,U_d)\sim\bigvee_i A^i$ and $Z\sim U(0,1)$ such that $(U_1,\ldots,U_d)|Z=t$ is comonotonic for all t. Then, we obtain

$$\begin{aligned} \overline{A^1\vee\cdots\vee A^d}(u) &= P(U_i>u_i\ \forall i) \\ &= \int_0^1 P(U_i>u_i\ \forall i\ |Z=t)\,\mathrm{d}t \\ &= \int_0^1 \min_i\left\{P(U_i>u_i|Z=t)\right\}\,\mathrm{d}t \\ &= 1-\int_0^1 \max\left\{P(U_i\leq u_i|Z=t)\right\}\,\mathrm{d}t \\ &= 1-\int_0^1 \max\left\{\partial_2 A^i(u_i,t)\right\}\,\mathrm{d}t\,, \end{aligned}$$

where the third equality holds due to the conditional comonotonicity. ■

Remark 2.5 *(a) From Proposition 2.4* (ii) *and Sklar's Theorem it follows that for $Z\sim U(0,1)$ the upper product describes the dependence structure of the solution of* (5)*, i.e.*

$$(F^{-1}_{X_i|Z}(U))_{1\leq i\leq d}\sim\bigvee_{i=1}^d C^i(F_1,\ldots,F_d) \tag{11}$$

for $X_i\sim F_i$ and $C_{X_i,Z}=C^i$. More generally, applying the transformation formula yields that (11) *holds true for all Z with continuous distribution function $G=F_Z$.*

(b) The continuity of G is decisive for (11)*. Assume for example that G follows a Dirac distribution. Then, any arbitrary copula C^i describes the dependence structure of (X_i,Z), $Z\sim G$, and hence, knowledge of $C_{X_i,Z}$ is no information. Thus, the worst case distribution in* (5) *must be given through the comonotonic random vector X^c (which coincides with X^c_Z in this case). But from Proposition 2.4* (v) *we obtain that $\bigvee_i C^i\neq M^d$ if not all C^i coincide.*

(c) Point (v) *of Proposition 2.4 induces that the upper product should take pointwise large values if the arguments are close to each other.*

The definition of the upper product yields an invariance property under Lebesgue-measure preserving transformations of the integrand.

Let λ be the Lebesgue measure on $\mathcal{B}([0,1])$. Denote by $\mathcal{T}$ the set of measurable transformations $T\colon ((0,1),\mathcal{B}((0,1)),\lambda) \to ((0,1),\mathcal{B}((0,1)),\lambda)$ that are *measure preserving*, i.e. $T*\lambda = \lambda$, where $T*\lambda(A) := \lambda(T^{-1}(A))$ for all $A \in \mathcal{B}((0,1))$ denotes the distribution of the image of λ under T. Let $\mathcal{T}_P$ be the set of all $T \in \mathcal{T}$ such that T is bijective and its inverse T^{-1} is measure preserving. Then, elements of $\mathcal{T}_P$ are denoted *shuffles*, see Durante and Sánchez (2012).

The following statement shows that the upper product is invariant under joint shuffles of the factor variable.

Proposition 2.6 *For all $T \in \mathcal{T}_P$ and $C \in \mathcal{C}_2$, the function $\mathcal{S}_T(C)\colon [0,1]^2 \to [0,1]$ given through*

$$\mathcal{S}_T(C)(u,v) := \int_0^v \partial_2 C(u, T^{-1}(t))\,\mathsf{d}t$$

is a bivariate copula. Furthermore, it holds that

$$\bigvee_{i=1}^d C^i = \bigvee_{i=1}^d \mathcal{S}_T(C^i)\,.$$

Proof: For $f_1, f_2 \in \mathcal{T}$ define the function $C_{f_1,f_2}\colon [0,1]^2 \to [0,1]$ through

$$C_{f_1,f_2}(u_1,u_2) := \lambda(f_1^{-1}([0,u_1]) \cap f_2^{-1}([0,u_2]))\,.$$

Let μ_C be the probability measure induced by C and denote by K_C the corresponding Markov kernel such that $\mu_C(\mathsf{d}s, \mathsf{d}t) = K_C(\mathsf{d}s, t)\,\mathsf{d}t$. Then, from the disintegration theorem it follows that $\partial_2 C(u,s) = K_C([0,u],s)$ almost surely. Denote by $(g_1,g_2) \in \mathcal{T} \times \mathcal{T}$ the measure-preserving decomposition of C according to Kolesárová et al. (2008, Theorem 3.1) such that $C_{g_1,g_2} = C$. Then, for all $(u,v) \in [0,1]^2$ holds

$$\begin{aligned}
\mathcal{S}_T(C)(u,v) &= \int_{[0,v]} \partial_2 C(u, T^{-1}(t))\,\mathsf{d}\lambda(t) = \int_{[0,v]} \partial_2 C(u, T^{-1}(t))\,\mathsf{d}\lambda^T(t) \\
&= \int_{T^{-1}([0,v])} \partial_2 C(u,s)\,\mathsf{d}\lambda(s) = \int_{T^{-1}([0,v])} K_C([0,u],s)\,\mathsf{d}s \\
&= \mu_C([0,u], T^{-1}([0,v])) = (g_1,g_2) * \lambda\left([0,u] \cap T^{-1}([0,v])\right) \\
&= (g_1, T\circ g_2) * \lambda([0,u] \cap [0,v]) \\
&= \lambda\left(g_1^{-1}([0,u]) \cap (T\circ g_2)^{-1}([0,v])\right) \\
&= C_{g_1, T\circ g_2}(u,v)\,,
\end{aligned}$$

where the second equality is true because T is λ-preserving, the third equality holds by the transformation formula, the fifth equality holds due to the disintegration theorem. The sixth equality holds because $C = C_{g_1,g_2}$. From Kolesárová et al. (2008, Theorem 3.1) we also get that $C_{g_1,T\circ g_2}$ defines a copula because $T \circ g_2$ is measure preserving. This proves the first statement.

Since $\mathcal{S}_T(C^i) \in \mathcal{C}_2$ for all i, the upper product $\bigvee \mathcal{S}_T(C^i)$ is well-defined.

Hence, the second statement follows from

$$\begin{aligned}\bigvee_{i=1}^{d} \mathcal{S}_T(C^i)\ (u_1,\dots,u_d) &= \int_0^1 \min\left\{\partial_2 C^i(u_i, T^{-1}(t))\right\}\,\mathsf{d}t \\ &= \int_{[0,1]} \min\left\{\partial_2 C^i(u_i, T^{-1}(t))\right\}\,\mathsf{d}\lambda^T(t) \\ &= \int_{[0,1]} \min\left\{\partial_2 C^i(u_i, s)\right\}\,\mathsf{d}\lambda(s) \\ &= \bigvee_{i=1}^{d} C^i\ (u_1,\dots,u_d)\end{aligned}$$

for all $(u_1,\dots,u_d) \in [0,1]^d$. ■

2.2 Approximation of upper products of copulas

The ordering properties developed in this paper depend strongly on the approximation of the upper products by upper products of discrete grid copulas. In the second part of this section we derive this kind of approximations. In the first part of this section we give some continuity results.

The upper product of copulas depends on the partial derivatives of its arguments. So, approximating the upper product also means approximating the partial derivatives. As we show in the following example uniform convergence of $(D_n^i)_n \subset \mathcal{C}_2$ is not sufficient for uniform convergence of $(\bigvee_i D_n^i)_n$.

Example 2.7 *Let $(T_n)_{n\in\mathbb{N}} \subset \mathcal{T}_P$ be a shuffle-of-min approximation of Π^2, i.e. $S_{T_n}(M^2) \to \Pi^2$ pointwise (and thus from Arzelà–Ascoli's Theorem also uniform), see Mikusinski et al. (1992, Theorem 3.1). Since $S_T(\Pi^2) = \Pi^2$ for all $T \in \mathcal{T}_P$, it follows that*

$$\begin{aligned}\Big(\lim_{n\to\infty} S_{T_n}(M^2)\Big) \vee \Big(\lim_{n\to\infty} S_{T_n}(\Pi^2)\Big) &= \Pi^2 \vee \Pi^2 = M^2 \\ &\neq \Pi^2 = M^2 \vee \Pi^2 \\ &= \lim_{n\to\infty} \left(S_{T_n}(M^2) \vee S_{T_n}(\Pi^2)\right),\end{aligned}$$

where the last equality follows from Proposition 2.6. Thus uniform convergence of $(D_n^i)_n$ does not imply in general (uniform) convergence of the upper products.

To establish continuity properties of upper products we consider the following metrics on $\mathcal{C}_2$ (see Trutschnig (2011, Lemma 4)).

$$D_1(A,B) := \int_0^1 \int_0^1 |\partial_2 A(u,t) - \partial_2 B(u,t)| \,\mathsf{d}t \,\mathsf{d}u\,, \tag{12}$$

$$D_2(A,B) := \left(\int_0^1 \int_0^1 |\partial_2 A(u,t) - \partial_2 B(u,t)|^2 \,\mathsf{d}t \,\mathsf{d}u \right)^{\frac{1}{2}}\,,$$

$$D_\infty(A,B) := \sup_{u\in[0,1]} \int_0^1 |\partial_2 A(u,t) - \partial_2 B(u,t)| \,\mathsf{d}t\,.$$

Let d_{sup} be the supremum metric on $\mathcal{C}_d$. Then, the following continuity result holds true.

Proposition 2.8 *Let D be one of the metrics D_1, D_2, and D_∞. Then, the upper product $\bigvee\colon (\mathcal{C}_2, D)^d \to (\mathcal{C}_d, d_{\text{sup}})$ is continuous in each place and also jointly continuous.*

Proof: Since the metrics D_1, D_2, and D_∞ are equivalent (see Trutschnig (2011, Theorem)) assume WLOG that $D = D_\infty$. Let $E_n^i, E^i \in \mathcal{C}_2$ be bivariate copulas for $n \in \mathbb{N}$ and $1 \leq i \leq d$ such that $D_\infty(E_n^i, E^i) \to 0$ for all i. Define $f_n^i(t) := \partial_2 E_n^i(u_i, t)$ and $f^i(t) := \partial_2 E^i(u_i,t)$. Then, $f_n^i \to f^i$ in L^1. Using the representation $\min(x,y) = \frac{1}{2}(x+y-|x-y|)$ it holds for $d=2$ that

$$\begin{aligned}
&2\int |\min\{f_n^1, f_n^2\} - \min\{f^1, f^2\}| \,\mathsf{d}t \\
&\quad = \int |(f_n^1 + f_n^2) - (f^1 + f^2) - (|f_n^1 - f_n^2| - |f^1 - f^2|)| \,\mathsf{d}t \\
&\quad \leq \int |f_n^1 - f^1| + |f_n^2 - f^2| + \big| |f_n^1 - f_n^2| - |f^1 - f^2| \big| \,\mathsf{d}t \\
&\quad \leq 2 \left(\int |f_n^1 - f^1| \,\mathsf{d}t + \int |f_n^2 - f^2| \,\mathsf{d}t \right) \to 0\,,
\end{aligned}$$

and thus $E_n^1 \vee E_n^2(u_1,u_2) \to E^1 \vee E^2(u_1,u_2)$. If $d > 2$, assume that $g_n^1 := \min_{i=1,\ldots,d-1}\{f_n^i\} \to \min_{i=1,\ldots,d-1}\{f^i\} =: g^1$ in L^1. With $g_n^2(t) := \partial_2 E_n^d(u_d,t)$ and $g^2(t) := \partial_2 E^d(u_d,t)$ it holds as above that

$$\int |\min\{g_n^1, g_n^2\} - \min\{g^1, g^2\}| \,\mathsf{d}t \to 0\,,$$

hence $\bigvee_{i=1}^d E_n^i(u_1,\ldots,u_d) \to \bigvee_{i=1}^d E^i(u_1,\ldots,u_d)$.

The assertion follows from Arzelà–Ascoli's Theorem with the equicontinuity of the set of copulas. ■

For $n \in \mathbb{N}$ and $d \geq 1$ denote by

$$\begin{aligned}
\mathbb{G}_n^d &:= \left\{ \left(\tfrac{i_1}{n}, \ldots, \tfrac{i_d}{n}\right) | i_k \in \{1,\ldots,n\}, k \in \{1,\ldots,d\} \right\} \quad \text{resp.} \\
\mathbb{G}_{n,0}^d &:= \left\{ \left(\tfrac{i_1}{n}, \ldots, \tfrac{i_d}{n}\right) | i_k \in \{0,\ldots,n\}, k \in \{1,\ldots,d\} \right\}
\end{aligned}$$

the (extended) uniform unit *n-grid* of dimension d with edge length $\frac{1}{n}$.

Let $C \in \mathcal{C}_d$ be a d-copula with associated probability measure μ_C. Let β_n be the probability measure on $[0,1]^d$ which distributes to each cell $[u - \frac{1}{n}, u]$, $u \in \mathbb{G}_n^d$, the mass $\mu_C([u - \frac{1}{n}, u])$ uniformly to the cell. Let C_n be the cumulative distribution function associated with β_n, i.e.

$$C_n(u_1, \ldots, u_d) = \beta_n([0, u_1] \times \cdots \times [0, u_d]), \quad u \in [0,1]^d.$$

Then, it holds that C_n is a copula for all n, $C_n(u) = C(u)$ for all $u \in \mathbb{G}_{n,0}^d$ and $C_n \to C$ uniformly. The sequence $(C_n)_n$ is called the *checkerboard approximation* of C and C_n is the *n-checkerboard copula* of C.

