Some Results on Multivariate Dependence Modeling

by

Yunran Wei

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Statistics

Waterloo, Ontario, Canada, 2015

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The goal of this thesis is to solve some problems in dependence modeling. Under special assumptions, we use Tankov [2011]'s result to give sharp bounds on variance of the sum of two random variables with partial information available and point out some drawbacks in his method. Thus, two different methods based on convex ordering are proposed. We show the one inspired by Bernard and Vanduffel [2014] may fail and provide an improved method. This thesis then discusses the compatible matrix problem. We characterize the covariance matrix for sums of normal distributed random variables to reach the minimum variance in dimensions three and four. This result is supported with application on variance bounds with background risk. The last part reviews some existing dependence measures and a new multivariate dependence measure focusing on the sum of random variables is introduced with properties and estimation method.

Each chapter ends with a conclusion and future research directions.

Acknowledgements

I would like to express my deep gratitude to my advisor, Professor Carole Bernard, for spending lots of time and energy on supervising me writing this thesis. Without her patient guidance and insightful advices, this thesis would not be completed or even started. She gave me a really good problem to work on. Thank Professor Ruodu Wang for reading this thesis, giving me lots of advice on my professional career and his help in Chapter 3. Thank Professor Marius Hofert for taking his precious time of reading this thesis and giving lots of constructive suggestions. Thank Professor Don McLeish for his help in Chapter 4.

Thank Ms. Mary Lou Dufton for her coordination and the Department of Statistics and Actuarial Science for the financial support.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation and Recent Works

Given a probability space $(\Omega, \mathcal{F}, P)$, a random vector $\mathbf{X} = (X_1, ..., X_d)$ is a mapping from Ω to $\mathbb{R}^d$ measurable with respect to $\mathcal{F}$ (see Billingsley [2008]). The dependence structure of $\mathbf{X}$ refers to the joint distribution function,

$$F(x_1, ..., x_d) := P(X_1 \leq x_1, ..., X_d \leq x_d)$$

for any $(x_1, ..., x_d) \in \mathbb{R}^d$.

This problem has been studied for centuries under different contexts, see Fréchet [1956], Hoeffding [1940], Hardy et al. [1952] and Tchen [1980] for early history.

Modeling multivariate dependence is important because data is multivariate (see Joe [1997] for problems in multivariate data modeling). Given marginal distributions, a function *copula* is introduced to model the multivariate dependence structure. Early references include Fréchet [1956], Hoeffding [1940], Kimeldori and Sampson [1975], see Schweizer [1991] and Dall'Aglio [1991] for a short history of copula, Embrechts [2009] for recent developments and potential future work.

Copula has applications in many practical problems, including engineering, biology, weather forecasting, Markov processes (see Darsow et al. [1992]), risk management (see

Embrechts et al. [2003b], Embrechts et al. [2002]) and finance (see Genest et al. [2009]). Its use in actuarial science is introduced by Carriere and Chan [1986].

One central problem of dependence modeling is to investigate the bounds on distribution of the sum $S = X_1 + ... + X_d$. Earliest references date back to 1981 by Makarov [1982] and Frank et al. [1987].

When dimension $d = 2$, this problem can be studied via the use of copula. The situation becomes more complex with the existence of partial information available on the copula. Tankov [2011] introduces the improved Fréchet bounds which is sharper than the classical Fréchet-Hoeffding bounds. This is a generalization of Nelsen [2007]'s result. Nelsen [2007] and Sadooghi-Alvandi et al. [2013] give the best-possible bounds when knowing the value of the copula at a given point and at several points separately; Nelsen et al. [2001] and Nelsen et al. [2004] give the improved bounds when a measure of association in terms of the copula and the diagonal sections are given, also see Beliakov et al. [2007], Nelsen and Ubeda-Flores [2005], Nelsen and Úbeda-Flores [2012] for related works. See Rachev and Rüschendorf [1998] and Rachev and Rüschendorf [1994] for earlier results on improvement of Fréchet bounds.

The improved Fréchet bounds have many applications in quantitative risk management such as the portfolio selection problems (see Bernard et al. [2014a] and Bernard et al. [2014b] for details), optimal investment strategies (see Bernard and Vanduffel [2011]), uniform distribution theory (Hofer and Iacò [2014]), option pricing (see Tankov [2011], Rapuch and Roncalli [2001]).

However, the problem is not fully solved as the result in Tankov [2011] does not give a sharp bounds on copulas in some cases. Bernard et al. [2012] extends Tankov [2011]'s result by giving a weaker sufficient condition for the quasi-copulas to be copulas and Bernard et al. [2013a] further extends Tankov [2011]'s result.

For $d \geq 3$, copula is not that useful as the bounds on $S = \sum_{i=1}^{d} X_i$ also depend on marginals (see Bernard et al. [2014]). Convex ordering is introduced here. The upper bounds for a general d and the lower bound for $d = 2$ is given in Denuit et al. [1999]. Related work include Dhaene et al. [2002] and Wang and Wang [2011]. The sharp lower bound when $n \geq 3$ is studied in Bernard et al. [2014] and generalized in Jakobsons et al. [2014], also see Cheung and Lo [2013], Cheung and Lo [2014]. Some numerical methods to approximate bounds are studied in Bernard and Mcleish [2014], Puccetti and Rüschendorf [2012], Embrechts et al. [2013], Puccetti [2013], Embrechts et al. [2014a], Bernard and

Vanduffel [2014], see Aas and Puccetti [2014] for a real case study. With partial information available in different situations, improved bounds on the sum are studied in Cheung and Vanduffel [2013] and Cheung [2008].

In quantitative risk management, by viewing risk factors as random vectors, we can use convex ordering to study bounds on the sum S, which is described as risk aggregation (see an overview in Embrechts and Puccetti [2010b]). This is useful for financial institutions to give more accurate risk assessment of portfolio and can be helpful for industry regulators and risk managers. There are mainly three problems being studied under two situations, with no or partial information available:

1. Bounds for the sum of dependent risk, namely

$$m_+(s) = \inf\{P(S < s) : X_i \text{ has marginal } F_i, i = 1, ..., d\}$$

and

$$M_+(s) = \sup\{P(S < s) : X_i \text{ has marginal } F_i, i = 1, ..., d\}.$$

See Rüschendorf [1982] and Embrechts et al. [2003a] for a short history of this problem, Wang et al. [2013] for a review on the existing results, also Embrechts and Puccetti [2006], Bernard et al. [2014], Rüschendorf [1991], Embrechts and Puccetti [2010a], Embrechts et al. [2013] for study of this problem under partial information available.

2. Bounds on the industry benchmark, *Value-at-risk* (VaR) (see Jorion [2007]), which is closely related to problem 1. See Embrechts et al. [2014a] for a short history, see Embrechts et al. [2003a], Wang et al. [2013], Wang and Wang [2013], Jakobsons et al. [2014] for the context of no information available and Bernard et al. [2013b], Kaas et al. [2009], Bernard et al. [2014], Bernard and Vanduffel [2014], Denuit et al. [1999], Bignozzi and Tsanakas [2013], Werner [2002], Mesfioui and Quessy [2005] with partial information available. See Alexander and Sarabia [2012] Kerkhof et al. [2010] for the worst VaR scenario. The issue of asymptotic equivalence of worst VaR and worst TVaR is studied in Embrechts et al. [2014b], Wang [2014], Embrechts et al. [2014a], Bernard et al. [2014] and Puccetti et al. [2013].

3. Bounds on other convex risk measures or coherent risk measures (see Artzner et al. [1999], Delbaen [2002], Kusuoka [2001]) Related work see Cont [2006], Kaas et al. [2009], Valdez et al. [2009], Bignozzi and Tsanakas [2013], Jakobsons et al. [2014], Bäuerle and Müller [2006].

Of course this is not a complete list and the problem is not fully solved especially when partial information is available.

Other application of convex ordering include bounds on option pricing: see Albrecher et al. [2008], d'Aspremont and El Ghaoui [2006], Hobson et al. [2005], Rogers and Shi [1995], Vanmaele et al. [2006], Keller-Ressel and Griessler [2011], Curran [1994], Vanduffel et al. [2008], Dhaene et al. [2005].

1.2 Setting and Notation

This section gives some definitions and theorems, which are used throughout the thesis. All the random vectors $\mathbf{X}$ in this thesis are continuous and we only focus on the sum $S = X_1 + ... + X_d$ where each $(X_i)_{1 \leq i \leq d}$ has fixed and known marginal.

1.2.1 Definitions

We first recall some definitions and theorems that can be found in Nelsen [2007]. A d-variate distribution function with $U[0,1]$ margins is called *copula*, denoted as: C. It has the formal definition as follows.

Definition 1.2.1. $C : [0,1]^d \longrightarrow [0,1]$ is a *copula* if and only if
(i) $C(\mathbf{u}) = 0$ if $\min(\mathbf{u}) = 0$
(ii) C is d-increasing, i.e. for all $\mathbf{a} < \mathbf{b}$, the volume of C in a hypercube $[\mathbf{a}, \mathbf{b}]$ is non-negative.
(iii) C has uniform margins, i.e. $C(\mathbf{u}) = u_i$ if $\mathbf{u} = (1, .., 1, u_i, 1, ..., 1)$ for $i = 1, ..., d$.

The *volume* of C at $[\mathbf{a}, \mathbf{b}]$ is given by,

$$\triangle_{[\mathbf{a},\mathbf{b}]} C = \sum_{i_d=1}^{2} ... \sum_{i_1=1}^{2} (-1)^{\sum_{j=1}^{d} i_j} C(u_{1,i_1}, ..., u_{d,i_d}) \tag{1.1}$$

where

$$u_{j,i_j} = \begin{cases} a_j, & \text{if} \quad i_j = 1 \\ b_j, & \text{if} \quad i_j = 2. \end{cases} \tag{1.2}$$

We focus on the bivariate copula $C : [0,1]^2 \longrightarrow [0,1]$ in the following. Then the volume is just $C(a_1, a_2) + C(b_1, b_2) - C(a_1, b_2) - C(b_1, a_2)$.

Given random variables $X_1, X_2, ...X_d$ following some joint distribution function H, where each X_i has marginal distribution F_i, $i = 1, ..., d$, we can use copula C to model the dependence structure between $(X_i)_{1\leq i\leq d}$. This is a result of *Sklar's theorem* as follows (see page 18 from Nelsen [2007]).