Corollary 2.9 *For $1 \leq i \leq d$, let $(D_n^i)_n$ be the checkerboard approximation of $D^i \in \mathcal{C}_2$. Then, it holds $\bigvee_{i=1}^d D_n^i \to \bigvee_{i=1}^d D^i$ uniformly.*

Proof: Defining ∂-convergence as in Mikusiński and Taylor (2010, Definition 3) it is shown in Trutschnig (2011, p. 695) that the topology of ∂-convergence is strictly finer than the topology of D_1. Then, the statement follows from Proposition 2.8 with the ∂-convergence of the checkerboard approximations as shown in Mikusiński and Taylor (2010, Theorem 5). ■

Similar results hold also true for checkmin approximations and Bernstein approximations of copulas (see Mikusiński and Taylor (2010, Theorem 6 and Theorem 7))

In the following, we make essential use of discrete approximations of the upper product by so-called grid copulas.

Definition 2.10 *For $d \in \mathbb{N}$, a (signed)* n-grid d-copula *(shortly* grid copula*) D is the (signed) distribution function of a (signed) probability distribution on $\mathbb{G}_{n,0}^d$ with* uniform univariate margins, *i.e. for all $i = 1, \ldots, d$ holds $D(u) = \frac{k}{n}$, for all $k = 0, \ldots, n$, if $u_i = \frac{k}{n}$ and $u_j = 1$ for all $j \neq i$.*

Denote by $\mathcal{C}_{d,n}$ ($\mathcal{C}_{d,n}^s$) the set of all (signed) d-dimensional n-grid copulas.

An $\frac{1}{n}$*-scaled doubly stochastic matrix* is defined as an $n \times n$-matrix with non-negative entries and row resp. column sums equal to $\frac{1}{n}$. By an *signed* $\frac{1}{n}$*-scaled doubly stochastic matrix* we mean an $\frac{1}{n}$-scaled doubly stochastic matrix where also negative entries are allowed.

The following statement is immediate.

Lemma 2.11 *There is a one-to-one correspondence between the set of (signed) n-grid 2-copulas and the set of (signed) $\frac{1}{n}$-scaled doubly stochastic matrices.*

Note that also bivariate n-checkerboard copulas can be represented by $\frac{1}{n}$-scaled doubly stochastic matrices.

For a bivariate (signed) n-grid copula $E \in \mathcal{C}_{2,n}$ ($\in \mathcal{C}_{2,n}^s$) let e, defined through

$$e(u,v) := \Delta_n^1 \Delta_n^2 E(u,v), \quad (u,v) \in \mathbb{G}_n^2,$$

be its corresponding (signed) probability mass function, where Δ_n^i denotes the difference operator of length $\frac{1}{n}$ with respect to the i-th variable, i.e. $\Delta_n^i g(u) := g(u) - g((u - \frac{1}{n}e_i) \vee 0)$ for $u \in \mathbb{G}_{n,0}^d$ and e_i being the unit vector with value 1 in the i-th component. Further, define its corresponding (signed) $\frac{1}{n}$-scaled doubly stochastic matrix $(e_{kl})_{1\leq k,l\leq n}$ by

$$e_{kl} = e(1 - \tfrac{k-1}{n}, \tfrac{l}{n})\,. \tag{13}$$

For every copula $D \in \mathcal{C}_d$ denote by $\mathbb{G}_n(D)$ its *canonical n-grid copula* defined through

$$\mathbb{G}_n(D)(u) := D(\tfrac{\lceil nu \rceil}{n})$$

for $u \in [0,1]^d$, where $\lceil \cdot \rceil$ denotes the componentwise ceiling function. Further, every $D_n \in \mathcal{C}_{d,n}$ ($\in \mathcal{C}_{d,n}^s$) can be extended to a (signed) distribution function D on $[0,1]^d$ via

$$D(u) := D_n(\tfrac{\lceil nu \rceil}{n}) \tag{14}$$

for $u \in [0,1]^d$.

Define the upper product $\bigvee\colon (\mathcal{C}_{2,n})^d \to \mathcal{C}_{d,n}$ for grid copulas $D_n^1, \ldots, D_n^d \in \mathcal{C}_{2,n}$ through

$$\bigvee_{i=1}^d D_n^i(u_1, \ldots, u_d) := \sum_{k=1}^n \min_{1\leq i\leq d} \left\{ \Delta_n^2 D^i(u_i, \tfrac{k}{n}) \right\}\,.$$

A version for signed grid copulas is defined analogously.

We show that the upper product of bivariate copulas can be uniformly approximated by the upper product of the corresponding grid copula approximations in the extended version given by (14).

Proposition 2.12 (Grid copula approximation of the upper product) *Let $D^1, \ldots, D^d \in \mathcal{C}_2$ be copulas. Then*

$$\bigvee_{i=1}^d \mathbb{G}_n(D^i) \quad \xrightarrow{\mathcal{D}} \quad \bigvee_{i=1}^d D^i \quad \textit{for } n \to \infty\,.$$

Proof: We need to show that $\bigvee_{i=1}^d \mathbb{G}_n(D^i)(\frac{\lceil nu \rceil}{n}) \to \bigvee_{i=1}^d D^i(u)$ for all $u \in [0,1]^d$ and $n \to \infty$. Define

$$D_n(u_1, \ldots, u_d) := \sum_{k=1}^n \min_{1\leq i\leq d} \left\{ \Delta_n^2 D^i(u_i, \tfrac{k}{n}) \right\}\,.$$

It can be shown that D_n is a copula for all n. We need to show that

$$D_n(u) \xrightarrow{n\to\infty} \bigvee_{i=1}^d D^i(u) \tag{15}$$

for all $u = (u_1, \dots, u_d) \in [0,1]^d$. Then, the statement follows from

$$\bigvee_{i=1}^d \mathbb{G}_n(D^i)(\tfrac{\lceil nu \rceil}{n}) = D_n(\tfrac{\lceil nu \rceil}{n}) \to \bigvee_{i=1}^d D^i(u)$$

for all $u \in [0,1]^d$ with the equicontinuity of $(D_n)_{n\in\mathbb{N}}$. The proof of the convergence in (15) is given in the appendix. ■

3 Ordering risk bounds for $\sum X_i$ in partially specified factor models

To solve maximization problem (6) for suitable sets $\mathcal{S}^i$ we aim to order solutions of the maximization problem (5) w.r.t. $\leq_{cx}$ for all marginal distributions F_i and in dependence on the constraints C^i. We first demonstrate that the usual ordering conditions (like supermodular ordering) for the constraints $C^i \in \mathcal{C}_2$ do not imply ordering of the upper product $\bigvee_i C^i$. We are, therefore, led to introduce a new type of orderings defined by the sign changes of the copula derivatives. The main result in this paper, Theorem 3.10, states that these new ordering conditions imply the desired ordering properties of the upper products.

It turns out that the supermodular ordering $\leq_{sm}$ of random vectors is sufficient for convex ordering of the sums independent of the marginal distributions whereas the weaker concordance ordering $\leq_c$ may lack this property. For an overview on stochastic orderings, see Müller and Stoyan (2002, Example 3.9.7) and Shaked and Shanthikumar (2007). Hence, the aim is to find conditions on the constraints $C^i, D^i \in \mathcal{C}_2$, $1 \leq i \leq d$, such that

$$\bigvee_{i=1}^d C^i \leq_{sm} \bigvee_{i=1}^d D^i \tag{16}$$

because this implies

$$\sum_{i=1}^d F^{-1}_{X_i|Z}(U) \leq_{cx} \sum_{i=1}^d F^{-1}_{Y_i|Z}(U)$$

for $C_{X_i,Z} = C^i$, $C_{Y_i,Z} = D^i$ and for all $X_i \sim Y_i$ and F_Z continuous.

A necessary condition for (16) is the lower orthant ordering, i.e.

$$\int_0^1 \min_i \{\partial_2 C^i(u_i,t)\} \, \mathsf{d}t \leq \int_0^1 \min_i \{\partial_2 D^i(u_i,t)\} \, \mathsf{d}t \quad \forall u \in [0,1]^d . \tag{17}$$

Ordering the constraints with respect to the supermodular ordering is not sufficient to obtain (16) as the following example illustrates.

Example 3.1 *(a) The upper product is not componentwise increasing w.r.t. the supermodular ordering, i.e. $A <_{sm} B$ for $A, B \in \mathcal{C}_2$ does not imply $C \vee A <_{sm} C \vee B$ for all $C \in \mathcal{C}_2$, because $C = A$ yields $C \vee A = M^2 >_{sm} C \vee B$ using Proposition 2.4*(v)*.*

(b) Consider the following bivariate 4-checkerboard copulas $A^1, A^2, A^3 \in \mathcal{C}_2$ given through the $\frac{1}{n}$-scaled doubly stochastic matrices

$$a^1 = \frac{1}{16} \cdot \begin{pmatrix} 0 & 2 & 0 & 2 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 2 & 0 & 2 & 0 \end{pmatrix}, \quad a^2 = \frac{1}{16} \cdot \begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 0 \end{pmatrix},$$

$$a^3 = \frac{1}{16} \cdot \begin{pmatrix} 0 & 0 & 2 & 2 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 2 & 2 & 0 & 0 \end{pmatrix}$$

as in (13)*. Then, it holds*

$$\Pi^2 <_{sm} A^1 <_{sm} A^2 <_{sm} A^3 \,. \tag{18}$$

Define the functions $h^i_{u_1,u_2} \colon (0,1) \to [-1,1]$ for $i = 1,2,3$ by

$$\begin{aligned} h^1_{u_1,u_2}(t) :&= \partial_2 A^1(u_2, t) - \partial_2 \Pi^2(u_1, t) \\ &= u_2 + \mathbb{1}_{\{t<\frac{1}{4}\}\cup\{\frac{1}{2}<t<\frac{3}{4}\}}[u_2 \vee \tfrac{1}{4} - (u_2 - \tfrac{3}{4}) \vee 0] \\ &\quad + \mathbb{1}_{\{\frac{1}{4}<t<\frac{1}{2}\}\cup\{\frac{3}{4}<t\}}[(u_2 - \tfrac{3}{4}) \vee 0 - u_2 \vee \tfrac{1}{4}] - u_1 \,, \\ h^2_{u_1,u_2}(t) :&= \partial_2 A^2(u_2, t) - \partial_2 \Pi^2(u_1, t) \\ &= u_2 + \mathbb{1}_{\{t<\frac{1}{4}\}}[u_2 \vee \tfrac{1}{4} - (u_2 - \tfrac{3}{4}) \vee 0] \\ &\quad + \mathbb{1}_{\{t>\frac{3}{4}\}}[(u_2 - \tfrac{3}{4}) \vee 0 - u_2 \vee \tfrac{1}{4}] - u_1 \,, \\ h^3_{u_1,u_2}(t) :&= \partial_2 A^3(u_2, t) - \partial_2 \Pi^2(u_1, t) \\ &= u_2 + \mathbb{1}_{\{t<\frac{1}{2}\}}[u_2 \vee \tfrac{1}{4} - (u_2 - \tfrac{3}{4}) \vee 0] \\ &\quad + \mathbb{1}_{\{t>\frac{1}{2}\}}[(u_2 - \tfrac{3}{4}) \vee 0 - u_2 \vee \tfrac{1}{4}] - u_1 \,. \end{aligned}$$

We observe that $h^2_{u_1,u_2} \prec_S h^1_{u_1,u_2} =_S h^3_{u_1,u_2}$ for all $u_1, u_2 \in [0,1]$, where $\prec_S$ denotes the Schur-ordering for functions. This implies with the Hardy-

Littlewood-Polya-Theorem (see Rüschendorf (2013, Theorem 3.21)) that

$$\begin{aligned}\Pi^2 \vee A^j(u_1,u_2) &= \int \min\{h^j_{u_1,u_2}(t),0\}\,\mathsf{d}t + u_1\\ &\geq \int \min\{h^2_{u_1,u_2}(t),0\}\,\mathsf{d}t + u_1\\ &= \Pi^2 \vee A^2(u_1,u_2)\,,\end{aligned}$$

for $j = 1,3$ and for all $(u_1,u_2) \in [0,1]^2$. Further, the inequality is strict, e.g. for $u_1 = u_2 = \frac{1}{4}$. Hence, we obtain

$$\Pi^2 \vee A^1 = \Pi^2 \vee A^3 >_{sm} \Pi^2 \vee A^2\,.$$

In consequence, the supermodular ordering of the constraints in (18) *does not yield ordering of the risk bounds in the natural way as described in Remark 2.5*(c)*.*

Note that also a pointwise ordering of the integrands in (17) is not possible. This demands to obtain ordering criteria for the whole integral. The identity

$$\int_0^1 \min\{h_1(t),h_2(t)\}\,\mathsf{d}t = \int_0^1 \min\{h_2(t)-h_1(t),0\}\,\mathsf{d}t + \int_0^1 h_2(t)\,\mathsf{d}t$$

motivates the following lemma.

Lemma 3.2 *Let $f,g\colon [0,1] \to \mathbb{R}$ be integrable functions with the properties that*

(i) $\int_0^1 f\,\mathsf{d}\lambda = \int_0^1 g\,\mathsf{d}\lambda$,

(ii) f,g have no $(-,+)$-sign change,

(iii) $g-f$ has no $(-,+)$-sign change.

Then it holds that

$$\int_0^1 f_-\,\mathsf{d}\lambda \geq \int_0^1 g_-\,\mathsf{d}\lambda \quad \textit{and} \quad \int_0^1 f_+\,\mathsf{d}\lambda \leq \int_0^1 g_+\,\mathsf{d}\lambda\,, \tag{19}$$

where h_- resp. h_+ denotes the negative resp. positive part of a function h. Further, every change of the sign sequence in (ii) *or in* (iii) *produces a change of the inequality signs in* (19)*.*

Proof: Conditions (i) and (iii) provide that there exists a point $s \in (0,1)$ such that $f \leq g$ on $(0,s)$ and $f \geq g$ on $(s,1)$. This implies $f_+ \leq g_+$ on $(0,s)$ and $f_+ \geq g_+$ on $(s,1)$.

If $g(s) < 0$, we obtain from condition (ii) that $f_+ = g_+ = 0$ on $(0,s)$, hence $\int f_+\,\mathsf{d}\lambda \geq \int g_+\,\mathsf{d}\lambda$.

If $g(s) \geq 0$, then condition (ii) provides $g_+ = g$ and thus $f_+ = f$ on $(s,1)$. Hence, it follows that

$$\int_s^1 (f_+ - g_+)\,\mathsf{d}\lambda = \int_s^1 (f-g)\,\mathsf{d}\lambda = \int_0^s (g-f)\,\mathsf{d}\lambda \geq \int_0^s (g_+ - f_+)\,\mathsf{d}\lambda$$

using Condition (i), and because $f \leq g$ on $(0, s)$ the inequality holds true due to

$$\{(x, y)|0 \leq x \leq s, f(x) \leq y \leq g(x)\} \supset \{(x, y)|0 \leq x \leq s, f(x) \leq y \leq g(x), y \geq 0\}.$$

If the sign sequence in condition (iii) is $(+, -)$, then the statement follows from the above one by changing the roles of f and g. The other cases follow by symmetry. ■

On the basis of the previous lemma, we introduce a new ordering on $\mathcal{C}_2$ and show in the sequel that this ordering provides supermodular ordering criteria for the upper product of bivariate copulas.

Definition 3.3 (Sign sequence ordering of derivative differences) *Let $D, E \in \mathcal{C}_2$ be bivariate copulas. Consider for $u, v \in [0, 1]$ the function $f_{u,v}(t) := \partial_2 E(v, t) - \partial_2 D(u, t)$ for almost all $t \in (0, 1)$,*

1. *Define that E is greater than D in the* sign sequence relation of derivative differences*, written $D \leq_{\partial\Delta} E$, if for all $u, v \in (0, 1)$ holds that*

$$f_{u,v} \textit{ has } \lambda\textit{-almost surely no } (-,+)\textit{-sign change}. \tag{20}$$

2. *A family $(C^\alpha)_{\alpha \in I} \subset \mathcal{C}_2$, $I \subset \mathbb{R}$, of bivariate copulas is* increasing with respect to the $\leq_{\partial\Delta}$-ordering *if $\alpha_1 < \alpha_2$, $\alpha_1, \alpha_2 \in I$, implies $C^{\alpha_1} \leq_{\partial\Delta} C^{\alpha_2}$.*
3. *For copulas $B^1, \ldots, B^d \in \mathcal{C}_2$ the d-order relation $B^1 \leq_{\partial\Delta} B^2 \leq_{\partial\Delta} \cdots \leq_{\partial\Delta} B^d$ is defined by $B^i \leq_{\partial\Delta} B^j$ for all $1 \leq i < j \leq d$.*
4. *Analogously, define the* symmetric sign sequence relation of derivative differences *$D \leq_{s\partial\Delta} E$ if (20) holds for all $u = v$.*

For bivariate grid copulas, the relations $\leq_{\partial\Delta}$ and $\leq_{s\partial\Delta}$ are defined in the same way.

The $\leq_{\partial\Delta}$-relation is a relation that is strictly stronger than the $\leq_{sm}$-relation. It can easily be verified that the reverse directions in the following result do not hold.

Proposition 3.4 *For $D, E \in \mathcal{C}_2$ holds that*

(i) *$D \leq_{\partial\Delta} E$ implies $D \leq_{s\partial\Delta} E$,*

(ii) *$D \leq_{s\partial\Delta} E$ implies $D \leq_{sm} E$.*

Proof: Statement (i) is trivial. Statement (ii) follows from

$$E(u, v) - D(u, v) = \int_0^v (\partial_2 E(u, t) - \partial_2 D(u, t)) \, \mathsf{d}t \geq 0$$

because the integrand has no $(-, +)$-sign change in t and the integral vanishes for $v = 1$. ■

Example 3.5 *(a) Elliptical copulas: Let* $(X_i, Z) \stackrel{\mathrm{d}}{=} RU^{(2)}A_i \sim \mathcal{EC}_2(0, \Sigma_i, \phi)$ *,* $i = 1, 2$ *, be elliptically distributed with* $A_i^T A_i = \Sigma_i = \left(\begin{smallmatrix} 1 & \rho_i \\ \rho_i & 1 \end{smallmatrix}\right)$ *. Assume that the radial part* R *has a continuous distribution function. Then the copula* $C_{X_i,Z}$ *of* (X_i, Z) *is uniquely determined. Assume that* $-1 < \rho_1 < \rho_2 < 1$ *. Then, from Cambanis et al. (1981, Corollary 5) we obtain*

$$F_{X_i|Z=z}(x_i) = F_{R_z^\pm}\left(\frac{x_i - \rho_i z}{\sqrt{1-\rho_i^2}}\right),$$

where $R_z^\pm \stackrel{\mathrm{d}}{=} R_z U^{(1)}$ *, with* $R_z \stackrel{\mathrm{d}}{=} \left(\sqrt{R^2 - z^2}|Z = z\right)$ *, does not depend on* ρ_i *,* $U^{(1)} \sim U(\{-1, 1\})$ *, and* $R_z, U^{(1)}$ *are independent for all* z *. This implies for all* x_1, x_2 *that*

$$\begin{aligned}
& F_{X_2|Z=z}(x_2) - F_{X_1|Z=z}(x_1) \geq 0 \\
\iff \quad & \frac{x_2 - \rho_2 z}{\sqrt{1-\rho_2^2}} \geq \frac{x_1 - \rho_1 z}{\sqrt{1-\rho_1^2}} \\
\iff \quad & z(\rho_1\sqrt{1-\rho_2^2} - \rho_2\sqrt{1-\rho_1^2}) \geq \sqrt{1-\rho_2^2}x_1 - \sqrt{1-\rho_1^2}x_2 \\
\iff \quad & z \leq \frac{\sqrt{1-\rho_2^2}x_1 - \sqrt{1-\rho_1^2}x_2}{\rho_1\sqrt{1-\rho_2^2} - \rho_2\sqrt{1-\rho_1^2}},
\end{aligned}$$

where the last equivalence holds because $\rho_1 < \rho_2$ *. Hence, we obtain*

$$\begin{aligned}
& f_{u_1,u_2}(t) := \partial_2 C_{X_2,Z}(u_2, t) - \partial_2 C_{X_1,Z}(u_1, t) \geq 0 \\
\iff \quad & t \leq F_{R^\pm}\left(\frac{\sqrt{1-\rho_2^2}F_{R^\pm}^{-1}(u_1) - \sqrt{1-\rho_1^2}F_{R^\pm}^{-1}(u_2)}{\rho_1\sqrt{1-\rho_2^2} - \rho_2\sqrt{1-\rho_1^2}}\right),
\end{aligned}$$

where $F_{R^\pm} = F_{X_i} = F_Z$ *is the distribution function of* $R^\pm := RU^{(1)}$ *. But this means that* $C_{X_1,Z} \leq_{\partial\Delta} C_{X_2,Z}$ *.*

(b) Archimedean copulas: As shown in Nelsen (2006, Section 4.4) the Clayton family (4.2.1), the Gumbel-Hougaard family (4.2.4), the Frank family (4.2.5) and the families (4.2.2) and (4.2.19) in Nelsen (2006) are ordered in concordance. Numerical results suggest that these families are even $\leq_{\partial\Delta}$*-increasing.*

In the following, we show that the $\leq_{\partial\Delta}$-ordering of the constraints implies the $\leq_{sm}$-ordering of the upper product if we substitute the greatest or smallest element in the $\leq_{\partial\Delta}$-increasing sequence of constraints, see Theorem 3.10. For the proof, we approximate the upper product by grid-copulas and use the lower orthant ordering result given in the following proposition.

Proposition 3.6 *Let* $A^1, \ldots, A^d, B^1, B^2 \in \mathcal{C}_2$ *be bivariate copulas such that*

$$A^j \leq_{\partial\Delta} B^1, B^2 \quad \text{and} \quad B^1 \leq_{s\partial\Delta} B^2, \quad 1 \leq j \leq d.$$

Then, it holds that

$$A^1 \vee \dots \vee A^d \vee B^1 \geq_c A^1 \vee \dots \vee A^d \vee B^2 \,. \tag{21}$$

Proof: Each function $f^{ij}_{u_i,v}(t) := \partial_2 B^j(v,t) - \partial_2 A^i(u_i,t)\,,\ i = 1,\dots,n\,,\ j = 1,2\,,$ has no $(-,+)$-sign change. Hence, also the pointwise defined functions $g^j := \min_i\{f^{ij}_{u_i,v}\}\,,\ j = 1,2\,,$ have no $(-,+)$-sign change. Assumption $B^1 \leq_{s\partial_\Delta} B^2$ ensures that the function $g^1 - g^2$ has no $(+,-)$-sign change. Since $\int g^1 \,\mathsf{d}\lambda = \int g^2 \,\mathsf{d}\lambda\,,$ we obtain from Lemma 3.2 that $\int g^1_- \,\mathsf{d}\lambda \geq \int g^2_- \,\mathsf{d}\lambda\,,$ hence

$$\begin{aligned}
&A^1 \vee \ldots \vee A^d \vee B^1(u_1,\dots,u_d,v) \\
&\quad = \int \min\Big\{0, \partial_2 B^1(v,t) - \min_i\{\partial_2 A^i(u_i,t)\}\Big\} \,\mathsf{d}t + \int \min_i\{\partial_2 A^i(u_i,t)\} \,\mathsf{d}t \\
&\quad = \int g^1_-(t) \,\mathsf{d}t + \int \min_i\{\partial_2 A^i(u_i,t)\} \,\mathsf{d}t \\
&\quad \geq \int g^2_-(t) \,\mathsf{d}t + \int \min_i\{\partial_2 A^i(u_i,t)\} \,\mathsf{d}t \\
&\quad = \int \min\Big\{0, \partial_2 B^2(v,t) - \min_i\{\partial_2 A^i(u_i,t)\}\Big\} \,\mathsf{d}t + \int \min_i\{\partial_2 A^i(u_i,t)\} \,\mathsf{d}t \\
&\quad = A^1 \vee \ldots \vee A^d \vee B^2(u_1,\dots,u_d,v)\,.
\end{aligned}$$

This holds for all $(u_1,\dots,u_d,v) \in (0,1)^{d+1}$, and thus $A^1 \vee \ldots \vee A^d \vee B^1 \geq_{lo} A^1 \vee \ldots \vee A^d \vee B^2$. The upper orthant ordering follows analogously with Proposition 2.4 (viii). ∎

To show the $\leq_{sm}$-ordering of the upper product it suffices to order the grid copula approximations w.r.t. $\leq_{sm}$ as the following result states.

Proposition 3.7 *Let $D^1,\dots,D^d,E^1,\dots,E^d \in \mathcal{C}_2$ be bivariate copulas. Then, it holds*

$$\bigvee_{i=1}^d \mathbb{G}_n(D^i) \leq_{sm} \bigvee_{i=1}^d \mathbb{G}_n(E^i) \quad \forall n \in \mathbb{N} \quad \Longrightarrow \quad \bigvee_{i=1}^d D^i \leq_{sm} \bigvee_{i=1}^d E^i\,.$$

Proof: Proposition 2.12 yields $\bigvee_i \mathbb{G}_n(D^i) \xrightarrow{\mathcal{D}} \bigvee_i D^i$ and $\bigvee_i \mathbb{G}_n(E^i) \xrightarrow{\mathcal{D}} \bigvee E^i$. Since the supermodular ordering is closed with respect to weak convergence (see Müller and Scarsini (2000, Theorem 3.5)), the statement follows. ∎

The grid-copula approximations define distributions with finite support. But the supermodular ordering of distributions with finite support has been characterized by supermodular transfers in Müller (2013, Theorem 2.5.4). It is clear that this result also holds for finite signed distributions with finite support:

Proposition 3.8 *Let μ and ν be finite signed distributions on $\mathbb{G}^d_n$. Then, $\mu \leq_{sm} \nu$ if and only if there exist a finite number $m \in \mathbb{N}_0$, weights $q_i > 0$ and points*

$x^i, y^i \in \mathbb{G}_n^d$, $1 \le i \le m$, *such that*

$$\mu + \sum_{i=1}^m \eta_i = \nu\,, \quad \text{where } \eta_i = q_i\left(\tfrac{1}{2}\delta_{x^i \wedge y^i} + \tfrac{1}{2}\delta_{x^i \vee y^i} - \left(\tfrac{1}{2}\delta_{x^i} + \tfrac{1}{2}\delta_{y^i}\right)\right)$$

The signed measures η_i are called supermodular transfers *and are* indicated *by*

$$q_i\left(\tfrac{1}{2}\delta_{x^i} + \tfrac{1}{2}\delta_{y^i}\right) \;\to\; q_i\left(\tfrac{1}{2}\delta_{x^i \wedge y^i} + \tfrac{1}{2}\delta_{x^i \vee y^i}\right)\,,$$

i.e. mass of size q_i is transferred from x^i and y^i to $x^i \wedge y^i$ and $x^i \vee y^i$.

The following result states that (21) even holds w.r.t. the $\le_{sm}$-ordering in the case of grid copulas. The technical proof is given in the appendix.

Proposition 3.9 *Let $A^1, \ldots, A^d, B^1, B^2 \in \mathcal{C}_{2,n}$ be bivariate grid copulas such that*

$$A^j \le_{\partial\Delta} B^l \quad \text{and} \quad B^1 \le_{s\partial\Delta} B^2\,, \quad 1 \le j \le d\,, l = 1, 2\,.$$

Then, there exists a finite sequence $(E^i)_{0 \le i \le m}$ of signed probability distribution functions on $\mathbb{G}_n^2$ such that $E^0 = B^1$, $E^m = B^2$ and

(i) for all $0 \le i \le m-1$, $P_{E^{i+1}} - P_{E^i}$ is a simple supermodular transfer,

(ii) $E^i \le_{s\partial\Delta} E^{i+1}$ for all $0 \le i \le m-1$,

(iii) $A^j \le_{\partial\Delta} E^i$ f.a. $1 \le j \le d$ and $0 \le i \le m$, and

(iv) $A^1 \vee \cdots \vee A^d \vee E^i \ge_{sm} A^1 \vee \cdots \vee A^d \vee E^{i+1}$ for all $0 \le i \le m-1$.