Theorem 1.2.2. *(Sklar's theorem)*
(i) For any joint distribution function H, with margins F_i, $i = 1, ..., d$, there exists a copula C such that

$$H(\mathbf{x}) = C(F_1(x_1), ..., F_d(x_d)), \mathbf{x} \in \mathbb{R}^{\mathbf{d}}. \tag{1.3}$$

Such C is unique if $F_i, i = 1, ..., d$ are continuous.
(ii) Given a d-variate copula C, and univariate distribution function $F_i, i = 1, ..., d$, H defined in equation (1.3) is a distribution function with margins $F_i, i = 1, .., d$.

The following is the definition of generalized inverse of an distribution function F.

Definition 1.2.3. (Generalized inverse of F, see more details in appendix of Embrechts et al. [2005])

$$F^{-1}(y) = \inf\{x \in \mathbb{R} : F(x) \geq y\} \qquad y \in [0,1] \tag{1.4}$$

and $\inf \emptyset = \infty$.

Definition 1.2.4. (Independence copula)
Given $\mathbf{u} \in [0,1]^d$, C is called an independence copula (denoted as $\prod$) if the following holds,

$$C(\mathbf{u}) = \prod_{i=1}^{d} u_i := \prod(\mathbf{u}).$$

Then we recall some definitions on stochastic orders from Müller and Stoyan [2002].

Definition 1.2.5. (Concordance order)
$\mathbf{X} = (X_1, ..., X_d) \prec \mathbf{Y} = (Y_1, ..., Y_d)$ means $\mathbf{X}$ is smaller than $\mathbf{Y}$ in concordance order, if both

$$P(X_1 \leq t_1, X_2 \leq t_2, ..., X_d \leq t_d) \leq P(Y_1 \leq t_1, Y_2 \leq t_2, ..., Y_d \leq t_d) \tag{1.5}$$

and

$$P(X_1 > t_1, X_2 > t_2, ..., X_d > t_d) \leq P(Y_1 > t_1, Y_2 > t_2, ..., Y_d > t_d)$$

hold for all $(t_1, t_2, ..., t_d) \in \mathbb{R}^d$.

Definition 1.2.6. (Convex order)
$\mathbf{X}$ is less than $\mathbf{Y}$ in convex order (written $\mathbf{X} \leq_{cx} \mathbf{Y}$), if $Ef(\mathbf{X}) \leq Ef(\mathbf{Y})$ for all convex functions $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ such that the expectations exist.

1.2.2 Simulation of a Copula in Two Dimensions

We define the derivative of the copula with respect to u and denote it as $C_u(v)$.

Definition 1.2.7.

$$C_u(v) = \frac{\partial C(u, v)}{\partial u} = \lim_{\triangle u \to 0} \frac{C(u + \triangle u, v) - C(u, v)}{\triangle u} = P[V \leq v | U = u] \qquad (1.6)$$

From Theorem 2.27 of Nelsen [2007], $C_u(v)$ exists for almost all u. $C_u(v)$ is also called conditional copula, denoted as $C(v|u)$.

The following is a procedure to simulate the copula C, which is used in the examples when we want to plot the support set (u, v) on the 2-dimensional u-v graph:

- Step 1. Generate a variate U that is uniform on $(0, 1)$;
- Step 2. Generate a variate T that is uniform on $(0, 1)$ and independent of U;
- Step 3. For n sufficiently large, $k = \lfloor Un \rfloor$, let $u = \frac{k}{n}$, $v = \frac{j}{n}$ for $j = 1, ..., n - 1$. We compute C_u in an approximate form as,

$$C_u(j) = \frac{C(\frac{k+1}{n}, \frac{j}{n}) - C(\frac{k}{n}, \frac{j}{n})}{\frac{1}{n}} \qquad (1.7)$$

  where C_u is a vector and $C_u(j)$ is the jth component;
- Step 4. We find the generalized inverse $V = C^{-1}(T)$ as follows. First find the index j such that the following holds, then we denote this j as $\tilde{j}$

$$\tilde{j} = \min\{j \in 1, .., n - 1 : C_u(j) > T\} \qquad (1.8)$$

Second,

$$V = \frac{\tilde{j}}{n+1}; \tag{1.9}$$

- Step 5. Plot (U, V) on a graph to visualize the support of the copula C.
- We repeat step 1 to 5 for a large number of times.

This procedure of copula simulation is known as the conditional distribution method, (see Section 2.9 of Nelsen [2007]) and it can easily be extended to higher dimensions (see Mai and Scherer [2012]).

1.2.3 Comonotonicity and Countermonotonicity

We now introduce the classical Fréchet-Hoeffding bounds, these are general bounds for all the copulas C when there is no information at all about C.

Theorem 1.2.8. *(Fréchet-Hoeffding bounds)*
Suppose C is a bivariate copula, for any $(u, v) \in [0,1]^2$, we have

$$W(u, v) := \max\{0, u + v - 1\} \leq C(u, v) \leq \min\{u, v\} := M(u, v). \tag{1.10}$$

Both W and M are copulas.

If random variables have copula M or W, we say they are *comonotonic* or *countermonotonic*.

When dimension $d \geq 3$, for any $(u_1, ..., u_d) \in [0,1]^d$, comonotonicity can still be defined by the upper Fréchet-Hoeffding bound $M(u_1, ..., u_d) = \min\{u_1, ..., u_d\}$, which is a copula such that

$$C(u_1, ..., u_d) \leq M(u_1, ..., u_d).$$

However, $W(u_1, ..., u_d) = \max\{\sum_{i=1}^{d} u_i - d + 1, 0\}$ is not a copula anymore. See Lee and Ahn [2014] for multidimensional extension of countermonotonicity. One special case of the most negative dependence structure is that $(X_1, ..., X_d)$ being $completely - mixable$, the followings are some basic concepts which are used in the thesis. They are cited from Wang et al. [2013].

Definition 1.2.9. (Completely mixable and jointly mixable distributions)
1. A univariate distribution function F is $n-$*completely mixable* (n-CM) if there exist n identically distributed random variables $X_1, ..., X_n$ with the same distribution F such that

$$P(X_1 + ... + X_n = C) = 1 \tag{1.11}$$

for some $C \in \mathbb{R}$.
2. The univariate distribution functions $F_1, ..., F_n$ are *jointly mixable* (JM) if there exist n random variables $X_1, ..., X_n$ with distribution functions $F_1, ..., F_n$, respectively, such that Equation (1.11) holds for some $C \in \mathbb{R}$.

Theorem 1.2.10. *1. Suppose $F_1, ..., F_n$ are JM with finite variances $\sigma_1^2, ..., \sigma_n^2$. Then*

$$\max_{1 \le i \le n} \sigma_i \le \frac{1}{2} \sum_{i=1}^{n} \sigma_i \tag{1.12}$$

2. Suppose F_i is $N(\mu_i, \sigma_i^2)$ for $i = 1, ..., n$. Then $F_1, ..., F_n$ are JM if and only if Equation (1.12) *holds.*

See Wang and Wang [2011], Puccetti et al. [2012], Wang and Wang [2014], Wang [2014], Puccetti and Wang [2014] for some recent developments of complete mixability.

The following proposition is first derived by Müller [1997], see Dhaene et al. [2002], Kaas et al. [2009] for more details.

Theorem 1.2.11. *For random variables $(X_i)_{1 \le i \le n}$ with marginals $(F_i)_{1 \le i \le n}$, the sharp upper convex ordering bound is $F_1^{-1}(U) + ... + F_n^{-1}(U)$, called the comonotonic dependence scenario,*

$$X_1 + X_2 + ... + X_n \le_{cx} F_1^{-1}(U) + ... + F^{-1}(U)$$

and the sharp lower convex ordering bound is obtained when $n = 2$, called the counter-monotonic dependence scenario,

$$F_1^{-1}(U) + F_2^{-1}(1 - U) \le_{cx} X_1 + X_2$$

where $U \sim U[0, 1]$.

1.3 Thesis Contributions and Overview

The main contributions of this thesis are as follows:

In Chapter 2, we study bounds on variance of the sum of two random variables $X+Y$ with partial information available under some special conditions. This part is based on the work of Tankov [2011]. We propose another method inspired by Bernard and Vanduffel [2014] and an improved one to get the same bounds with Tankov [2011]'s method. In particular, the central problem is as follows.

(**A**) With partial information available on the dependence structure ((X,Y)'s copula C) of (X,Y) on area $\mathcal{F} \subset [0,1]^2$, Tankov [2011] gives sharp upper and lower bounds (denoted as A and B respectively) of C. Then we get sharp upper and lower bounds of $\mathrm{Var}(X+Y)$ as,

$$\mathrm{Var}(X^* + Y^*) \leq \mathrm{Var}(X+Y) \leq \mathrm{Var}(\tilde{X} + \tilde{Y})$$

where (X^*, Y^*) has copula B and $(\tilde{X}, \tilde{Y})$ has copula A. There are two drawbacks of Tankov [2011]'s method. First, these copula bounds A and B are only sharp under some special conditions and we do not know the necessary conditions for them to be sharp. Second, the computation to get $\mathrm{Var}(X^*+Y^*)$ and $\mathrm{Var}(\tilde{X}+\tilde{Y})$ can be very lengthy, which is shown in Chapter 2. Third, it is not straightforward to extend this method to dimension $n \geq 3$.

Thus we propose another method as follows.

(**B**) Since we already know the dependence structure of (X,Y) on area $\mathcal{F}$, we only need to get an upper and lower bound of $(X+Y)|(X,Y) \notin \mathcal{F}$. By Theorem 1.2.11, do the following convex order bounds give sharp bounds (the same result with method A) of $\mathrm{Var}(X+Y)$?

$$F^{-1}_{X|(X,Y)\notin\mathcal{F}}(U) + F^{-1}_{Y|(X,Y)\notin\mathcal{F}}(1-U) \leq_{cx} ((X+Y)|(X,Y) \notin \mathcal{F}) \leq_{cx} F^{-1}_{X|(X,Y)\notin\mathcal{F}}(U) + F^{-1}_{Y|(X,Y)\notin\mathcal{F}}(U),$$

where $U \sim U[0,1]$.