It follows that

(v) $A^1 \vee \cdots \vee A^d \vee B^1 \ge_{sm} A^1 \vee \cdots \vee A^d \vee B^2$.

Now, we can formulate the main result of this article which provides some important properties of the $\le_{\partial\Delta}$-ordering. In contrast to the $\le_{sm}$-ordering on $\mathcal{C}_2$ (see Example 3.1(b)), the $\le_{\partial\Delta}$-ordering on $\mathcal{C}_2$ is sufficient for the supermodular ordering of the upper product.

Theorem 3.10 *Let $A^1, \ldots, A^d, B^1, B^2 \in \mathcal{C}_2$ be bivariate copulas such that either*

(i) $A^j \le_{\partial\Delta} B^i$ and $B^1 \le_{s\partial\Delta} B^2$, $1 \le j \le d$, $i = 1, 2$, or

(ii) $A^j \ge_{\partial\Delta} B^i$ and $B^1 \ge_{s\partial\Delta} B^2$, $1 \le j \le d$, $i = 1, 2$.

Then, it holds that

$$A^1 \vee \cdots \vee A^d \vee B^1 \ge_{sm} A^1 \vee \cdots \vee A^d \vee B^2\,. \tag{22}$$

Proof: Assume that (i) holds. Then, we obtain $\mathbb{G}_n(A^j) \leq_{\partial\Delta} \mathbb{G}_n(B^i)$ and $\mathbb{G}_n(B^1) \leq_{s\partial\Delta} \mathbb{G}_n(B^2)$ for all $1 \leq j \leq d$, $i = 1,2$ and $n \in \mathbb{N}$. Thus, the statement follows from Proposition 3.9 (v) and Proposition 3.7.
If (ii) holds, then the statement follows from (i) and Proposition 2.6 with $T(t) = 1 - t$. ■

It can be shown analogously that (22) can be generalized to

$$A^1 \vee \cdots \vee A^d \vee \underbrace{B^1 \vee \cdots \vee B^1}_{\delta\text{-times}} \geq_{sm} A^1 \vee \cdots \vee A^d \vee \underbrace{B^2 \vee \cdots \vee B^2}_{\delta\text{-times}} \tag{23}$$

for every $\delta \in \mathbb{N}_0$. Applying (23) repeatedly, we obtain together with Proposition 3.4 the following corollary.

Corollary 3.11 *Let $C^1, \ldots, C^d \in \mathcal{C}_2$ be bivariate copulas such that $C^1 \leq_{\partial\Delta} \cdots \leq_{\partial\Delta} C^d$. Then, for $1 \leq d_1 \leq d_2 \leq d$ holds*

$$\bigvee_{i=1}^{d} C^i \leq_{sm} \bigvee_{i=1}^{d} D^i$$

where $D^i := C^{d_1}$ for $1 \leq i \leq d_1$, $D^i := C^i$ for $d_1 < i < d_2$ and $D^i := C^{d_2}$ for $d_2 \leq i \leq d$.

Remark 3.12 *(a) Theorem 3.10 and Corollary 3.11 indicate: The closer the elements are together w.r.t the $\leq_{\partial\Delta}$-ordering the greater is their upper product w.r.t. the supermodular ordering. Note that we only modify the most extreme elements keeping the others fixed.*

(b) Corollary 3.11 is a generalization of Ansari and Rüschendorf (2016, Corollary 3 and Proposition 6) to general classes of copulas and to the supermodular ordering.

Coming back to the comparison of solutions of (5) w.r.t. the constraints C^i we get the following result.

Corollary 3.13 *Let $W_1, \ldots, W_d, Z$ be real random variables such that the sequence of copulas $(C_{W_i,Z})_{1\leq i \leq d}$ is $\leq_{\partial\Delta}$-increasing. Assume that Z has a continuous distribution function. Let $X_i := g_i(W_i)$ for g_i increasing. Then, for $1 \leq d_1 \leq d_2 \leq d$ holds*

$$\sum_{i=1}^{d} X_i \ \leq_{cx} \ Y_1 \ + \sum_{i=d_1+1}^{d_2-1} g_i(\xi_i) \ + \ Y_3 \,, \tag{24}$$

where $Y_1 = \sum_{i=1}^{d_1} g_i(F^{-1}_{W_{d_1}|Z}(U))$, $\xi_i = F^{-1}_{W_i|Z}(U)$ for $i = d_1+1, \ldots, d_2 - 1$, and $Y_3 = \sum_{i=d_2}^{d} g_i(F^{-1}_{W_{d_2}|Z}(U))$, $U \sim U(0,1)$ independent of Z.

If $d_1 = d_2$, then (24) simplifies to

$$\sum_{i=1}^{d} X_i \ \leq_{cx} \ Y_1 \ + \ Y_3 \, .$$

Proof: This follows from Remark 2.2, Proposition 2.4 (i) and Corollary 3.11. ∎

Remark 3.14 *(a) In Corollary 3.13, both Y_1 and Y_3 are comonotonic sums, but $Y_1 + Y_3$ is only conditionally comonotonic.*

(b) Let $(C^\gamma)_{\gamma \in \mathbb{R}} \subset \mathcal{C}_2$ be a $\leq_{\partial\Delta}$-increasing family of bivariate copulas. Denote by $\mathcal{F}^1$ the set of all univariate distribution functions and by $\mathcal{F}_\uparrow$ the set of all increasing functions. As a consequence of Corollary 3.13 we obtain solutions of maximization problem (6) resp. (7) for $\mathcal{F}_i = \mathcal{F}^1$ resp. $\mathcal{G}_i = \mathcal{F}_\uparrow$ for all continuous G if we choose the sets $\mathcal{S}^i = \{C^\gamma | \gamma \leq a\}$ for $i \leq d_1$, $\mathcal{S}^i = \{C^\gamma | \gamma = b_i\}$ for $d_1 < i < d_2$ and $\mathcal{S}^i = \{C^\gamma | \gamma \geq c\}$ for $i \geq d_2$ where $1 \leq d_1 \leq d_2 \leq d$ and $a \leq b_1 \leq \ldots \leq b_{d_2 - d_1 - 1} \leq c$.

4 Application

As application we consider a portfolio $\Sigma_t := \sum_{i=1}^{6} Y_t^i$ of calls and puts on different assets. More specifically, let $Y_t^i := (S_t^i - K^i)_+$ be calls for $i = 1, 2, 3$ and $Y_t^i := (K^i - S_t^i)_+$ be puts for $i = 4, 5, 6$, on assets S_t^i for different strikes $K^i > 0$, where $(S_t^i)_{t \geq 0}$ denotes the asset price process of Allianz ($i = 1$), Daimler ($i = 2$), Siemens ($i = 3$), Deutsche Bank ($i = 4$), SAP ($i = 5$) resp. Adidas ($i = 6$). For times to maturity $T = 15$ trading days resp. $T = 50$ trading days resp. $T = 100$ trading days, we aim to get improved risk bounds (w.r.t the standard comonotonic risk bound) for Σ_T applying Corollary 3.13 where daily historical data are given. Denote by $(S_t^0)_{t \geq 0}$ the risk factor process which is the DAX in our case.

We model $S_t = (S_t^0, \ldots, S_t^6)$ by an exponential process $S_t = S_0 \exp(L_t)$, t in trading days, under the following assumptions.

Let $0 = t_0 < t_1 < t_2 < \ldots$ with $t_i - t_{i-1} = T$ for all i.

(I) The component processes $(L_t^i)_{t \geq 0}$ are Lévy processes for all i.

(II) The time T-increments $(\xi_k^0, \xi_k^i) := (L_{t_k}^0 - L_{t_{k-1}}^0, L_{t_k}^i - L_{t_{k-1}}^i)$, $1 \leq k \leq n$ are identically distributed in k and independent in k for all $1 \leq i \leq 6$.

(III) There exists a $\leq_{\partial\Delta}$-increasing family $(C^\alpha)_{\alpha \in I}$ of bivariate copulas such that for all $1 \leq i \leq 6$, $C_{\xi_1^i, \xi_1^0} \in \{C^\alpha | \alpha \in I^i\}$ for some intervals $I^i \subset I$ (which are specified later).

Assumptions (I) – (III) are consistent. Assumption (I) is a standard assumption on the log-increments of $(S_t^i)_{t \geq 0}$ while Assumption (II) generalizes the dependence assumptions for multivariate Lévy models because neither multivariate stationarity nor independence for all increments is claimed. Assumption (III) describes the dependence structure of (ξ_1^i, ξ_1^0) by subfamily of a $\leq_{\partial\Delta}$-increasing family of copulas (see Example 3.5) which can be chosen arbitrarily.

For the estimation of the distribution of S_T^i, we distinguish between the following two specifications of Assumption (I):

1. (a) Each $(S_t^i)_{t\geq 0}$, $i = 0, \ldots, 7$, follows a geometric Brownian motion, i.e.

$$S_t^i = S_0^i \exp(L_t^i)\,, \quad L_t^i = \sigma_i B_t^i + (\mu_i - \tfrac{\sigma_i^2}{2})t\,, \quad t \geq 0$$

   where $(B_t^i)_{t\geq 0}$ is a Brownian motion, $S_0^i > 0$, $\sigma_i > 0$, $\mu_i \in \mathbb{R}$.

   (b) Each $(S_t^i)_{t\geq 0}$, $i = 0, \ldots, 7$, follows an exponential NIG process, i.e.

$$S_t^i = S_0^i \exp(L_t^i) \quad t \geq 0\,,$$

   where each $(L_t^i)_{t\geq 0}$ is an NIG process, $S_0^i > 0$.

For the estimation of upper bounds for the time T-increments $(\xi_1^1, \ldots, \xi_1^7)$ in supermodular ordering, we specify Assumption (III) as follows:

3. For fixed $\nu \in (2, \infty]$, the dependence structure of (ξ_1^i, ξ_1^0) is described by a family $(C_\nu^\rho)_{\rho\in I^i}$ of t-copulas with unknown correlation parameter $\rho \in I^i$ and ν degrees of freedom for some intervals $I^i \subset [-1, 1]$ (which we specify later), i.e. $C_{\xi_1^i, \xi_1^0} \in (C_\nu^\rho)_{\rho\in I^i}$ for all $1 \leq i \leq 6$.

For the estimation of the intervals I^i, we use the i.i.d. assumption in Assumption (II) to determine (one-sided) confidence intervals for the correlation of (ξ_1^i, ξ_1^0) from historical log-return data.

Compared to the basic assumptions underlying multivariate exponential Lévy models the above assumptions are quite weak. The dependence structure among the components is not uniquely determined. For larger values of T (which we consider in this application), the set of historical data is too small to determine the unknown correlation parameter reliably. Thus, we need to solve maximization problem (6) instead of maximization problem (5).

Such solutions lead to improved risk bounds for the portfolio Σ_T given the observed starting values $(S_0^1, \ldots, S_0^7)$ and constraints $\mathcal{S}^i$. We speak about *Model Gauss* if S_T is modeled by Assumptions (1a),(II) and (3) and about *Model NIG* if S_T is modeled by Assumptions (1b),(II) and (3).

The *normal inverse Gaussian* (NIG) distribution has density

$$d_{NIG(\alpha,\beta,\delta,\nu)}(x) = \frac{\alpha\delta}{\nu} \frac{K_1(\alpha\sqrt{\delta^2 + (x-\nu)^2})}{\sqrt{\delta^2 + (x-\nu)^2}} e^{\delta\sqrt{\alpha^2-\beta^2}+\beta(x-\nu)}\,, \quad x \in \mathbb{R}$$

and convolution property for the characteristic functions given through

$$\varphi_{NIG(\alpha,\beta,\delta,\nu)}(ts) = \varphi_{NIG(\alpha,\beta,t\delta,t\nu)}(s), \quad t > 0\,, s \in \mathbb{R}$$

where K_1 denotes the modified Bessel function of third kind of order 1.

	$i=0$	$i=1$	$i=2$	$i=3$
K^i		180	60	105
S_0^i	12540.5	193.5399	65.12	108.3
μ_i	-0.0001370	-0.0001442	-0.0000121	-0.0001247
σ_i	0.0143478	0.0210187	0.0221977	0.0187162
α_i	50.7164	26.015486	32.6959583	42.06498
β_i	4.9547093	1.0846819	1.0591878	-0.6504885
δ_i	0.0103575	0.0110434	0.0157737	0.0131712
ν_i	-0.0012570	-0.0008272	-0.0007756	-0.0000870

	$i=4$	$i=5$	$i=6$
K^i	13	90	230
S_0^i	11.5740	86.5199	212.1999
μ_i	0.0009175	-0.0003305	-0.0005275
σ_i	0.0286615	0.0150850	0.0187999
α_i	25.37659	57.057185	44.6539402
β_i	0.3022331	4.0909649	-3.2505242
δ_i	0.0202047	0.0120536	0.0151163
ν_i	0.0002623	-0.0013079	0.0004002

Table 1: Strikes K^i (chosen), initial values $S_0^i = s_{2540}^i$ (observed) and parameters μ_i , σ_i , α_i , β_i , δ_i , ν_i describing the daily log-increments of $(S_t^i)_t$ under Assumption (1a) resp. (1b) (estimated).

Note that for $\nu \to \infty$ the t-copula passes into a Gaussian copula. In contrast to Gaussian copulas, t-copulas exhibit tail-dependencies with equal coefficients of lower resp. upper tail dependence

$$\lambda_l = \lambda_u = 2t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{1-\rho}/\sqrt{1+\rho}\right) \tag{25}$$

where t_ν denotes the standard univariate Student's t-distribution function with ν degrees of freedom (see Demarta and McNeil (2005)).

Application to real market data

As data set, we take the daily adjusted close data from yahoo finance from 23/04/2008 to 20/04/2018. It contains the values of 2540 trading days for 7 assets (with some missing data) which we denote by $(s_k^0, s_k^1, \dots, s_k^6)_{1\le k\le 2540}$, see Figure 1. More precisely, $(s_k^0)_k$ are the adjusted close data of "DAX PERFORMANCE-INDEX (GDAXI)", $(s_k^1)_k$ of "Allianz SE (ALV.DE)", $(s_k^2)_k$ of "Daimler AG (DAI.DE)", $(s_k^3)_k$ of "Siemens Aktiengesellschaft (SIE.DE)", $(s_k^4)_k$ of "Deutsche Bank Aktiengesellschaft (DBK.DE)", $(s_k^5)_k$ from "SAP SE (SAP.DE)" and $(s_k^6)_k$ of "adidas AG (ADS.DE)".

Denote by $(x_k^i)_k$, $x_k^i := \log s^i_{2540-t_{k-1}} - \log s^i_{2540-t_k}$, the historical time T-log-returns of the i-th asset (see Figure 1 in the case $T=1$). Hence, for $T=15$ resp. $T=50$ resp. $T=100$, the sequence $(x_k^i, x_k^0)_k$ consists of 169 resp. 50 resp.

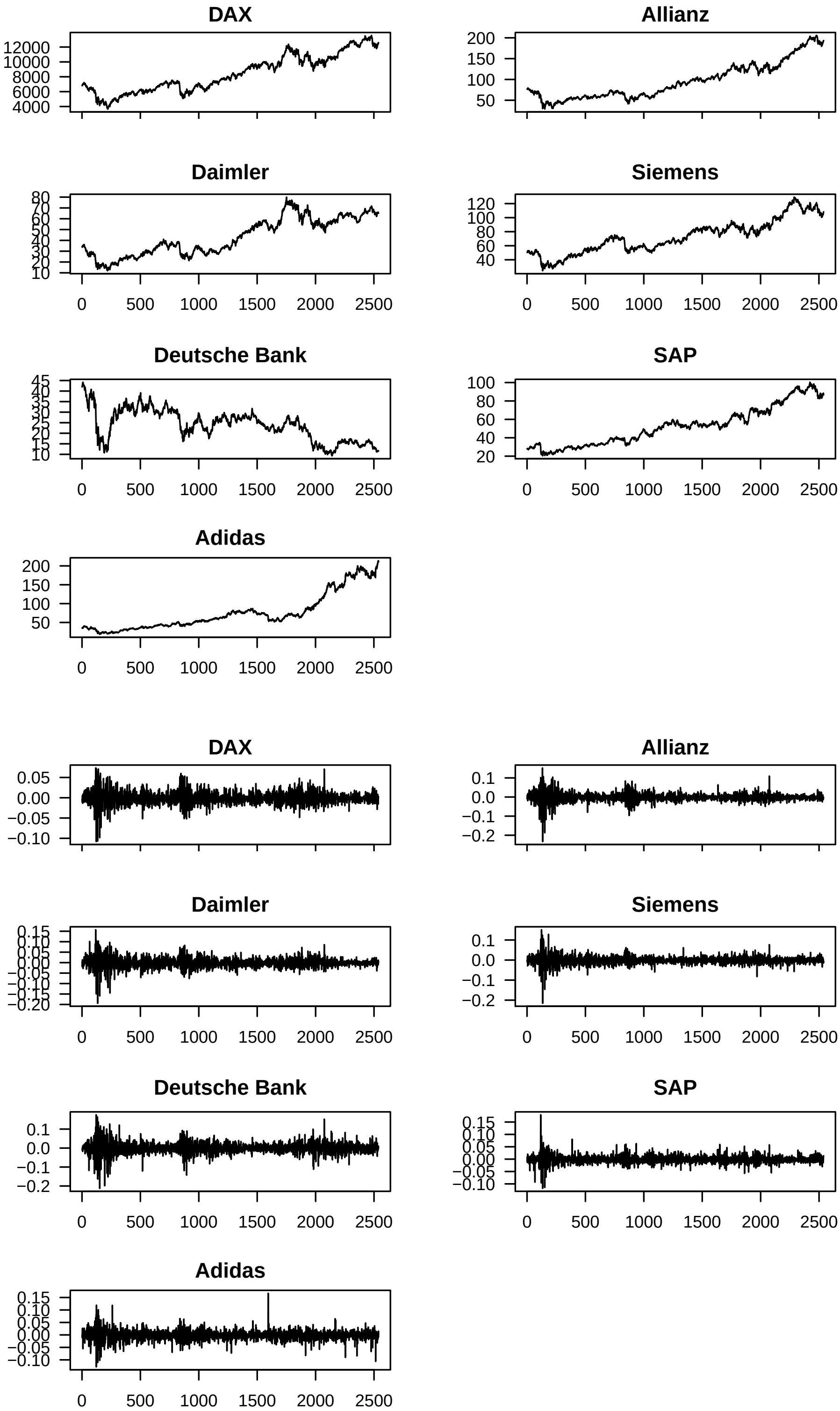

Figure 1: Daily adjusted close data resp. the log-returns of the underlying assets from 23/04/2008 to 20/04/2018; x-axis: n-th trading day beginning with 23/04/2008; y-axis: adjusted closing price resp. log-return of the underlying asset.

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
$\hat{\rho}_i^{15}$	0.8225	0.8808	0.8573	0.7595	0.7193	0.6466
$\underline{\rho}_i^{15}$	0.7767	0.8487	0.8196	0.7000	0.6519	0.5661
$\hat{\rho}_i^{50}$	0.8308	0.8827	0.8264	0.6445	0.7545	0.6246
$\underline{\rho}_i^{50}$	0.7401	0.8171	0.7338	0.4823	0.6312	0.4563
$\hat{\rho}_i^{100}$	0.8437	0.9317	0.8217	0.7374	0.7320	0.5716
$\underline{\rho}_i^{100}$	0.7081	0.8671	0.6703	0.5328	0.5244	0.2905

Table 2: Empirical Pearson correlation $\hat{\rho}_i^T$ between the T-days log-returns of the i-th underlying asset and the DAX estimated from log-return data $(x_k^i, x_k^0)_k$ over T days for $T = 15$, $T = 50$ resp. $T = 100$ trading days; $\underline{\rho}_i^T$ denotes the lower bound of the 95%-confidence interval for $\hat{\rho}_i^T$ under a bivariate normality assumption.

25 pairs of data. Table 2 shows the empirical correlation $\hat{\rho}_i^T$ of $(x_k^i, x_k^0)_k$ (which estimates the correlation of $(\xi_1^i, \xi_0^i))$ and a lower bound $\underline{\rho}_i^T$ for the one-sided 95%-confidence interval for $\hat{\rho}_i^T$. This justifies a determination of $I^i := [\underline{\rho}_i^T, 1]$ for the unspecified intervals I^i in Assumption (3).

Since Y_T^i is an increasing resp. decreasing transformation of S_T^i we can choose

$$\begin{aligned} C_{Y_T^i, S_T^0} &= C_{S_T^i, S_T^0} = C_{\xi_1^i, \xi_1^0} && \text{for } i = 1,2,3 \quad \text{resp.} \\ C_{Y_T^i, S_T^0} &= C_{-S_T^i, S_T^0} = C_{-\xi_1^i, \xi_1^0} && \text{for } i = 4,5,6\,. \end{aligned}$$

Note that the copula $C_{Y_T^i, S_T^0}$ may not be uniquely determined. This leads to the sets of constraints $\mathcal{S}^i$ for $C_{Y_T^i, S_T^0}$ given by

$$\begin{aligned} \mathcal{S}^i &= \{C_\nu^\rho | \rho \geq \underline{\rho}^T\} && \text{for } i = 1,2,3\,, \text{ resp.} \\ \mathcal{S}^i &= \{C_\nu^\rho | \rho \leq \overline{\rho}^T\} && \text{for } i = 4,5,6\,, \end{aligned}$$

where $\underline{\rho}^T := \min_{i=1,2,3} \underline{\rho}_i^T$ and $\overline{\rho}^T := \max_{i=4,5,6} -\underline{\rho}_i^T$.

Now, Corollary 3.13 and (2) yield

$$\begin{aligned} \Sigma_T &\leq_{cx} \sum_{i=1}^6 F_{Y_T^i}^{-1}\left(f_i(Z,\varepsilon)\right) =: \Sigma_{T,\underline{\rho}^T,\overline{\rho}^T,\nu}^c \\ &\leq_{cx} \sum_{i=1}^6 F_{Y_T^i}^{-1}(U) =: \Sigma_T^c \end{aligned} \tag{26}$$

for $U \sim U(0,1)$ where

$$F_{Y_T^i}^{-1}(z) = \begin{cases} \left(S_0^i \exp\left(F_{L_T^i}^{-1}(z)\right) - K_i\right)_+ & \text{for } i \in \{1,2,3\}\,, \\ \left(K_i - S_0^i \exp\left(F_{L_T^i}^{-1}(1-z)\right)\right)_+ & \text{for } i \in \{4,5,6\} \end{cases}$$

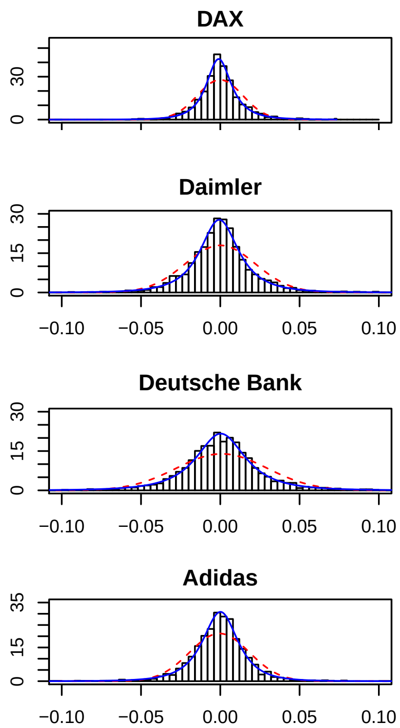

Figure 2: Histograms of the daily log-returns and fitted Gaussian (dashed) and NIG (solid) density with estimated parameters given in Table 1.

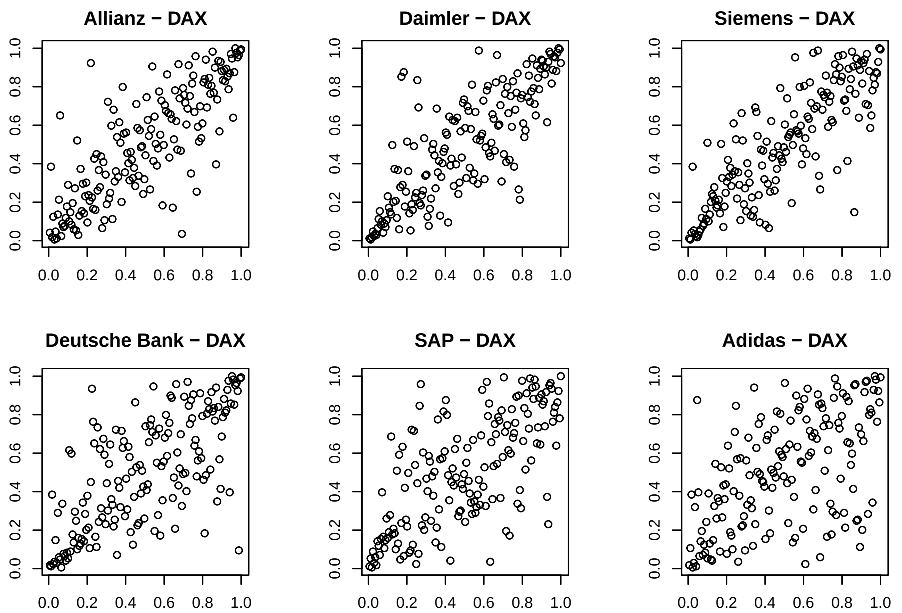

Figure 3: Plots of the empirical copulas between the T-day log-returns of each asset and the DAX for $T = 15$ trading days.

are the quantile functions of the calls resp. puts Y_T^i. Further, $(f_i(Z,\varepsilon))_i$ given by

$$f_i(Z,\varepsilon) := f(\eta_i,\nu,Z,\varepsilon) := t_\nu\left(\eta_i\, t_\nu^{-1}(Z) + \sqrt{\frac{(\nu + t_\nu^{-1}(Z)^2)(1-\eta_i^2)}{\nu+1}}\; t_{\nu+1}^{-1}(\varepsilon)\right)$$

with $\eta_i = \underline{\rho}^T$ if $i = 1,2,3$ and $\eta_i = \overline{\rho}^T$ if $i = 4,5,6$ is the conditionally on Z comonotonic random vector for random variables $Z,\varepsilon \sim U(0,1)$ that are independent. Note that the distribution function of $(f(\rho,\nu,Z,\varepsilon),Z)$ is the t-copula with correlation ρ and ν degrees of freedom (see Aas et al. (2009)). Further, the marginalization property of elliptical distributions implies that $(f(\rho_1,\nu,Z,\varepsilon), f(\rho_2,\nu,Z,\varepsilon))$ follows a t-copula with correlation parameter

$$M(\rho_1,\rho_2) := \rho_1\rho_2 + \sqrt{1-\rho_1^2}\sqrt{1-\rho_2^2}$$

and ν degrees of freedom. Hence, $(f_i(Z,\varepsilon), f_j(Z,\varepsilon))$ is comonotonic if $1 \le i,j \le 3$ or $4 \le i,j \le 6$ (cp. Remark 3.14(a)). Otherwise, it follows a t-copula with correlation $M(\underline{\rho}^T, \overline{\rho}^T)$.