We show that this method may fail to give sharp bounds and propose an improved version based on the similar technique by splitting the area $[0,1]\backslash\mathcal{F}$. The difficulty of this problem is to choose the right way to split area $[0,1]\backslash\mathcal{F}$.

In Chapter 3, this thesis discusses the compatible covariance matrix problem for sums of normal distributed random variables $S = X_1 + ... + X_n$ to reach the minimum variance.

Under the assumption of complete mixability (see Definition 1.2.9), we characterize such matrix when $n = 3$ and 4. This result is supported with applications on variance bounds with background risk.

In Chapter 4, we give an overview on existing dependence measures. A new multivariate dependence measure focusing on the sum of random variables is then introduced with properties and estimation method.

We give conclusions and future research directions at the end of each chapter and a small summary of the whole thesis in Chapter 5.

Chapter 2

Bounds on Variance with Partial Information on Dependence

This chapter is organized as follows. In Section 2.1, we give some background knowledge, including some theorems and definitions. In Section 2.2, we give one special example of calculating upper and lower bounds of $\mathrm{Var}(X+Y)$ using improved Fréchet bounds from Tankov [2011] and two conjectures. In Section 2.3, we try a first attempt in getting bounds of $\mathrm{Var}(X+Y)$ using Bernard and Vanduffel [2014]. We try a second attempt in Section 2.4 and propose an improved bounds using convex order. Section 2.5 gives some conjectures and an example on how to get these improved bounds. Section 2.6 gives the conclusion and future research directions while Section 2.7 is an appendix, including some simulation details from Sections 2.3 - 2.5.

2.1 Background of Chapter 2

We start this chapter with some useful lemmas and a corollary.

Lemma 2.1.1. *If $U \geq 0, V \geq 0$ are two random variables with distribution functions F_U and F_V, then*

$$E[UV] = \int_0^{F_V^{-1}(1)} \int_0^{F_U^{-1}(1)} P(U \geq u, V \geq v) du dv \tag{2.1}$$

Proof.

$$\begin{aligned}
&\int_0^{F_V^{-1}(1)}\int_0^{F_U^{-1}(1)} P(U \geq u, V \geq v) dudv \\
=&\int_0^{F_V^{-1}(1)}\int_0^{F_U^{-1}(1)} E[\mathbb{I}_{(U\geq u, V\geq v)}] dudv \\
=&E[\int_0^{F_V^{-1}(1)}\int_0^{F_U^{-1}(1)} \mathbb{I}_{(U\geq u, V\geq v)} dudv] \text{ by Fubini theorem} \\
=&E[\int_0^{U}\int_0^{V} dudv] \qquad \text{since} \quad U \leq F_U^{-1}(1), V \leq F_V^{-1}(1) \\
=&E[UV].
\end{aligned} \tag{2.2}$$

□

Corollary 2.1.2. *(Corollary of Lemma 2.1.1) Given random vector (X, Y) having copula C, if we have point-wise upper and lower copula bounds A and B such that*

$$B(u, v) \leq C(u, v) \leq A(u, v)$$

for any $(u, v) \subset [0, 1]^2$. Then, the following inequality holds,

$$\text{Var}(X^* + Y^*) \leq \text{Var}(X + Y) \leq \text{Var}(\tilde{X} + \tilde{Y})$$

where $X^*, \tilde{X}, X$ and $Y^*, \tilde{Y}, Y$ have same marginals respectively, (X^*, Y^*) has copula B and $(\tilde{X}, \tilde{Y})$ has copula A.

Proof.

$$\begin{aligned}
\text{Var}(X + Y) &= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y) \\
&= \text{Var}(X) + \text{Var}(Y) + 2\left[E(XY) - E(X)E(Y)\right].
\end{aligned}$$

Let F_X, F_Y be the marginal distribution functions for X, Y, and we know $F_{X^*} = F_{\tilde{X}} = F_X, F_{Y^*} = F_{\tilde{Y}} = F_Y$. By Lemma 2.1.1

$$\begin{aligned} E[XY] &= \int_0^{F_X^{-1}(1)} \int_0^{F_Y^{-1}(1)} P(X \geq x, Y \geq y) dx dy \\ &= \int_0^{F_X^{-1}(1)} \int_0^{F_Y^{-1}(1)} (1 - P(X \leq x) - P(Y \leq y) + P(X \leq x, Y \leq y)) \, dx dy \\ &= \int_0^{F_X^{-1}(1)} \int_0^{F_Y^{-1}(1)} (1 - F_X(x) - F_Y(y) + C(F_X(x), F_Y(y))) \, dx dy. \end{aligned}$$

Since

$$B(F_{X^*}(x), F_{Y^*}(y)) = B(F_X(x), F_Y(y)) \leq C(F_X(x), F_Y(y)) \leq A(F_X(x), F_Y(y)) = A(F_{\tilde{X}}(x), F_{\tilde{Y}}(y)),$$

the inequality follows. □

Lemma 2.1.3. *For a copula C and $\forall x, y, u, v \in [0, 1]$, the following inequality holds,*

$$C(x, y) - (x - u)_+ - (y - v)_+ \leq C(u, v) \leq C(x, y) + (u - x)_+ + (v - y)_+ \tag{2.3}$$

Proof. See page 71 of Nelsen [2007] for a proof. □

We make use of the concept of copulas introduced in Section 1.2. When some partial information on dependence is available, Tankov [2011] finds the improved Fréchet bounds. To present his result, we first define quasi-copulas.

Definition 2.1.4. (Quasi-copula)
A two-dimensional quasi-copula is a function $Q : [0, 1]^2 \to [0, 1]$ with the following properties,
(i) Q satisfies the boundary conditions: $Q(0, u) = Q(u, 0) = 0$ and $Q(1, u) = Q(u, 1) = u$ for all $u \in [0, 1]$.
(ii) Q is increasing in each argument.
(iii) Q has the Lipschitz property: $|Q(u_2, v_2) - Q(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|$ for all $(u_1, v_1, u_2, v_2) \in [0, 1]^4$.

Theorem 2.1.5. *(Improved Fréchet bounds from Tankov [2011])*
Let S be a compact subset of $[0, 1]^2$, and let Q be a quasi-copula. Let

$$\mathcal{C}_S = \{C' \textit{ is a copula} \mid C'(a, b) = Q(a, b), \quad \forall (a, b) \in S\} \tag{2.4}$$

and

$$\mathcal{Q}_S = \{Q' \text{ is a quasi-copula} \mid Q'(a,b) = Q(a,b), \quad \forall (a,b) \in S\}. \tag{2.5}$$

Define

$$\begin{aligned} A^{S,Q}(u,v) &:= \min\{u, v, \min_{(a,b)\in S}\{Q(a,b) + (u-a)^+ + (v-b)^+\}\}, \\ B^{S,Q}(u,v) &:= \max\{0, u+v-1, \max_{(a,b)\in S}\{Q(a,b) - (a-u)^+ - (b-v)^+\}\}. \end{aligned} \tag{2.6}$$

Then
(i) $A^{S,Q}$ and $B^{S,Q}$ are quasi-copulas satisfying

$$B^{S,Q}(u,v) \leq Q'(u,v) \leq A^{S,Q}(u,v) \quad \forall (u,v) \in [0,1]^2 \tag{2.7}$$

for every $Q' \in \mathcal{Q}_S$ and

$$A^{S,Q}(a,b) = B^{S,Q}(a,b) = Q(a,b) \quad \forall (a,b) \in S. \tag{2.8}$$

This means that $A^{S,Q}$ and $B^{S,Q}$ are the best-possible bounds of the set $\mathcal{Q}_S$. Since a copula is also a quasi-copula, $\mathcal{C}_S \subset \mathcal{Q}_S$, then $A^{S,Q}$ and $B^{S,Q}$ are also bounds of the set $\mathcal{C}_S$, which is

$$B^{S,Q}(u,v) \leq C'(u,v) \leq A^{S,Q}(u,v) \quad \forall (u,v) \in [0,1]^2 \tag{2.9}$$

for every $C' \in \mathcal{C}_S$.
(ii) If the set S is increasing (i.e, for all $(a_1,b_1) \in S$ and $(a_2,b_2) \in S$, either $a_1 \leq a_2$ and $b_1 \leq b_2$ or $a_1 \geq a_2$ and $b_1 \geq b_2$.), then $B^{S,Q}$ is a copula; if the set S is decreasing (i.e, for all $(a_1,b_1) \in S$ and $(a_2,b_2) \in S$, either $a_1 \leq a_2$ and $b_1 \geq b_2$ or $a_1 \geq a_2$ and $b_1 \leq b_2$.), then $A^{S,Q}$ is a copula. This gives the best-possible bounds of $\mathcal{C}_S$.

The following theorem from Bernard et al. [2013a] is an extension of Theorem 2.1.5, which gives weaker sufficient conditions for $A^{S,Q}$ and $B^{S,Q}$ to be copulas.

Theorem 2.1.6. *If S is a compact set satisfying the following property:*

$$\forall (a_0,b_0) \in S, \forall (a_1,b_1) \in S, \quad (a_0,b_1) \in S, (a_1,b_0) \in S. \tag{2.10}$$

Furthermore, suppose Q is a quasi-copula such that $\forall (a_0,b_0),(a_1,b_1) \in S$ with $a_0 < a_1$,

$b_0 < b_1$, we have

$$Q(a_1, b_1) + Q(a_0, b_0) - Q(a_0, b_1) - Q(a_1, b_0) \geq 0, \tag{2.11}$$

then $A^{S,Q}$ and $B^{S,Q}$ are copulas. Note Equation (2.11) is automatically satisfied when Q is a copula.

2.2 Upper and Lower Bounds of $\mathrm{Var}(X+Y)$ Using Improved Fréchet Bounds from Tankov [2011]

2.2.1 When Knowing Dependence Structure on $[0, a] \times [0, b]$

Proposition 2.2.1. *Fix $a, b \in (0, 1)$ and assume random vector (X, Y) has $U[0, 1]$ marginals. Let $\mathcal{F}$ denote a subset $[0, a] \times [0, b]$ of $[0, 1]^2$, if we know (X, Y) has independence copula C when $(X, Y) = (F_X(X), F_Y(Y)) \in \mathcal{F}$. Then by Corollary 2.1.2, the lower bound of $\mathrm{Var}(X + Y)$ is:*

$$\mathrm{Var}(U+V) = \frac{a^3b^3}{3} - \frac{2a^3b^2}{3} + \frac{a^3b}{3} - \frac{2a^2b^3}{3} + \frac{3a^2b^2}{2} - a^2b + \frac{ab^3}{3} - ab^2 + ab,$$

where (U, V) has copula $B^{\mathcal{F},C}$ with marginals $U[0, 1]$. Note that this is the improved Fréchet lower bound as in Theorem 2.1.5.