As a consequence of (1) and (26) we obtain

$$\Psi(\Sigma_T) \le \Psi(\Sigma^c_{T,\underline{\rho}^T,\overline{\rho}^T,\nu}) \le \Psi(\Sigma_T^c)\,. \tag{27}$$

More specifically, let Ψ be the *Average Value-at-Risk* at level λ (also known as *Expected Shortfall*) defined by

$$\mathrm{AVaR}_\lambda(S) := \frac{1}{1-\lambda}\int_\lambda^1 F_S^{-1}(t)\,\mathrm{d}t\,, \quad \lambda \in (0,1)\,.$$

It is well-known that AVaR_λ is a convex, law-invariant risk measure. In Tables 3, (4) resp. (5), we compare the improved risk bound $\mathrm{AVaR}_\lambda(\Sigma^c_{T,Z,(\eta_i),\nu})$ given by (27) with the standard comonotonic bound $\mathrm{AVaR}_\lambda(\Sigma^c_T)$ in Models Gauss and NIG (7million simulated points) for different λ and ν and for $T = 15$ resp. $T = 50$ resp. $T = 100$ trading days.

We observe that both the improved and the standard portfolio risk bounds $\mathrm{AVaR}_\lambda(\Sigma^c_{T,\underline{\rho}^T,\overline{\rho}^T,\nu})$ resp. $\mathrm{AVaR}_\lambda(\Sigma^c_T)$ depend for high levels λ on the model for the univariate margins of the summands and their tails. The fatter tails of the NIG distribution yield higher risks. But for larger times T to maturity, we see that the differences are less significant. This can be explained by the fact that the parameters $\delta_i' = T\delta_i$ and $\alpha_i' = \alpha_i$ (see Table 1) of L_T^i are quite large for large T and thus $F_{L_T^i}$ is approximately normal with variance δ_i'/α_i' (see Barndorff-Nielsen (1978, p.153)). In our application, Model NIG fits the data better than Model Gauss (see Figure 2). In contrast, for levels $\lambda \le 0.95$ the results in this application nearly coincide for Models Gauss and NIG.

Further, we observe that the improvement of the risk bounds depends on

Model Gauss $\mathbb{E}[\Sigma_{15}] = 51.5$	$\mathrm{AVaR}_\lambda\left(\Sigma^c_{15,\underline{\rho}^{15},\overline{\rho}^{15},\nu}\right)$ $\nu = 3$		$\nu = 10$		$\nu = \infty$		$\mathrm{AVaR}_\lambda(\Sigma^c_{15})$
$\lambda = 0.5$	76.3	(29.0%)	76.7	(28.1%)	76.8	(27.7%)	86.5
$\lambda = 0.8$	99.2	(28.1%)	99.4	(27.5%)	99.4	(27.5%)	117.5
$\lambda = 0.9$	114.8	(25.2%)	113.8	(26.3%)	113.3	(26.9%)	136.1
$\lambda = 0.95$	129.6	(22.5%)	126.9	(25.2%)	125.6	(26.5%)	152.3
$\lambda = 0.99$	161.9	(17.0%)	154.3	(22.6%)	150.1	(25.8%)	184.4
$\lambda = 0.995$	174.8	(15.0%)	165.3	(21.6%)	159.4	(25.6%)	196.6
$\lambda = 0.999$	202.4	(11.8%)	189.3	(19.5%)	179.4	(25.2%)	222.6
Model NIG $\mathbb{E}[\Sigma_{15}] = 51.1$	$\mathrm{AVaR}_\lambda\left(\Sigma^c_{15,\underline{\rho}^{15},\overline{\rho}^{15},\nu}\right)$ $\nu = 3$		$\nu = 10$		$\nu = \infty$		$\mathrm{AVaR}_\lambda(\Sigma^c_{15})$
$\lambda = 0.5$	75.2	(28.7%)	75.6	(27.7%)	75.8	(27.1%)	85.0
$\lambda = 0.8$	98.0	(27.6%)	98.2	(27.4%)	98.3	(27.2%)	115.9
$\lambda = 0.9$	114.2	(25.5%)	113.2	(26.6%)	112.8	(27.0%)	135.7
$\lambda = 0.95$	130.0	(23.2%)	127.3	(25.8%)	126.2	(27.0%)	153.9
$\lambda = 0.99$	167.0	(18.5%)	159.0	(24.1%)	155.0	(26.9%)	193.4
$\lambda = 0.995$	183.1	(16.8%)	172.6	(23.4%)	167.2	(26.9%)	209.8
$\lambda = 0.999$	221.2	(13.6%)	205.1	(21.8%)	195.9	(26.5%)	248.0

Table 3: Comparison of the improved risk bound $\mathrm{AVaR}_\lambda(\Sigma^c_{15,\underline{\rho}^{15},\overline{\rho}^{15},\nu})$ with the standard comonotonic risk bound $\mathrm{AVaR}_\lambda(\Sigma^c_{15})$ for $\mathrm{AVaR}_\lambda(\Sigma_{15})$ in Model Gauss resp. NIG for $T = 15$ trading days for different levels λ, for different ν and for fixed $\underline{\rho}^{15} = .7767$ and $\overline{\rho}^{15} = -0.5661$. The relative DU-improvement given by $1 - (\mathrm{AVaR}_\lambda(\Sigma^c_{15,\underline{\rho}^{15},\overline{\rho}^{15},\nu}) - \mathbb{E}[\Sigma_{15}])/(\mathrm{AVaR}_\lambda(\Sigma^c_{15}) - \mathbb{E}[\Sigma_{15}])$ is displayed in brackets.

Model Gauss $\mathbb{E}[\Sigma_{50}] = 67.1$	$\mathrm{AVaR}_\lambda\left(\Sigma^c_{50,\underline{\rho}^{50},\overline{\rho}^{50},\nu}\right)$ $\nu = 3$		$\nu = 10$		$\nu = \infty$		$\mathrm{AVaR}_\lambda(\Sigma^c_{50})$
$\lambda = 0.5$	110.1	(21.9%)	110.4	(21.3%)	110.5	(21.1%)	122.1
$\lambda = 0.8$	154.4	(22.1%)	154.6	(22.0%)	154.6	(21.9%)	179.2
$\lambda = 0.9$	185.0	(19.6%)	183.7	(20.5%)	182.9	(21.0%)	213.7
$\lambda = 0.95$	213.8	(17.2%)	210.2	(19.2%)	208.3	(20.3%)	244.3
$\lambda = 0.99$	275.7	(12.6%)	266.0	(16.7%)	260.3	(19.1%)	306.0
$\lambda = 0.995$	300.1	(11.1%)	288.4	(15.8%)	280.8	(18.7%)	329.9
$\lambda = 0.999$	354.8	(8.5%)	338.1	(13.8%)	325.4	(17.8%)	381.4
Model NIG $\mathbb{E}[\Sigma_{50}] = 66.2$	$\mathrm{AVaR}_\lambda\left(\Sigma^c_{50,\underline{\rho}^{50},\overline{\rho}^{50},\nu}\right)$ $\nu = 3$		$\nu = 10$		$\nu = \infty$		$\mathrm{AVaR}_\lambda(\Sigma^c_{50})$
$\lambda = 0.5$	108.4	(21.8%)	108.7	(21.2%)	108.7	(21.1%)	120.1
$\lambda = 0.8$	152.1	(22.1%)	152.3	(21.9%)	152.3	(21.9%)	176.4
$\lambda = 0.9$	182.8	(19.7%)	181.5	(20.6%)	180.8	(21.1%)	211.4
$\lambda = 0.95$	212.2	(17.4%)	208.6	(19.4%)	206.7	(20.5%)	242.9
$\lambda = 0.99$	277.4	(13.0%)	267.4	(17.1%)	261.7	(19.5%)	309.1
$\lambda = 0.995$	304.5	(11.6%)	291.9	(16.3%)	284.3	(19.1%)	335.8
$\lambda = 0.999$	366.0	(9.2%)	349.1	(14.3%)	336.0	(18.3%)	396.5

Table 4: Comparison of the improved risk bound $\mathrm{AVaR}_\lambda(\Sigma^c_{50,\underline{\rho}^{50},\overline{\rho}^{50},\nu})$ with the standard comonotonic risk bound $\mathrm{AVaR}_\lambda(\Sigma^c_{50})$ for $\mathrm{AVaR}_\lambda(\Sigma_{50})$ in Model Gauss resp. NIG for $T = 50$ trading days for different levels λ, for different ν and for fixed $\underline{\rho}^{50} = .7338$ and $\overline{\rho}^{50} = -0.4563$. The relative DU-improvement given by $1 - (\mathrm{AVaR}_\lambda(\Sigma^c_{50,\underline{\rho}^{50},\overline{\rho}^{50},\nu}) - \mathbb{E}[\Sigma_{50}])/(\mathrm{AVaR}_\lambda(\Sigma^c_{50}) - \mathbb{E}[\Sigma_{50}])$ is displayed in brackets.

Model Gauss $\mathbb{E}[\Sigma_{100}] = 83.8$	$\mathrm{AVaR}_\lambda\left(\Sigma^c_{100,\underline{\rho}^{100},\overline{\rho}^{100},\nu}\right)$ $\nu=3$		$\nu=10$		$\nu=\infty$		$\mathrm{AVaR}_\lambda(\Sigma^c_{100})$
$\lambda=0.5$	145.1	(14.0%)	145.3	(13.6%)	145.5	(13.4%)	155.1
$\lambda=0.8$	213.7	(14.6%)	214.0	(14.4%)	214.2	(14.3%)	235.9
$\lambda=0.9$	260.5	(12.5%)	259.5	(13.0%)	259.0	(13.3%)	285.8
$\lambda=0.95$	304.1	(10.7%)	301.2	(11.9%)	299.5	(12.5%)	330.4
$\lambda=0.99$	397.2	(7.5%)	389.4	(9.8%)	384.5	(11.3%)	422.7
$\lambda=0.995$	434.8	(6.7%)	425.1	(9.1%)	418.5	(10.8%)	459.2
$\lambda=0.999$	518.0	(4.7%)	504.6	(7.7%)	494.4	(9.9%)	539.5
Model NIG $\mathbb{E}[\Sigma_{100}] = 82.5$	$\mathrm{AVaR}_\lambda\left(\Sigma^c_{100,\underline{\rho}^{100},\overline{\rho}^{100},\nu}\right)$ $\nu=3$		$\nu=10$		$\nu=\infty$		$\mathrm{AVaR}_\lambda(\Sigma^c_{100})$
$\lambda=0.5$	142.6	(14.0%)	142.9	(13.6%)	143.0	(13.4%)	152.4
$\lambda=0.8$	210.1	(14.6%)	210.5	(14.4%)	210.6	(14.3%)	231.9
$\lambda=0.9$	256.7	(12.5%)	255.7	(13.0%)	255.0	(13.3%)	281.6
$\lambda=0.95$	300.5	(10.7%)	297.6	(11.9%)	295.7	(12.6%)	326.6
$\lambda=0.99$	395.7	(7.6%)	388.2	(9.9%)	382.7	(11.5%)	421.6
$\lambda=0.995$	435.1	(6.7%)	425.7	(9.1%)	418.7	(11.0%)	460.2
$\lambda=0.999$	524.4	(5.1%)	513.2	(7.5%)	501.1	(10.1%)	548.3

Table 5: Comparison of the improved risk bound $\mathrm{AVaR}_\lambda(\Sigma^c_{100,\underline{\rho}^{100},\overline{\rho}^{100},\nu})$ with the standard comonotonic risk bound $\mathrm{AVaR}_\lambda(\Sigma^c_{100})$ for $\mathrm{AVaR}_\lambda(\Sigma_{100})$ in Model Gauss resp. NIG for $T=100$ trading days for different levels λ, for different ν and for fixed $\underline{\rho}^{100} = .6703$ and $\overline{\rho}^{100} = -0.2905$. The relative DU-improvement given by $1-(\mathrm{AVaR}_\lambda(\Sigma^c_{100,\underline{\rho}^{100},\overline{\rho}^{100},\nu}) - \mathbb{E}[\Sigma_{100}])/(\mathrm{AVaR}_\lambda(\Sigma^c_{100}) - \mathbb{E}[\Sigma_{100}])$ is displayed in brackets.

the degree of freedom ν of the constraining t-copula families $\mathcal{S}^i$. The smaller the parameter ν the higher is the tail-dependence of the (t-)copula of

$$\left(\sum_{i=1}^{3} F^{-1}_{Y^i_T}(f_i(Z,\varepsilon)), \sum_{i=4}^{6} F^{-1}_{Y^i_T}(f_i(Z,\varepsilon))\right), \tag{28}$$

see(25). This means that extreme tail events occur more often simultaneously in the components which leads to higher risks. The empirical data exhibit tail-dependencies, see Figure 3. Thus, a t-copula with degree of freedom ν not too large should be preferred to a Gaussian copula in this application.

We see that the improvement of the standard DU-interval $[\mathbb{E}\Sigma_T, \mathrm{AVaR}_\lambda(\Sigma^c_T)]$ is largest for $T=15$ trading days (about 20% to 30%) and smallest for $T=100$ trading days (about 10% to 15%). A large improvement means a small correlation parameter for (28) which is achieved if T is small, i.e.

$$M\left(\underline{\rho}^{15},\overline{\rho}^{15}\right) = 0.0795 < M\left(\underline{\rho}^{50},\overline{\rho}^{50}\right) = 0.2697 < M\left(\underline{\rho}^{100},\overline{\rho}^{100}\right) = 0.5154,$$

because in this case the underlying data sets $(x^i_k)_k$ are larger such that the lower bounds $\underline{\rho}^T_i$ for the 95%-confidence intervals for $\hat{\rho}^T_i$ are larger. Thus, the intervals I^i could be chosen tighter for smaller T.

5 Appendix

Proof of the convergence in (15)**:** By some technical but standard arguments of integration theory one can show that

$$\sum_{k=1}^{n} \min\left\{\int_{(\frac{k-1}{n},\frac{k}{n})} h(t)\,\mathrm{d}\lambda(t), 0\right\} \xrightarrow{n\to\infty} \int_{(0,1)} \min\{h(t),0\}\,\mathrm{d}\lambda(t) \tag{29}$$

for any $\mathcal{B}((0,1))$-measurable function h with $-1 \leq h \leq 1$. Then, we deduce the statement in (15) by induction as follows: Consider measurable functions $h_1, \ldots, h_d$ with $0 \leq h_i \leq 1$. For the base case $d = 2$ set $h = h_2 - h_1$. Due to the convergence in (29) it holds

$$\begin{aligned}
&\sum_{k=1}^{n} \min\left\{\int_{(\frac{k-1}{n},\frac{k}{n})} h_1(t)\,\mathrm{d}t, \int_{(\frac{k-1}{n},\frac{k}{n})} h_2(t)\,\mathrm{d}t\right\} \\
&= \sum_{k=1}^{n} \min\left\{\int_{(\frac{k-1}{n},\frac{k}{n})} h(t)\,\mathrm{d}t, 0\right\} + \int_0^1 h_1(t)\,\mathrm{d}t \\
&\xrightarrow{n\to\infty} \int_0^1 \min\{h(t),0\}\,\mathrm{d}t + \int_0^1 h_1(t)\,\mathrm{d}t = \int_0^1 \min\{h_1(t),h_2(t)\}\,\mathrm{d}t\,.
\end{aligned}$$

For the induction step set $h = h_d - \min\{h_1, \ldots, h_{d-1}\}$. Then, we obtain again with (29) and the induction hypothesis that

$$\begin{aligned}
&\int_0^1 \min_{1\leq i\leq d}\{h_i(t)\}\,\mathrm{d}t \leq \sum_{k=1}^{n} \min_{1\leq i\leq d}\left\{\int_{(\frac{k-1}{n},\frac{k}{n})} h_i(t)\,\mathrm{d}t\right\} \\
&\leq \sum_{k=1}^{n} \min\left\{\int_{(\frac{k-1}{n},\frac{k}{n})} h(t)\,\mathrm{d}t, 0\right\} + \sum_{k=1}^{n} \min_{1\leq i\leq d-1}\left\{\int_{(\frac{k-1}{n},\frac{k}{n})} h_i(t)\,\mathrm{d}t\right\} \\
&\xrightarrow{n\to\infty} \int_0^1 \min\{h(t),0\}\,\mathrm{d}t + \int_0^1 \min_{1\leq i\leq d-1}\{h_i(t)\}\,\mathrm{d}t = \int_0^1 \min_{1\leq i\leq d}\{h_i(t)\}\,\mathrm{d}t\,,
\end{aligned}$$

where both inequalities hold true due to Jensen's inequality. ■

Proof of Proposition 3.9: Denote by $b^\iota = (b^\iota_{kl})_{1\leq k,l\leq n}$ the corresponding $\frac{1}{n}$-scaled doubly stochastic matrix of B^ι, $\iota = 1,2$. Consider the following algorithm that constructs the sequence $(E^i)_{1\leq i\leq m}$ adjusting in each step the $\frac{1}{n}$-scaled doubly stochastic matrix b^1 to b^2 by a simple supermodular transfer that preserves the $\leq_{\partial\Delta}$-relation with respect to each A^j.

1. Define $E^0 = B^1$ with $\frac{1}{n}$-scaled doubly stochastic matrix $e^0 = b^1$. Set $i = 0$ and $k = n$.

2. Mass compensation in line k : If $e^i_{kl} = b^2_{kl}$ for all $1 \leq l \leq n$, go to step (3). Otherwise let $l_* := \min\{l | e^i_{kl} < b^2_{kl}\}$, $l^* := \max\{l | e^i_{kl} > b^2_{kl}\}$, $t_* := \frac{l_*}{n}$,

$t^* := \frac{l^*}{n}$, $v_* := 1 - \frac{k-1}{n}$. Define $u_* := (u_{*j})_{1\leq j\leq d}$ and $u^* := (u^*_j)_{1\leq j\leq d}$ by

$$u_{*j} : = \min\{u_j \in \mathbb{G}^1_n | \Delta^2_n E^i(v_*, t^*) \leq \Delta^2_n A^j(u_j, t^*)\}\,, \tag{30}$$

$$u^*_j : = \max\{u_j \in \mathbb{G}^1_n | \Delta^2_n E^i(v_*, t_*) \geq \Delta^2_n A^j(u_j, t_*)\} + \tfrac{1}{n}\,. \tag{31}$$

Define the transferred mass

$$\begin{aligned}\eta : &= \eta(e^i_{kl_*}, e^i_{kl^*}, b^2_{kl_*}, b^2_{kl^*}, (\Delta^2_n A^j(u^*_j, t_*))_j, (\Delta^2_n A^1(u_{*1} - \tfrac{1}{n}, t^*))_j, \\ &\qquad\qquad \Delta^2_n E^i(v_*, t_*), \Delta^2_n E^i(v_*, t^*)) \\ : &= \min\Big\{ b^2_{kl_*} - e^i_{kl_*}, e^i_{kl^*} - b^2_{kl^*}, \min_j\{\Delta^2_n A^j(u^*_j, t_*)\} - \Delta^2_n E^i(v_*, t_*), \\ &\qquad\qquad \Delta^2_n E^i(v_*, t^*) - \max_j\{\Delta^2_n A^j((u_{*j} - \tfrac{1}{n}) \vee 0, t^*)\}\Big\}\end{aligned}$$

Define E^{i+1} via the $\frac{1}{n}$-scaled doubly stochastic matrix

$$e^{i+1}_{\iota\kappa} = \begin{cases} e^i_{\iota\kappa} - \eta & \text{if } (\iota,\kappa) \in \{(k,l^*), (k-1,l_*)\}\,, \\ e^i_{\iota\kappa} + \eta & \text{if } (\iota,\kappa) \in \{(k,l_*), (k-1,l^*)\}\,, \\ e^i_{\iota\kappa} & \text{else} \end{cases} \tag{32}$$

for $\iota, \kappa \in \{1, \ldots, n\}$. Set $i = i + 1$. Repeat step (2).

3. If $k = 2$ set $m = i$ and stop the algorithm. Otherwise set $k = k - 1$ and go to step (2).

First, we show that $\eta > 0$. From the definition of l_* resp. l^* it holds $b^2_{kl_*} - e^i_{kl_*} > 0$ resp. $e^i_{kl^*} - b^2_{kl^*} > 0$. Further, since for $\kappa > k$ holds $e^i_{\kappa l} = b^2_{\kappa l}$ for all $1 \leq l \leq n$, we obtain

$$\begin{aligned}&\Delta^2_n E^i(v_*, t^*) > \Delta^2_n B^2(v_*, t^*) \geq 0 \quad \text{and} \\ &\Delta^2_n E^i(v_*, t_*) < \Delta^2_n B^2(v_*, t_*) \leq 1\,.\end{aligned}$$

This yields with the definition of u^*_j resp. u_{*j}

$$\begin{aligned}\Delta^2_n A^j(u^*_j, t_*) - \Delta^2_n E^i(v_*, t_*) > 0 \quad \text{resp.} \\ \Delta^2_n E^i(v_*, t^*) - \Delta^2_n A^j((u_{*j} - \tfrac{1}{n}) \vee 0, t^*) > 0\end{aligned}$$

for all $1 \leq j \leq d$.

Secondly, we observe that for each $(u, t) \in \mathbb{G}^2_n \setminus \{(v_*, t_*), (v_*, t^*)\}$ holds by construction of E^{i+1} that

$$\Delta^2_n E^{i+1}(u, t) = \sum_{u' \leq u} e^{i+1}(u', t) = \sum_{u' \leq u} e^i(u', t) = \Delta^2_n E^i(u, t)\,. \tag{33}$$

Thirdly, we show

$$E^i \leq_{s\partial\Delta} B^2 \tag{34}$$

for $0 \le i \le m$ by induction. For $i = 0$, the statement is given by the assumption that $B^1 \le_{s\partial\Delta} B^2$. Suppose that $E^i \le_{s\partial\Delta} B^2$ for an $i \in \{0, 1, \dots, m-1\}$. We obtain with (33) that

$$\Delta_n^2 B^2(u,t) - \Delta_n^2 E^{i+1}(u,t) = \Delta_n^2 B^2(u,t) - \Delta_n^2 E^i(u,t) = 0 \quad \text{for } u < v_* . \tag{35}$$

The last equality holds because the lines of $(e^i_{\kappa\iota})$ are for $\kappa > k$, i.e. $u < v_*$, already adjusted to the lines of $(b^2_{\kappa\iota})$.

For $u > v_*$ we obtain from (33) that $\Delta_n^2 B^2(u,t) - \Delta_n^2 E^{i+1}(u,t) = \Delta_n^2 B^2(u,t) - \Delta_n^2 E^i(u,t)$ where the latter has no $(-,+)$-sign change as assumed.

Consider the case $u = v_*$. Then $\Delta_n^2 B^2(u,t) - \Delta_n^2 E^i(u,t) = b^2_{kl} - e^i_{kl}$ with $l = tn$ has exactly one sign change in t which is from $+$ to $-$ as assumed. Hence, it follows

$$l_* < l^* \quad \text{resp.} \quad t_* < t^* \tag{36}$$

and $\Delta_n^2 B^2(u,t_*) - \Delta_n^2 E^i(u,t_*) > 0$ resp. $\Delta_n^2 B^2(u,t^*) - \Delta_n^2 E^i(u,t^*) < 0$. Since $\eta \le \min\{b^2_{kl_*} - e^i_{kl_*}, e^i_{kl^*} - b^2_{kl^*}\}$ we get together with (35) and (32) that

$$\Delta_n^2 B^2(u,t_*) - \Delta_n^2 E^{i+1}(u,t_*) = b^2_{kl_*} - e^{i+1}_{kl_*} = b^2_{kl_*} - e^i_{kl_*} - \eta \ge 0$$

and

$$\Delta_n^2 B^2(u,t^*) - \Delta_n^2 E^{i+1}(u,t^*) = b^2_{kl^*} - e^{i+1}_{kl^*} = b^2_{kl^*} - e^i_{kl^*} + \eta \le 0 .$$

Hence, also $\Delta_n^2 B^2(u,t) - \Delta_n^2 E^{i+1}(u,t)$ has no $(-,+)$-sign change in t.

Fourthly, we observe from the proof of (34) that there exists a finite $i \in \mathbb{N}$ such that mass in the lines of $(e^i_{kl})_{kl}$ has been adjusted to $(c_{kl})_{kl}$ for all $k = n, \dots, 2$. Then, since both e^i and b^2 are (signed) $\frac{1}{n}$-scaled doubly stochastic also $e^i_{1l} = b^2_{1l}$ holds. Thus, it is sufficient to stop the algorithm setting $m := i$ if $k = 2$ and $e^i_{kl} = b^2_{kl}$ for all $1 \le l \le n$. This proves $E^m = B^2$.

(i): For each $i \in \{0, \dots, m\}$ it follows by construction that $\sum_{\iota=1}^n e^i_{\iota\kappa} = 1$ for all κ and $\sum_{\kappa=1}^n e^i_{\iota\kappa} = 1$ for all ι. Note that elements of e^i can get negative. Thus P_{E^i} defines a signed probability measure on $\mathbb{G}_n^2$ for all i. Since $0 \ne P_{E^{i+1}} - P_{E^i}(x) = \pm\eta$ for exactly 4 points $x \in \mathbb{G}_n^2$ and mass is transferred from the off-diagonal onto the diagonal, see (32) and (36), $P_{E^{i+1}} - P_{E^i}$ indicates a simple supermodular transfer.

(ii): For each $u \ne v_*$ holds due to (33) that $\Delta_n^2 E^{i+1}(u,t) - \Delta_n^2 E^i(u,t) = 0$, and the left hand-side trivially has as a function in t no sign change. Due to (36) and the definition of E^{i+1} in(32) it follows that

$$\Delta_n^2 E^{i+1}(v_*,t_*) = \Delta_n^2 E^i(v_*,t_*) + \eta , \quad \text{and}$$
$$\Delta_n^2 E^{i+1}(v_*,t^*) = \Delta_n^2 E^i(v_*,t^*) - \eta .$$

This means that $\Delta_n^2 E^{i+1}(v_*, t) - \Delta_n^2 E^i(v_*, t)$ has exactly one sign change in t which is from $+$ to $-$.

(iii): We show the statement by induction. For $i = 0$ there is nothing to show. Let $i \in \{0, \dots, m-1\}$ and suppose that $A^j \leq_{\partial\Delta} E^i$. Then we immediately obtain with (33) in the case $u \neq v_*$ that

$$\Delta_n^2 E^{i+1}(u,t) - \Delta_n^2 A^j(z,t) = \Delta_n^2 E^i(u,t) - \Delta_n^2 A^j(z,t)$$

has no $(-,+)$-sign change in t for all $z \in \mathbb{G}_n^1$.

Consider the case $u = v_*$. Define the functions

$$\begin{aligned} f(t) &:= f_{u,z}^j(t) : = \Delta_n^2 B^2(u,t) - \Delta_n^2 A^j(z,t)\,, \\ g(t) &:= g_{u,z}^j(t) : = \Delta_n^2 E^{i+1}(u,t) - \Delta_n^2 A^j(z,t) \\ h(t) &:= h_{u,z}^j(t) : = \Delta_n^2 E^i(u,t) - \Delta_n^2 A^j(z,t)\,, \end{aligned}$$

for $t, z \in G_n^1$. Due to (36) and the definition of t_* and t^* mass in line k has already been adjusted to b^2 for $t < t_*$ and $t > t^*$, i.e. $e_{kl}^i = e_{kl}^{i+1} = b_{kl}^2$ for all $l < l_*$ and $l > l^*$. Hence, it holds

$$f(t) = g(t) = h(t) \quad \text{for } t < t_* \quad \text{or} \quad t > t^*\,.$$

Since $\eta \leq \min\{c_{kl_*} - e_{kl_*}^i, e_{kl^*}^i - c_{kl^*}\}$ we obtain due to the construction of E^{i+1} that

$$\begin{aligned} f(t_*) &\geq h(t_*) + \eta = g(t_*)\,, \quad \text{and} \\ f(t^*) &\leq h(t^*) - \eta = g(t^*)\,. \end{aligned} \tag{37}$$

Again by construction, it holds that

$$g(t) = h(t) \quad \text{for all } t_* < t < t^*\,.$$

We need to show that g has no $(-,+)$-sign change.