Proof. From Equation (2.6), $B^{\mathcal{F},C}$ is defined as

$$B^{\mathcal{F},C}(u, v) = \max\{0, u + v - 1, \max_{(x,y)\in\mathcal{F}}\{C(x, y) - (x - u)^+ - (y - v)^+\}\}. \tag{2.12}$$

From Theorem 2.1.6, we know $B^{\mathcal{F},C}$ is a copula since C is a copula and area $\mathcal{F}$ meets the sufficient conditions. Since

$$\begin{aligned}\mathrm{Var}(U + V) &= \mathrm{Var}(U) + \mathrm{Var}(V) + 2\mathrm{Cov}(U, V) \\ &= \frac{1}{12} + \frac{1}{12} + 2[E(UV) - E(U)E(V)] \\ &= 2E(UV) - \frac{1}{3}.\end{aligned} \tag{2.13}$$

Since $F_U^{-1}(1) = 1, F_V^{-1}(1) = 1$, by Lemma 2.1.1,

$$
\begin{aligned}
E[UV] &= \int_0^1 \int_0^1 P(U \geq u, V \geq v) du dv \\
&= \int_0^1 \int_0^1 \left(1 - P(U \leq u) - P(V \leq v) + P(U \leq u, V \leq v)\right) du dv \\
&= \int_0^1 \int_0^1 \Big(1 - u - v + B^{\mathcal{F},C}(u,v)\Big) du dv \\
&= \int_0^1 \int_0^1 B^{\mathcal{F},C}(u,v) du dv \qquad (2.14) \\
&= \underbrace{\int_0^1 \int_0^1 uv \mathbb{I}_{(u,v)\in\mathcal{F}} du dv}_{①} + \underbrace{\int_0^1 \int_0^1 B^{\mathcal{F},C}(u,v) \mathbb{I}_{(u,v)\in\mathcal{U}_1} du dv}_{②} \\
&\quad + \underbrace{\int_0^1 \int_0^1 B^{\mathcal{F},C}(u,v) \mathbb{I}_{(u,v)\in\mathcal{U}_2} du dv}_{③} + \underbrace{\int_0^1 \int_0^1 B^{\mathcal{F},C}(u,v) \mathbb{I}_{(u,v)\in\mathcal{U}_3} du dv}_{④},
\end{aligned}
$$

where the ①, ②, ③, ④ parts are shown in Figure 2.1.

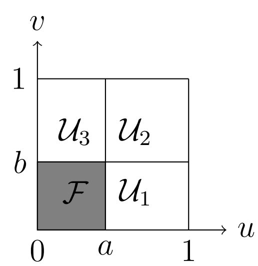


Figure 2.1: When $\mathcal{F} = [0, a] \times [0, b]$

We need to get the support set to simplify the expressions ① − ④.

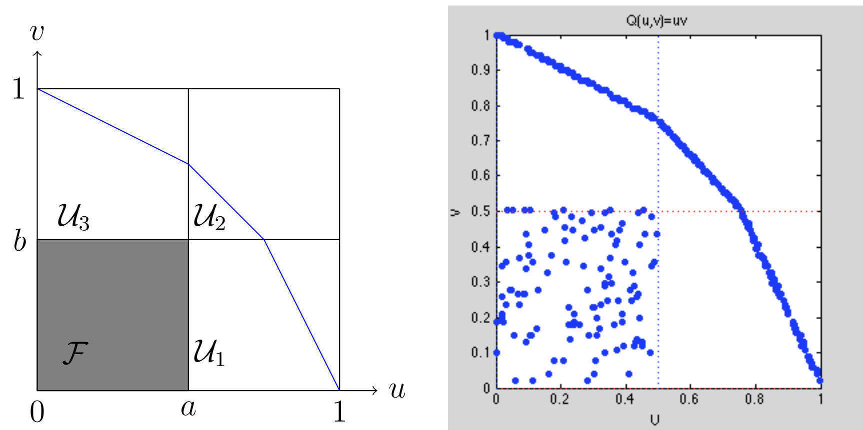
The followings are three cases to simplify parts ② − ④ respectively, treating C as a general copula. Denote $C(x, y) - (x - u)_+ - (y - v)_+$ by $H(x, y)$.

Case 1: $(u, v) \in \mathcal{U}_1$: When $(u, v) \in \mathcal{U}_1$, we have $a \leq u \leq 1, 0 \leq v \leq b$. Since $0 \leq x \leq a, 0 \leq y \leq b$, from Equation (2.3), we can see $H(x, y)$ can not achieve its maximum

$H(u,v)$, but as $H(x,y) = C(x,y) - (y-v)_+$ in this case, we can let $y = v$, pick $x = a$. So $\max_{0\le x\le a, 0\le y\le b} H(x,y) = C(a,v)$.
Thus we get $B^{\mathcal{F},C}(u,v) = \max\{0, u+v-1, C(a,v)\}$.
We further simplify $B^{\mathcal{F},C}(u,v)$ and focus on cases when $B^{\mathcal{F},C}(u,v)$ is not 0. Since if it is 0, then $B_u(v) = 0$, then $B_u^{-1}(0) = 0$. There is no support set here.
Case 1.1: $B^{\mathcal{F},C}(u,v) = u+v-1$
This is when $u+v-1 \ge \max\{0, C(a,v)\} = C(a,v)$. Then $B_u(v) = 1$.
Case 1.2: $B^{\mathcal{F},C} = C(a,v)$.
This is when $C(a,v) \ge \max\{0, u+v-1\}$, then $B_u(v) = 0$.

Case 2: $(u,v) \in \mathcal{U}_2$: When $(u,v) \in \mathcal{U}_2$ we have $a \le u \le 1, b \le v \le 1$. Since $0 \le x \le a, 0 \le y \le b$, then $(x-u)_+ = (y-v)_+ = 0$, so $H(x,y) = C(x,y)$. Thus $\max_{0\le x\le a, 0\le y\le b} H(x,y) = C(a,b)$ as C is increasing in x and y.
So $B^{\mathcal{F},C}(u,v) = \max\{0, u+v-1, C(a,b)\}$.
Case 2.1: $B^{\mathcal{F},C}(u,v) = u+v-1$
This is when $u+v-1 \ge \max\{0, C(a,b)\} = C(a,b)$, then $B_u(v) = 1$.
Case 2.2: $B^{\mathcal{F},C}(u,v) = C(a,b)$
This is when $C(a,b) \ge \max\{0, u+v-1\}$, then $B_u(v) = 0$.

Case 3: $(u,v) \in \mathcal{U}_3$: When $(u,v) \in \mathcal{U}_3$, we have $0 \le u \le a, b \le v \le 1$. Similar to case 1, we get $\max_{0\le x\le a, 0\le y\le b} H(x,y) = C(u,b)$.
So $B^{\mathcal{F},C}(u,v) = \max\{0, u+v-1, C(u,b)\}$.
Case 3.1: $B^{\mathcal{F},C}(u,v) = u+v-1$
This is when $u+v-1 \ge \max\{0, C(u,b)\}$, then $B_u(v) = 1$.
Case 3.2: $B^{\mathcal{F},C}(u,v) = C(u,b)$
This is when $C(u,b) \ge \max\{0, u+v-1\}$, then $B_u(v) = \frac{\partial C(u,b)}{\partial u}$.
Now we take $C(u,v) = uv$ as the independence copula. The support set is plotted in Panel A of Figure 2.2, which is the same as the simulation result in Panel B of Figure 2.2. The simulation result is generated by a MATLAB program following the simulation procedure stated in Section 1.2.2 where a=b=0.5 .
The detail in calculating the support set is as follows.
For $B^{\mathcal{F},C}$, when $(u,v) \in \mathcal{F}, B^{\mathcal{F},C}(u,v) = C(u,v) = uv$.
Case 1: $(u,v) \in \mathcal{U}_1$:



Panel A Panel B

Figure 2.2: Panel A: Plot of the support set. Panel B: Simulation of the support set when $a = b = 0.5$.

Follows from the general case in the beginning, we have

$$\begin{aligned} B^{\mathcal{F},C}(u,v)\mathbb{I}_{(u,v)\in\mathcal{U}_1} &= \begin{cases} av & \text{if} \quad u+v-1 \leq C(a,v) = av \\ \max\{0, u+v-1\} & \text{if} \quad C(a,v) = av \leq \max\{0, u+v-1\} \end{cases} \\ \Longrightarrow B_u(v) &= \begin{cases} 1 & \text{if} \quad u+v-1 \geq C(a,v) = av \\ 0 & \text{if} \quad C(a,v) = av \geq \max\{0, u+v-1\}. \end{cases} \end{aligned} \tag{2.15}$$

This is just

$$B_u(v) = \begin{cases} 1 & \text{if} \quad v \geq \frac{1-u}{1-a} \\ 0 & \text{if} \quad v \leq \frac{1}{a}\max\{0, u+v-1\}. \end{cases} \tag{2.16}$$

So we get

$$v = B_u^{-1}(t) = \begin{cases} \frac{1-u}{1-a} & \text{if} \quad t > 0 \\ 0 & \text{if} \quad t = 0. \end{cases} \tag{2.17}$$

So the support set lies on $v = \frac{1-u}{1-a}$. When $(u,v) \in \mathcal{U}_1$, we have $a \leq u \leq 1, 0 \leq v \leq b$. To keep $0 \leq v \leq b$, we need $0 \leq \frac{1-u}{1-a} \leq b \Longrightarrow 1-b+ab \leq u \leq 1$. Since $(1-b+ab)-a =$

$(1-a)(1-b) \geq 0$, we get

$$\begin{cases} 1-b+ab \leq u \leq 1 \\ a \leq u \leq 1 \end{cases} \implies 1-b+ab \leq u \leq 1. \tag{2.18}$$

So the support set here is $\Gamma_{\mathcal{U}_1} := \{1-b+ab \leq u \leq 1, 0 \leq v \leq b | v = \frac{1-u}{1-a}\}$ and $B^{\mathcal{F},C}$ on $\mathcal{U}_1$ is shown in Figure 2.3.