Assume that g has exactly one sign change *immediately after* $s \in \mathbb{G}_n^1$, i.e. $g(t) \leq 0$ for $t \leq s$, $g(t) < 0$ for an $t \leq s$, $g(t) \geq 0$ for $t > s$ and $g(t) > 0$ for an $t > s$. If $t_* \leq s < t^*$, then $h(t_*) < g(t_*) \leq 0 \leq g(t^*) < h(t^*)$, which is a contradiction because h has no $(-,+)$-sign change.

If $t_* < t^* \leq s$, then $h(t_*) < g(t_*) \leq 0$ and $0 < g(t) = h(t)$ for an $t > s$, which again is a contradiction. The case $s \leq t_* < t^*$ is analogous. Hence, the sign change of g cannot be from $-$ to $+$.

Assume that g has (at least) two sign changes, say immediately after s resp. $s' \in \mathbb{G}_n^1$. Then there exist $t_1 \leq s < t_2 \leq s' < t_3$ such that

$$g(t_1) < 0 < g(t_2) > 0 > g(t_3) \quad \text{or} \quad g(t_1) > 0 > g(t_2) < 0 < g(t_3)\,. \tag{38}$$

Consider the left case. Since $h(t) = g(t)$ for all $t \neq t_*, t^*$ and h has no $(-,+)$-sign change, we obtain from (37) that $t_1 = t^*$ or $t_2 = t_*$. If $t_1 = t^*$, then

$f(t_1) \leq g(t_1) < 0$, but $f(t_2) = g(t_2) > 0$ which is a contradiction to the assumption that f has no $(-,+)$-sign change. If $t_2 = t_*$, then $f(t_2) \geq g(t_2) > 0$ and $f(t_1) = g(t_1) < 0$, which again is a contradiction. The second case in (38) follows analogously. This completes the prove of (iii).

(iv): Define $F^i := A^1 \vee \cdots \vee A^d \vee E^i$ for $0 \leq i \leq m$. We show that $P_{F^{i+1}} - P_{F^i}$ is indicated by a simple submodular transfer for all $0 \leq i \leq m-1$. Consider the set

$$
\begin{aligned}
\mathcal{S}^i &:= \{(u, u_{d+1}) \in \mathbb{G}_n^{d+1} | F^i(u, u_{d+1}) > F^{i+1}(u, u_{d+1})\} \\
&= \Bigg\{ (u, u_{d+1}) | \ \frac{1}{n} \sum_{t \in \mathbb{G}_n^1} \min\{ \min_{1 \leq j \leq d} \{\Delta_n^2 A^j(u_j, t), \Delta_n^2 E^i(u_{d+1}, t)\} \\
&\qquad > \frac{1}{n} \sum_{t \in \mathbb{G}_n^1} \min\{ \min_{1 \leq j \leq d} \{\Delta_n^2 A^j(u_j, t)\}, \Delta_n^2 E^{i+1}(u_{d+1}, t)\} \Bigg\} \\
&= \Bigg\{ (u, u_{d+1}) | \sum_{t \in \{t_*, t^*\}} \min\{\min_j \{\Delta_n^2 A^j(u_j, t)\}, \Delta_n^2 E^i(u_{d+1}, t)\} \\
&\qquad > \sum_{t \in \{t_*, t^*\}} \min\{\min_j \{\Delta_n^2 A^j(u_j, t)\}, \Delta_n^2 E^{i+1}(u_{d+1}, t)\} \Bigg\} \\
&= \Big\{ (u, u_{d+1}) | \ u_{d+1} = v_*, \ \Delta_n^2 E^{i+1}(v_*, t_*) > \min_j \{\Delta_n^2 A^j(u_j, t_*)\}, \text{ and} \\
&\qquad \Delta_n^2 E^{i+1}(v_*, t^*) < \min_j \{\Delta_n^2 A^j(u_j, t^*)\} \Big\} \qquad (39) \\
&= \Big\{ (u, u_{d+1}) | \ u_{d+1} = v_*, \ \Delta_n^2 E^i(v_*, t_*) \geq \min_j \{\Delta_n^2 A^j(u_j, t_*)\}, \text{ and} \\
&\qquad \Delta_n^2 E^i(v_*, t^*) \leq \min_j \{\Delta_n^2 A^j(u_j, t^*)\} \Big\}. \qquad (40)
\end{aligned}
$$

The first equality holds by the definition of the discrete upper product. For the second equality we use that only the summands on the left and right hand-side depending on t_* and t^* differ.

From (iii) and (ii) we obtain with Proposition 3.6 that $F^i(x) - F^{i+1}(x) \geq 0$ for all $x \in \mathbb{G}_n^{d+1}$, and furthermore holds $\mathcal{S}^i \neq \emptyset$. Further, the set $\mathcal{S}^i$ is restricted to $u_{d+1} = v_*$ because $\Delta_n^2 E^i(v, t) = \Delta_n^2 E^{i+1}(v, t)$ for all $v \neq v_*$ and for all t, see (33). Then, the third equality holds due to (iii) and due to the fact that on the one hand $\Delta_n^2 E^{i+1}(v_*, t_*) + \Delta_n^2 E^{i+1}(v_*, t^*) = \Delta_n^2 E^i(v_*, t_*) + \Delta_n^2 E^i(v_*, t^*)$, and on the other hand $\Delta_n^2 E^{i+1}(v_*, t_*) > \Delta_n^2 E^i(v_*, t_*)$ and $\Delta_n^2 E^{i+1}(v_*, t^*) < \Delta_n^2 E^i(v_*, t^*)$.

For the fourth equality we show that

$$
\begin{aligned}
&\Delta_n^2 E^{i+1}(v_*, t^*) < \min_j \{\Delta_n^2 A^j(u_j, t^*)\} \\
\iff \quad &\Delta_n^2 E^i(v_*, t^*) \le \min_j \{\Delta_n^2 A^j(u_j, t^*)\}
\end{aligned}
\tag{41}
$$

and

$$
\begin{aligned}
&\Delta_n^2 E^{i+1}(v_*, t_*) > \min_j \{\Delta_n^2 A^j(u_j, t_*)\} \\
\iff \quad &\Delta_n^2 E^i(v_*, t_*) \ge \min_j \{\Delta_n^2 A^j(u_j, t_*)\} \,.
\end{aligned}
\tag{42}
$$

Assume that the right side in (41) holds. Then the left side follows directly from $\Delta_n^2 E^{i+1}(v_*, t^*) = \Delta_n^2 E^i(v_*, t^*) - \eta$. Conversely, assume that there exist j and u_j such that $\Delta_n^2 E^{i+1}(v_*, t^*) < \Delta_n^2 A^j(u_j, t^*) < \Delta_n^2 E^i(v_*, t^*)$. Then we obtain

$$
\begin{aligned}
\Delta_n^2 E^i(v_*, t^*) - \Delta_n^2 A^j(u_j, t^*) &\ge \Delta_n^2 E^i(v_*, t^*) - \Delta_n^2 A^j(u_{*,j} - \tfrac{1}{n}, t^*) \\
&\ge \eta = \Delta_n^2 E^i(v_*, t^*) - \Delta_n^2 E^{i+1}(v_*, t^*) \,,
\end{aligned}
$$

which is a contradiction to the assumption. The first inequality follows from the definition of u_{*j} using that $\Delta_n^2 A^j(\,\cdot\,, t^*)$ is increasing, and the second inequality follows from the choice of η .
Statement (42) can be shown analogously.

For u_j^* defined in (31) holds $u_j^* \le 1$ because otherwise, if $u_j = 1$ for a j , we would have that

$$
1 \ge \Delta_n^2 E^{i+1}(v_*, t_*) = \Delta_n^2 E^i(v_*, t_*) + \eta \ge \Delta_n^2 A^j(1, t_*) + \eta = 1 + \eta \,,
$$

which is a contradiction. Since the set $\mathcal{S}^i$ is non-empty we obtain from (40) that $u_* < u^*$.

Next, we observe that

$$
F^{i+1}(u) - F^i(u) = \begin{cases} -\frac{\eta}{n} & \text{if } u \in \mathcal{S}^i \,, \\ 0 & \text{else} \end{cases}
$$

holds by construction of E^{i+1} . Hence, $P_{F^{i+1}}$ is obtained from P_{F^i} by a finite number of reverse Δ-antitone transfers, see Müller (2013, Theorem 2.5.7). We show that these transfers can, in particular, be expressed by a reverse supermodular transfer indicated by

$$
\frac{\eta}{n}\left(\delta_{(u_*, v_*)} + \delta_{(u^*, v^*)}\right) \to \frac{\eta}{n}\left(\delta_{(u_*, v^*)} + \delta_{(u^*, v_*)}\right) \,.
$$

Define $\mu := \frac{\eta}{n}(\delta_{(u_*, v_*)} + \delta_{(u^*, v^*)} - (\delta_{(u_*, v^*)} + \delta_{(u^*, v_*)}))$. Then, we need to show that $\mu([0, x]) = -\frac{\eta}{n}$ for all $x \in \mathcal{S}$ and $\mu([0, x]) = 0$ for all $x \in (\mathcal{S}^i)^c \cap \mathbb{G}_n^{d+1}$.

Let $y \in \mathcal{S}^i$. Then, $y_{d+1} = v_*$. Further, from (40) we see that $y_j \ge u_{j,*}$ for all $1 \le j \le d$, and there exists an $j' \in \{1, \ldots, d\}$ such that $y_{j'} < u_{j'}^*$. Hence, we calculate $\mu([0, y]) = -\frac{\eta}{n}\delta_{(u_*, v_*)}([0, y]) = -\frac{\eta}{n}$.

Now, assume that $y_{d+1} = v_*$ but $y \notin \mathcal{S}^i$. If $y_j \geq u_j^*$ for all $1 \leq j \leq d$, then $\mu([0,y]) = \frac{\eta}{n}\left(\delta_{(u^*,v_*)}([0,y]) - \delta_{(u_*,v_*)}([0,y])\right) = 0$. If $y_{j'} < u_{*,j'}$ for a $j' \in \{1,\dots,d\}$, we obtain $\mu([0,y]) = 0$.

Suppose that $y_{d+1} \neq v_*$. If $y_{d+1} < v_*$ it immediately holds $\mu([0,y]) = 0$. If $y_{d+1} > v_*$, then $y_{d+1} \geq v^*$. But this also yields $\mu([0,y]) = 0$ independent of y_j, $1 \leq j \leq d$.

Now, since μ is a reverse supermodular transfer, the statement follows from Proposition 3.8. ■

References

Kjersti Aas, Claudia Czado, Arnoldo Frigessi, and Henrik Bakken. Pair-copula constructions of multiple dependence. *Insur. Math. Econ.*, 44(2), 2009.

Jonathan Ansari and Ludger Rüschendorf. Ordering results for risk bounds and cost-efficient payoffs in partially specified risk factor models. *Methodology and Computing in Applied Probability*, 2016.

O. Barndorff-Nielsen. Hyperbolic distributions and distributions on hyperbolae. *Scand. J. Stat.*, 5, 1978.

Carole Bernard, Ludger Rüschendorf, Steven Vanduffel, and Ruodu Wang. Risk bounds for factor models. *Finance Stoch.*, 21(3), 2017.

Stamatis Cambanis, Steel Huang, and Gordon Simons. On the theory of elliptically contoured distributions. *J. Multivariate Anal.*, 11, 1981.

William F. Darsow, Bao Nguyen, and Elwood T. Olsen. Copulas and Markov processes. *Ill. J. Math.*, 36(4), 1992.

Stefano Demarta and Alexander J. McNeil. The t copula and related copulas. *Int. Stat. Rev.*, 73(1), 2005.

Fabrizio Durante and Juan Fernández Sánchez. On the approximation of copulas via shuffles of Min. *Stat. Probab. Lett.*, 82(10), 2012.

Fabrizio Durante, Erich Peter Klement, and José Juan Quesada-Molina. Copulas: compatibility and fréchet classes. *arXiv preprint arXiv:0711.2409*, 2007.

Anna Kolesárová, Radko Mesiar, and Carlo Sempi. Measure-preserving transformations, copulæ and compatibility. *Mediterr. J. Math.*, 5(3), 2008.

Piotr Mikusiński and Michael D. Taylor. Some approximations of n-copulas. *Metrika*, 72(3), 2010.

Piotr Mikusinski, Howard Sherwood, and Michael D. Taylor. Shuffles of Min. *Stochastica*, 13(1), 1992.

Alfred Müller. Duality theory and transfers for stochastic order relations. In *Stochastic orders in reliability and risk.* New York, NY: Springer, 2013.

Alfred Müller and Marco Scarsini. Some remarks on the supermodular order. *Journal of Multivariate Analysis*, 73(1), 2000.

Alfred Müller and Dietrich Stoyan. *Comparison Methods for Stochastic Models and Risks.* Chichester: Wiley, 2002.

Roger B. Nelsen. *An introduction to copulas. 2nd ed.* New York, NY: Springer, 2nd ed. edition, 2006.

Ludger Rüschendorf. Solution of a statistical optimization problem by rearrangement methods. *Metrika*, 30, 1983.

Ludger Rüschendorf. *Mathematical Risk Analysis.* Springer, 2013.

Moshe Shaked and J. George Shanthikumar. *Stochastic Orders.* Springer, 2007.

Andre H. Tchen. Inequalities for distributions with given marginals. *Ann. Probab.*, 8, 1980.

Wolfgang Trutschnig. On a strong metric on the space of copulas and its induced dependence measure. *J. Math. Anal. Appl.*, 384(2), 2011.