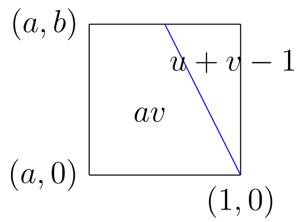


Figure 2.3: Case 1: $(u, v) \in \mathcal{U}_1$

So

$$B^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_1}(u,v) = \begin{cases} av & \text{if} \quad a \leq u \leq 1-b+ab, 0 \leq v \leq b \\ av & \text{if} \quad 1-b+ab \leq u \leq 1, 0 \leq v \leq \frac{1-u}{1-a} \\ u+v-1 & \text{if} \quad 1-b+ab \leq u \leq 1, \frac{1-u}{1-a} \leq v \leq b. \end{cases} \tag{2.19}$$

Case 2: $(u, v) \in \mathcal{U}_2$:
Since

$$\begin{aligned} B^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_2} &= \begin{cases} u+v-1 & \text{if} \quad u+v-1 \geq C(a,b) = ab \\ C(a,b) & \text{if} \quad C(a,b) = ab \geq \max\{0, u+v-1\} \end{cases} \\ \implies B_{u,3}(v) &= \begin{cases} 1 & \text{if} \quad u+v-1 \geq C(a,b) = ab \\ 0 & \text{if} \quad C(a,b) = ab \geq \max\{0, u+v-1\}. \end{cases} \end{aligned} \tag{2.20}$$

Then

$$B_u(v) = \begin{cases} 1 & \text{if} \quad v \geq ab + 1 - u \\ 0 & \text{if} \quad C(a,b) = ab \geq \max\{0, u+v-1\}. \end{cases} \tag{2.21}$$

Similar to case 1,

$$v = B_u^{-1}(t) = \begin{cases} ab + 1 - u & \text{if} \quad t > 0 \\ 0 & \text{if} \quad t = 0. \end{cases} \tag{2.22}$$

So the support set lies on $v = ab + 1 - u$. When $(u, v) \in \mathcal{U}_2$, we have $a \leq u \leq 1, b \leq v \leq 1$. To keep $b \leq v \leq 1$, we need $b \leq ab + 1 - u \leq 1 \implies ab \leq u \leq ab + 1 - b$. Since $ab + 1 - b = 1 + \underbrace{b(a-1)}_{\leq 0} \leq 1$ and $ab \leq a$, then

$$\begin{cases} ab \leq u \leq ab + 1 - b \\ a \leq u \leq 1 \end{cases} \implies a \leq u \leq ab + 1 - b. \tag{2.23}$$

So the support set here is $\Gamma_{\mathcal{U}_2} := \{a \leq u \leq ab + 1 - b, b \leq v \leq 1 | v = ab + 1 - u\}$, and $B^{\mathcal{F},C}$ on $\mathcal{U}_2$ is shown in Figure 2.4.

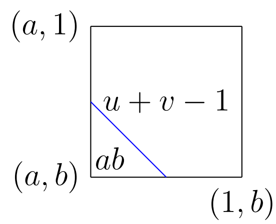


Figure 2.4: Case 2: $(u, v) \in \mathcal{U}_2$

So

$$B^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_2}(u, v) = \begin{cases} u + v - 1 & \text{if} \quad a \leq u \leq 1, ab + 1 - a \leq v \leq 1 \\ u + v - 1 & \text{if} \quad ab + 1 - v \leq u \leq 1, b \leq v \leq ab + 1 - a \\ ab & \text{if} \quad a \leq u \leq ab + 1 - v, b \leq v \leq ab + 1 - a. \end{cases} \tag{2.24}$$

<u>Case 3: $(u, v) \in \mathcal{U}_3$:</u>

Since

$$
\begin{aligned}
B^{\mathcal{F},C}(u,v) &= \begin{cases} u+v-1 & \text{if} \quad u+v-1 \geq \max\{0, C(u,b)\} = ub \\ C(u,b) & \text{if} \quad C(u,b) = ub \geq \max\{0, u+v-1\} \end{cases} \\
\Longrightarrow B^{\mathcal{F},C}(u,v) &= \begin{cases} u+v-1 & \text{if} \quad u+v-1 \geq \max\{0, C(u,b)\} = ub \\ C(u,b) = ub & \text{if} \quad C(u,b) = ub \geq \max\{0, u+v-1\}. \end{cases}
\end{aligned} \tag{2.25}
$$

Then

$$
B_u(v) = \begin{cases} 1 & \text{if} \quad v \geq 1 + ub - u \\ b & \text{if} \quad ub \geq \max\{0, u+v-1\}. \end{cases} \tag{2.26}
$$

Since $ub \geq \max\{0, u+v-1\}$ implies $ub \geq u+v-1$ and $v \geq 1-u$; or $ub \geq 0$ and $v < 1-u$. This is equivalent to $ub - u + 1 \geq v \geq 1-u$ or $v < 1-u$, which is equivalent to $1 + ub - u \geq v$.

So

$$
B_u(v) = \begin{cases} 1 & \text{if} \quad v \geq 1 + ub - u \\ b & \text{if} \quad v \leq 1 + ub - u. \end{cases} \tag{2.27}
$$

Then

$$
v = B_u^{-1}(t) = \begin{cases} 0 & \text{if} \quad t < b \\ 1 + ub - u & \text{if} \quad t \geq b. \end{cases} \tag{2.28}
$$

So the support set lies on $v = 1 + ub - u$. When $(u,v) \in \mathcal{U}_3$, we have $0 \leq u \leq a, b \leq v \leq 1$. To keep $b \leq v \leq 1$, we need $b \leq 1 + ub - u \leq 1 \Longrightarrow 0 \leq u \leq 1$.

So the support set here is $\Gamma_{\mathcal{U}_3} := \{(u,v) \in \mathcal{U}_3 | v = 1 + ub - u\}$, and $B^{\mathcal{F},C_3}$ on $\mathcal{U}_3$ is shown in Figure 2.5.

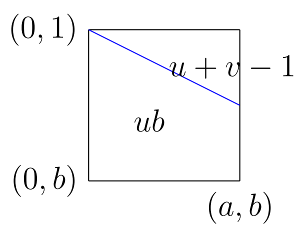


Figure 2.5: Case 3: $(u,v) \in \mathcal{U}_3$

So

$$B^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_3}(u,v)=\begin{cases}u+v-1 & \text{if} \quad 0\le u\le a, 1+ub-u\le v\le 1\\ ub & \text{if} \quad 0\le u\le a, 1+ab-a\le v\le 1+ub-u\\ ub & \text{if} \quad 0\le u\le a, b\le v\le ab+1-a.\end{cases} \tag{2.29}$$

In summary, $B^{\mathcal{F},C}=uv\mathbb{I}_{(u,v)\in\mathcal{F}}+B^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_1}+B^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_2}+B^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_3}$ and the last three terms are given in Equations (2.19), (2.24) and (2.29).

By Equation (2.14),

$$① = \int_0^b\int_0^a uv\,du\,dv=\frac{1}{4}a^2b^2. \tag{2.30}$$

By Equation (2.19),

$$\begin{aligned} ② &= \int_0^b\int_a^{1-b+ab} av\,du\,dv+\int_{1-b+ab}^1\int_0^{\frac{1-u}{1-a}} av\,dv\,du+\int_{1-b+ab}^1\int_{\frac{1-u}{1-a}}^b u+v-1\,dv\,du\\ &=\frac{a^2b^3}{2}-\frac{a^2b^2}{2}-\frac{ab^3}{2}+\frac{ab^2}{2}+\frac{ab^3}{6}-\frac{a^2b^3}{6}+\frac{b^3}{6}-\frac{a^2b^3}{6}\\ &=\frac{a^2b^3}{6}-\frac{a^2b^2}{2}-\frac{ab^3}{3}+\frac{ab^2}{2}+\frac{b^3}{6}.\end{aligned} \tag{2.31}$$

By Equation (2.24),

$$\begin{aligned} ③ &= \int_{ab+1-a}^1\int_a^1(u+v-1)\,du\,dv+\int_b^{ab+1-a}\int_{ab+1-v}^1(u+v-1)\,du\,dv\\ &\quad+\int_b^{ab+1-a}\int_a^{ab+1-v}(ab)\,du\,dv\\ &=\frac{a^3b^3}{6}-\frac{a^3b^2}{2}+\frac{a^3b}{2}-\frac{a^3}{6}-\frac{a^2b^3}{2}+\frac{3a^2b^2}{2}-a^2b+\frac{ab^3}{2}-ab^2+\frac{ab}{2}-\frac{b^3}{6}+\frac{1}{6}.\end{aligned} \tag{2.32}$$

By Equation (2.29),

$$\begin{aligned}
④ &= \int_0^a \int_{1+ub-u}^{1} (u+v-1)dvdu + \int_0^a \int_{1+ab-a}^{1+ub-u} (ub)dvdu \\
&+ \int_0^a \int_b^{ab+1-a} (ub)dvdu \\
&= \frac{a^3b^2}{6} - \frac{a^3b}{3} + \frac{a^3}{6} - \frac{a^2b^2}{2} + \frac{a^2b}{2}.
\end{aligned} \tag{2.33}$$

Then

$$\begin{aligned}
E(UV) &= ① + ② + ③ + ④ \\
&= \frac{a^3b^3}{6} - \frac{a^3b^2}{3} + \frac{a^3b}{6} - \frac{a^2b^3}{3} + \frac{3a^2b^2}{4} - \frac{a^2b}{2} + \frac{ab^3}{6} - \frac{ab^2}{2} + \frac{ab}{2} + \frac{1}{6},
\end{aligned} \tag{2.34}$$

so

$$\mathrm{Var}(U+V) = \frac{a^3b^3}{3} - \frac{2a^3b^2}{3} + \frac{a^3b}{3} - \frac{2a^2b^3}{3} + \frac{3a^2b^2}{2} - a^2b + \frac{ab^3}{3} - ab^2 + ab. \tag{2.35}$$

□

Example 2.2.2. To check whether our calculated variance is right, we compare Equation (2.35) with the simulation result. We simulate pairs of points (U, V) which has copula $B^{\mathcal{F},C}$ according to the simulation procedure in Section 1.2.2 and calculate the variance of $(U+V)$. The sample size is n here. Denote the simulated variance of $U + V$ as $\widehat{\mathrm{Var}(U+V)}$. We take several choices of a, b and look at the error as the absolute difference of the simulated $\widehat{\mathrm{Var}(U+V)}$ and the calculated $\mathrm{Var}(U + V)$ when n is some large number. The error is denoted as

$$\varepsilon_n = |\widehat{\mathrm{Var}(U+V)} - \mathrm{Var}(U+V)|. \tag{2.36}$$

Proposition 2.2.3. *Fix $a, b \in (0, 1)$ and assume random vector (X, Y) has $U[0, 1]$ marginals. Let $\mathcal{F}$ denote a subset $[0, a] \times [0, b]$ of $[0, 1]^2$, if we know (X, Y) has independence copula C when $(X, Y) = (F_X(X), F_Y(Y)) \in \mathcal{F}$. Then by Corollary 2.1.2, the upper bound of* $\mathrm{Var}(X + Y)$ *is:*

$$\mathrm{Var}(U+V) = \frac{2a^3b^3}{3} - \frac{4a^3b^2}{3} + \frac{2a^3b}{3} - \frac{4a^2b^3}{3} + \frac{5a^2b^2}{2} - a^2b + \frac{2ab^3}{3} - ab^2 + \frac{1}{3},$$

a	b	ε_{50}	ε_{500}	ε_{5000}
0.1	0.3	0.0085	0.0032	0.0029
0.4	0.6	0.0149	0.0031	0.0015
0.5	0.5	0.0137	0.0059	0.0001
0.9	0.63	0.0103	0.0078	0.0003

Table 2.1: Simulation error of Var$(U+V)$ when (U,V) has copula $B^{\mathcal{F},C}$

where (U,V) *has copula* $A^{\mathcal{F},C}$ *with marginals* $U[0,1]$*. Note that this is the improved Fréchet upper bound as in Theorem 2.1.5.*

Proof. From Equation (2.6), $A^{\mathcal{F},C}$ is defined as

$$A^{\mathcal{F},C}(u,v)=\min\{u,v,\min_{(x,y)\in\mathcal{F}}\{C(x,y)+(u-x)^{+}+(v-y)^{+}\}\}. \tag{2.37}$$

From Theorem 2.1.6, we know $A^{\mathcal{F},C}$ is a copula since C is a copula and area $\mathcal{F}$ meets the sufficient conditions. Similarly,

$$\mathrm{Var}(U+V)=2E(UV)-\frac{1}{3} \tag{2.38}$$

and

$$\begin{aligned}E[UV]=&\underbrace{\int_0^1\int_0^1 uv\mathbb{I}_{(u,v)\in\mathcal{F}}dudv}_{⑤}+\underbrace{\int_0^1\int_0^1 A^{\mathcal{F},C}(u,v)\mathbb{I}_{(u,v)\in\mathcal{U}_1}dudv}_{⑥}\\&+\underbrace{\int_0^1\int_0^1 A^{\mathcal{F},C}(u,v)\mathbb{I}_{(u,v)\in\mathcal{U}_2}dudv}_{⑦}+\underbrace{\int_0^1\int_0^1 A^{\mathcal{F},C}(u,v)\mathbb{I}_{(u,v)\in\mathcal{U}_3}dudv}_{⑧}.\end{aligned} \tag{2.39}$$

We need to get the support set to simplify the expressions ⑤ − ⑧.
Similar to the proof in Proposition 2.2.1, the followings are three cases to simplify parts ⑥ − ⑧ respectively, treating C as a general copula. Denote $C(x,y)-(x-u)_{+}-(y-v)_{+}$ by $G(x,y)$.

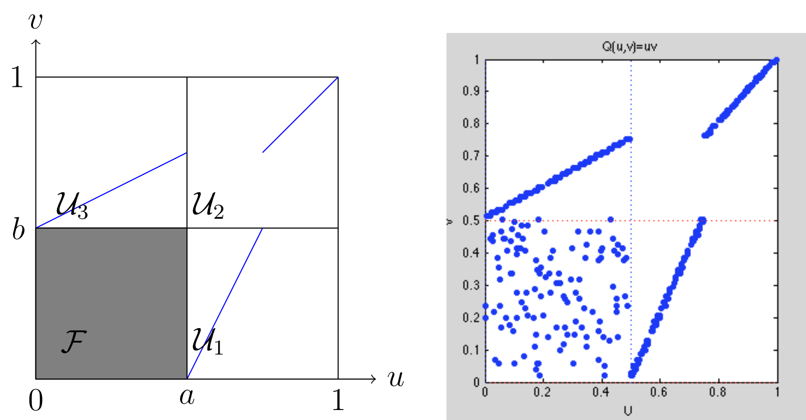
Case 1: $(u,v)\in\mathcal{U}_1$:

This is when $a \leq u \leq 1, 0 \leq v \leq b$. By Equation (2.3), $G(x, y)$ can not achieve its minimum at $(u, v) = (x, y)$ since $0 \leq x \leq a$. Fix $y = v$, $G(x, y) = C(x, v) + u - x$, its minimum depends on the choice of C.

Case 2: $(u, v) \in \mathcal{U}_2$:
This is when $a \leq u \leq 1, b \leq v \leq 1$. As $0 \leq x \leq a, 0 \leq y \leq b$, we have $G(x, y) = C(x, y) + u - x + v - y$. $\min_{0 \leq x \leq a, 0 \leq y \leq b} G(x, y)$ depends on the choice of C.

Case 3: $(u, v) \in \mathcal{U}_3$:
This is when $0 \leq u \leq a, b \leq v \leq 1$. Similar to case 1, $(x, y) = (u, v)$ is not achievable. Fix $x = u$, $G(x, y) = C(u, y) + v - y$.

Now we take $C(u, v) = uv$ as the independence copula. The support set is plotted in Panel A of Figure 2.6, which is the same as the simulation result in Panel B of Figure 2.6.



Panel A Panel B

Figure 2.6: Panel A: Plot of the support set. Panel B: Simulation of the support set when $a = b = 0.5$.

The detail in calculating the support set is as follows.
For $A^{\mathcal{F},C}$, when $(u, v) \in \mathcal{F}$, $A^{\mathcal{F},C}(u, v) = C(u, v) = uv$.

Case 1: $(u, v) \in \mathcal{U}_1$:
Follows from the general case, we get $A^{\mathcal{F},C}(u, v) = \min\{u, v, \min_{0\le x\le a}\{xv + u - x\}\} = \min\{u, v, av + u - a\} = \min\{v, av + u - a\}$ since $av + u - a \le u$.
So

$$\begin{aligned} A^{\mathcal{F},C}(u,v)\mathbb{I}_{(u,v)\in\mathcal{U}_1} &= \begin{cases} av+u-a & \text{if} \quad av+u-a \le v \\ v & \text{if} \quad v \le av+u-a \end{cases} \\ \Longrightarrow A_u(v) &= \begin{cases} 1 & \text{if} \quad av+u-a \le v \\ 0 & \text{if} \quad v \le av+u-a. \end{cases} \end{aligned} \tag{2.40}$$

Then

$$v = A_u^{-1}(t) = \begin{cases} \frac{u-a}{1-a} & \text{if} \quad t > 0 \\ 0 & \text{if} \quad t = 0. \end{cases} \tag{2.41}$$

When $(u, v) \in \mathcal{U}_1$, we have $0 \le v \le b$, so $0 \le \frac{u-a}{1-a} \le b \Longrightarrow a \le u \le a + b - ab$, since

$1-(a+b-ab)=(1-a)(1-b)\geq 0$, so

$$\begin{cases} a \leq u \leq 1 \\ a \leq u \leq a+b-ab \\ a+b-ab \leq 1 \end{cases} \Longrightarrow a \leq u \leq a+b-ab. \tag{2.42}$$

So the support set here is $\Gamma_{\mathcal{U}_1} := \{a \leq u \leq a+b-ab | v = \frac{u-a}{1-a}\}$ and $A^{\mathcal{F},C}$ on $\mathcal{U}_1$ is shown in Figure 2.7.

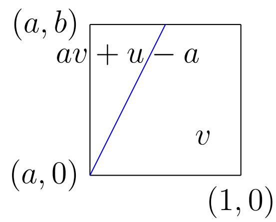


Figure 2.7: Case 1: $(u,v) \in \mathcal{U}_1$

So

$$A^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_1}(u,v) = \begin{cases} av+u-a & \text{if} \quad a \leq u \leq a+b-ab, \frac{u-a}{1-a} \leq v \leq b \\ v & \text{if} \quad a \leq u \leq a+b-ab, 0 \leq v \leq \frac{u-a}{1-a} \\ v & \text{if} \quad a+b-ab \leq u \leq 1, 0 \leq v \leq b. \end{cases} \tag{2.43}$$

<u>Case 2: $(u,v) \in \mathcal{U}_2$:</u>
Follows from the general case, we get

$$A^{\mathcal{F},C}(u,v) = \min\{u, v, \min_{0\leq x\leq a, 0\leq y\leq b}\{xy+u-x+v-y\}\} = \min\{u, v, ab+u-a+v-b\}. \tag{2.44}$$

Since

$$A^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_2} = \begin{cases} u & \text{if} \quad u \le \min\{v, ab+u-a+v-b\} \\ ab+u-a+v-b & \text{if} \quad ab+u-a+v-b \le \min\{u,v\} \\ v & \text{if} \quad v \le \min\{u, ab+u-a+v-b\} \end{cases}$$
$$\Longrightarrow A_u(v) = \begin{cases} 0 & \text{if} \quad v \le \min\{u, ab+u-a+v-b\} \\ 1 & \text{if} \quad v > \min\{u, ab+u-a+v-b\}. \end{cases} \tag{2.45}$$

which is

$$A_u(v) = \begin{cases} 1 & \text{if} \quad v \ge u, v \ge a+b-ab \quad \mathit{or} \quad v \le a+b-ab, u \le a+b-ab \\ 0 & \text{if} \quad \text{otherwise.} \end{cases} \tag{2.46}$$

Then

$$v = A_u^{-1}(t) = \begin{cases} u & \text{if} \quad t > 0, v \ge a+b-ab \\ 0 & \text{if} \quad \text{otherwise.} \end{cases} \tag{2.47}$$

So the support set here lies on $v = u$ when $v \ge a+b-ab$. When $(u,v) \in \mathcal{U}_2$, we have $b \le v \le 1, a \le u \le 1$. Since $a+b-ab \ge a$ and $a+b-ab \ge b$, then the support set is $\Gamma_{\mathcal{U}_2} := \{(u,v) \in \mathcal{U}_2 | v = u, v \ge a+b-ab\}$ and $A^{\mathcal{F},C}$ on $\mathcal{U}_2$ is shown in Figure 2.8.

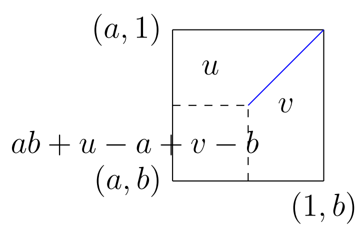


Figure 2.8: Case 2: $(u,v) \in \mathcal{U}_2$

So

$$A^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_2}(u,v)=\begin{cases}u & \text{if} \quad a\leq u\leq a+b-ab, a+b-ab\leq v\leq 1\\ u & \text{if} \quad a+b-ab\leq u\leq v, a+b-ab\leq v\leq 1\\ v & \text{if} \quad a+b-ab\leq u\leq 1, b\leq v\leq a+b-ab\\ v & \text{if} \quad v\leq u\leq 1, a+b-ab\leq v\leq 1\\ ab+u-a+v-b & \text{if} \quad a\leq u\leq a+b-ab, b\leq v\leq a+b-ab.\end{cases} \tag{2.48}$$

Case 3: $(u,v)\in\mathcal{U}_3$:

Follows from the general case, since $v\leq ub+v-b$ for $u\in[0,a]$ and $v\in[b,1]$, we get

$$A^{\mathcal{F},C}(u,v)=\min\{u,v,\min_{0\leq y\leq b}\{uy+v-y\}\}=\min\{u,v,ub+v-b\}=\min\{u,ub+v-b\}. \tag{2.49}$$

Since

$$\begin{aligned}A^{\mathcal{F},C}(u,v)&=\begin{cases}u & \text{if} \quad u\leq ub+v-b\\ ub+v-b & \text{if} \quad u>ub+v-b\end{cases}\\ \implies A_u(v)&=\begin{cases}1 & \text{if} \quad u\leq ub+v-b\\ b & \text{if} \quad u>ub+v-b,\end{cases}\end{aligned} \tag{2.50}$$

which is

$$A_u(v)=\begin{cases}1 & \text{if} \quad v\geq u-ub+b\\ b & \text{if} \quad v<u-ub+b.\end{cases} \tag{2.51}$$

Then

$$v=A_u^{-1}(t)=\begin{cases}u-ub+b & \text{if} \quad t>b\\ 0 & \text{if} \quad \text{otherwise.}\end{cases} \tag{2.52}$$

So the support set here is $\Gamma_{\mathcal{U}_3}:=\{(u,v)\in\mathcal{U}_3|v=u-ub+b\}$ and $A^{\mathcal{F},C}$ on $\mathcal{U}_3$ is shown in Figure 2.9.

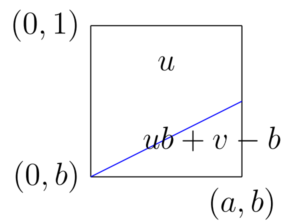


Figure 2.9: Case 3: $(u, v) \in \mathcal{U}_3$

So

$$A^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_3}(u,v) = \begin{cases} u & \text{if} \quad 0 \le u \le a, a+b-ab \le v \le 1 \\ u & \text{if} \quad 0 \le u \le \frac{v-b}{1-b}, b \le v \le a+b-ab \\ ub+v-b & \text{if} \quad \frac{v-b}{1-b} \le u \le a, b \le v \le a+b-ab. \end{cases} \tag{2.53}$$

In summary, $A^{\mathcal{F},C} = uv\mathbb{I}_{(u,v)\in\mathcal{F}} + A^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_1} + A^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_2} + A^{\mathcal{F},C}\mathbb{I}_{(u,v)\in\mathcal{U}_3}$ and the last three terms are given in Equations (2.43), (2.48) and (2.53).

By Equation (2.39),

$$⑤ = \int_0^b \int_0^a uvdudv = \frac{1}{4}a^2b^2. \tag{2.54}$$

By Equation (2.43),

$$\begin{aligned} ⑥ &= \int_a^{a+b-ab} \int_{\frac{u-a}{1-a}}^b (av+u-a)dvdu + \int_a^{a+b-ab} \int_0^{\frac{u-a}{1-a}} vdvdu + \int_{a+b-ab}^1 \int_0^b vdvdu \\ &= -\frac{1}{6}(a-1)(a+1)b^3 + \frac{b^3}{6} - \frac{ab^3}{6} + \frac{1}{2}b^2(ab-a-b+1) \\ &= -\frac{1}{6}a^2b^3 + \frac{ab^3}{3} - \frac{ab^2}{2} - \frac{b^3}{6} + \frac{b^2}{2}. \end{aligned} \tag{2.55}$$

By Equation (2.48),

$$
\begin{aligned}
⑦ &= \int_a^{a+b-ab}\int_{a+b-ab}^{1} udvdu + \int_{a+b-ab}^{1}\int_{a+b-ab}^{v} ududv \\
&+ \int_{a+b-ab}^{1}\int_{b}^{a+b-ab} vdvdu + \int_{a+b-ab}^{1}\int_{v}^{1} vdudv + \int_{a}^{a+b-ab}\int_{b}^{a+b-ab} (ab+u-a+v-b)dvdu \\
&= \frac{1}{2}(a-1)^2(b-1)b(a(b-2)-b) - \frac{1}{6}(a-1)^2(b-1)^2(2a(b-1)-2b-1) \\
&+ \frac{1}{2}(a-1)a(b-1)^2(a(b-1)-2b) - \frac{1}{6}(a-1)^2(b-1)^2(2a(b-1)-2b-1) \\
&+ \frac{1}{2}(a-1)a(b-1)b(a+b) \\
&= \frac{a^3b^3}{3} - \frac{a^3b^2}{2} + \frac{a^3}{6} - \frac{a^2b^3}{2} + a^2b^2 - \frac{a^2}{2} + \frac{b^3}{6} - \frac{b^2}{2} + \frac{1}{3}.
\end{aligned}
\tag{2.56}
$$

By Equation (2.53),

$$
\begin{aligned}
⑧ &= \int_0^a\int_{a+b-ab}^{1} udvdu + \int_b^{a+b-ab}\int_0^{\frac{v-b}{1-b}} ududv + \int_b^{a+b-ab}\int_{\frac{v-b}{1-b}}^{a} (ub+v-b)dudv \\
&= \frac{a^3b}{2} - \frac{a^3}{2} - \frac{a^2b}{2} + \frac{a^2}{2} + \frac{a^3}{6} - \frac{a^3b}{6} - \frac{1}{6}a^3(b-1)(b+1) \qquad (2.57) \\
&= -\frac{1}{6}a^3b^2 + \frac{a^3b}{3} - \frac{a^3}{6} - \frac{a^2b}{2} + \frac{a^2}{2}.
\end{aligned}
$$

Then

$$
\begin{aligned}
E(UV) &= ⑤ + ⑥ + ⑦ + ⑧ \\
&= \frac{a^3b^3}{3} - \frac{2a^3b^2}{3} + \frac{a^3b}{3} - \frac{2a^2b^3}{3} + \frac{5a^2b^2}{4} - \frac{a^2b}{2} + \frac{ab^3}{3} - \frac{ab^2}{2} + \frac{1}{3},
\end{aligned}
\tag{2.58}
$$

so

$$
\mathrm{Var}(X+Y) = \frac{2a^3b^3}{3} - \frac{4a^3b^2}{3} + \frac{2a^3b}{3} - \frac{4a^2b^3}{3} + \frac{5a^2b^2}{2} - a^2b + \frac{2ab^3}{3} - ab^2 + \frac{1}{3}. \tag{2.59}
$$

□

Example 2.2.4. To check this calculated variance of $(U+V)$, we fixed the same pairs of a, b as in Example 2.2.2 and compare the theoretical variance with the simulated variance in Table 2.2.

a	b	ε_{50}	ε_{500}	ε_{5000}
0.1	0.3	0.0222	0.0093	0.0074
0.4	0.6	0.0319	0.0072	0.0020
0.5	0.5	0.0240	0.0129	0.0011
0.9	0.63	0.0268	0.0030	0.0008

Table 2.2: Simulation error of $\mathrm{Var}(U+V)$ when (U,V) has copula $A^{\mathcal{F},C}$

Corollary 2.2.5. *(Variance bounds using improved Fréchet bounds)*
Fix $a,b\in(0,1)$ and assume random vector (X,Y) has $U[0,1]$ marginals. Let $\mathcal{F}$ denote a subset $[0,a]\times[0,b]$ of $[0,1]^2$, if we know (X,Y) has independence copula C when $(X,Y)=(F_X(X),F_Y(Y))\in\mathcal{F}$. Then sharp upper and lower bounds of $\mathrm{Var}(X+Y)$ are as follows,

$$\begin{aligned}&\frac{a^3b^3}{3}-\frac{2a^3b^2}{3}+\frac{a^3b}{3}-\frac{2a^2b^3}{3}+\frac{3a^2b^2}{2}-a^2b+\frac{ab^3}{3}-ab^2+ab\\ &\leq\mathrm{Var}(X+Y)\leq\\ &\frac{2a^3b^3}{3}-\frac{4a^3b^2}{3}+\frac{2a^3b}{3}-\frac{4a^2b^3}{3}+\frac{5a^2b^2}{2}-a^2b+\frac{2ab^3}{3}-ab^2+\frac{1}{3}.\end{aligned}$$

Proof. This directly follows from Propositions 2.2.1, 2.2.3 and Corollary 2.1.2. □

2.2.2 When Knowing Dependence Structure on $[0, a_1] \times [0, b_1]$ and $[a_2, 1] \times [b_2, 1]$

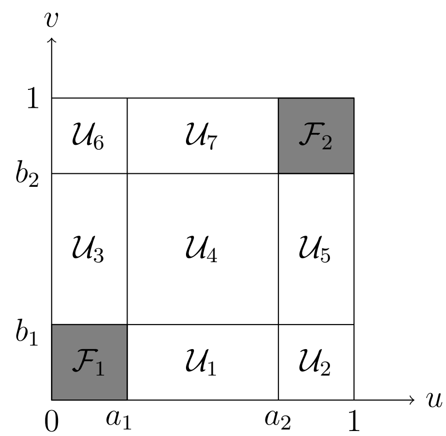
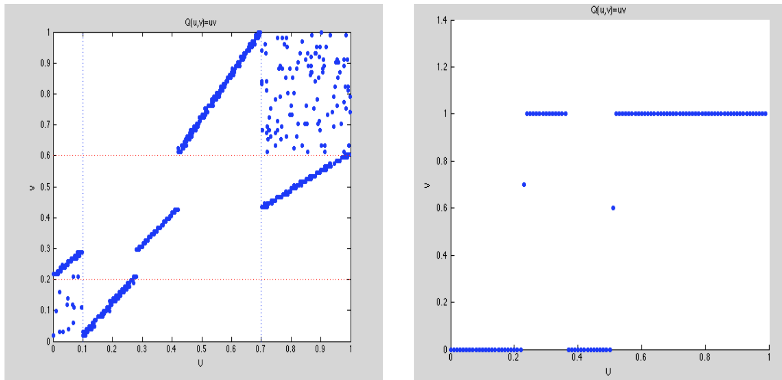


Figure 2.10: When $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$

Remark 2.2.6. Fix $a_1, b_1, a_2, b_2 \in (0, 1)$ where $a_1 < a_2, b_1 < b_2$. Let $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$ denote a subset $[0, a_1] \times [0, b_1] \cup [a_2, 1] \times [b_2, 1]$ of $[0, 1]^2$ (see Figure 2.10), let $C(u, v) = uv$ be an independence copula. From Equation (2.6) of Theorem 2.1.5, $A^{\mathcal{F},C}$ is defined as

$$A^{\mathcal{F},C}(u, v) = \min\{u, v, \min_{(x,y)\in\mathcal{F}}\{C(x, y) + (u - x)^+ + (v - y)^+\}\}. \tag{2.60}$$

Under this case, whether $A^{\mathcal{F},C}$ is a copula depends on the choice of a_1, b_1, a_2, b_2. For instance, when $a_1 = 0.1, b_1 = 0.2, a_2 = 0.7, b_2 = 0.6$, $A^{\mathcal{F},C}$ is a copula, see Panel A of Figure 2.11. When $a_1 = 0.1, b_1 = 0.3, a_2 = 0.6, b_2 = 0.5$, $A^{\mathcal{F},C}$ is not a copula, see Panel B of Figure 2.11, where $\frac{\partial C(u,v)}{\partial u}$ is decreasing when $v = 0.3$. (See Equation (1.6) for details.)



Panel A Panel B

Figure 2.11: Panel A: Simulation plot when $a_1 = 0.1, b_1 = 0.2, a_2 = 0.7, b_2 = 0.6$. Panel B: Simulation plot of $\frac{\partial C(u,v)}{\partial u}$ when $v = 0.3$ where $a_1 = 0.1, b_1 = 0.3, a_2 = 0.6, b_2 = 0.5$.

Following Remark 2.2.6, we give two conjectures which give weaker sufficient conditions for $A^{\mathcal{F},C}$ and $B^{\mathcal{F},C}$ to be copulas.

Conjecture 2.2.7. *Fix $a_1, b_1, a_2, b_2 \in (0,1)$ where $a_1 < a_2, b_1 < b_2$. Let $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$ denote a subset $[0,a_1] \times [0,b_1] \cup [a_2,1] \times [b_2,1]$ of $[0,1]^2$(see Figure 2.10), let $C(u,v)$ be a given copula. When*

$$C(a_2,b_2) + C(a_1,b_1) \geq a_1 + b_1$$[1],

then $A^{\mathcal{F},C}$ (see Equation 2.60) is a copula.

Conjecture 2.2.8. *Fix $a_1, b_1, a_2, b_2 \in (0,1)$ where $a_1 < a_2, b_1 < b_2$. Let $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$ denote a subset $[0,a_1] \times [b_2,1] \cup [a_2,1] \times [0,b_1]$ of $[0,1]^2$ (see Figure 2.12), let $C(u,v)$ be a given copula. When*

$$a_2 + b_2 \geq C(a_1,b_2) + C(a_2,b_1) + 1,$$

then $B^{\mathcal{F},C}$ (see Equation 2.12) is a copula.

[1]This actually comes from the condition to make Figure 2.18 a valid plot.

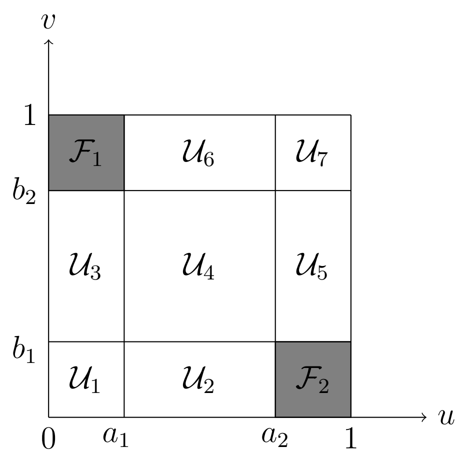


Figure 2.12: When $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$

2.3 Upper and Lower Bounds of $\mathrm{Var}(X+Y)$ Using Bernard and Vanduffel [2014]

The following proposition will be used in both Sections 2.3 and 2.4.

Proposition 2.3.1. *Given random variables (X,Y) having $U[0,1]$ marginals, let $\mathbb{I}_S$ be the indicator variable corresponding to the event $(X,Y) \in S$, assume that we know the copula of (X,Y) on area $\mathcal{F} \subset [0,1]^2$ and conditional distributions $F_{X|(X,Y)\in\mathcal{U}_i}, F_{Y|(X,Y)\in\mathcal{U}_i}, i = 1, ..., m$ where each $\mathcal{U}_i \subset [0,1]^2, \cup_{i=1}^m \mathcal{U}_i = [0,1]\backslash\mathcal{F}, \mathcal{U}_i \cap \mathcal{U}_j = \emptyset, i \neq j$. Then we have the following convex order bounds:*

$$\mathbb{I}_{\mathcal{F}}(X+Y) + \sum_{i=1}^{m} \mathbb{I}_{\mathcal{U}_i} \left(F^{-1}_{X|(X,Y)\in\mathcal{U}_i}(U) + F^{-1}_{Y|(X,Y)\in\mathcal{U}_i}(1-U) \right) \leq_{cx} X+Y$$
$$\leq_{cx} \mathbb{I}_{\mathcal{F}}(X+Y) + \sum_{i=1}^{m} \mathbb{I}_{\mathcal{U}_i} \left(F^{-1}_{X|(X,Y)\in\mathcal{U}_i}(U) + F^{-1}_{Y|(X,Y)\in\mathcal{U}_i}(U) \right), \quad (2.61)$$

where $U \sim U[0,1]$ is independent of all $\mathbb{I}_{\mathcal{F}}, \mathbb{I}_{\mathcal{U}_i}, i = 1, ..., m$.

Proof. Let $F := \{(X,Y) \in \mathcal{F}\}$, $U_i := \{(X,Y) \in \mathcal{U}_i\}$, $G_{2i-1} := F_{X|U_i}, G_{2i} := F_{Y|U_i}, i = 1, ..., m$.

For any convex function $v(x)$ and each $i = 1, ..., m$, by Theorem 1.2.11,

$$E(v(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(1-U))) \leq E(v(X+Y)|U_i) \leq E(v(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(U))). \quad (2.62)$$

Since U_i and U are independent, this is just,

$$E(v(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(1-U))|U_i) \leq E(v(X+Y)|U_i) \leq E(v(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(U))|U_i). \quad (2.63)$$

Then

$$E(v(X+Y)) = E(v(X+Y)|F)P(F) + \sum_{i=1}^{m} E(v(X+Y)|U_i)P(U_i). \quad (2.64)$$

Since we have

$$\begin{aligned}
&E(v(\mathbb{I}_{\mathcal{F}}(X+Y) + \sum_{i=1}^{m} \mathbb{I}_{\mathcal{U}_i}(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(1-U)))) \\
=&E(v(\mathbb{I}_{\mathcal{F}}(X+Y) + \sum_{i=1}^{m} \mathbb{I}_{\mathcal{U}_i}(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(1-U)))|F)P(F) \\
+&\sum_{i=1}^{m} E(v(\mathbb{I}_{\mathcal{F}}(X+Y) + \sum_{i=1}^{m} \mathbb{I}_{\mathcal{U}_i}(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(1-U)))|U_i)P(U_i) \\
=&E(v(X+Y)|F)P(F) + \sum_{i=1}^{m} E(v(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(1-U))|U_i)P(U_i),
\end{aligned}$$

and similarly,

$$\begin{aligned}
&E(v(\mathbb{I}_{\mathcal{F}}(X+Y) + \sum_{i=1}^{m} \mathbb{I}_{\mathcal{U}_i}(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(U)))) \\
=&E(v(X+Y)|F)P(F) + \sum_{i=1}^{m} E(v(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(U))|U_i)P(U_i).
\end{aligned}$$

Then by Equations (2.63) and (2.64),

$$\begin{aligned}
&E(v(\mathbb{I}_{\mathcal{F}}(X+Y) + \sum_{i=1}^{m} \mathbb{I}_{\mathcal{U}_i}(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(1-U)))) \leq E(v(X+Y)) \leq \\
&E(v(\mathbb{I}_{\mathcal{F}}(X+Y) + \sum_{i=1}^{m} \mathbb{I}_{\mathcal{U}_i}(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(U))))
\end{aligned}$$

So we get Equation (2.61). □

Remark 2.3.2. Note in Equation (2.62), the inequality is a result of Theorem 1.2.11, but not of Proposition 2 from Kaas et al. [2000].

The following proposition is inspired by Bernard and Vanduffel [2014].

Proposition 2.3.3. *Fix some* $0 < a < 1$ *and* $0 < b < 1$ *such that* $\mathcal{F} = [0, a] \times [0, b]$ *and* $\mathcal{U} = [0, 1]^2 \backslash \mathcal{F}$ *(See Figure 2.13).*

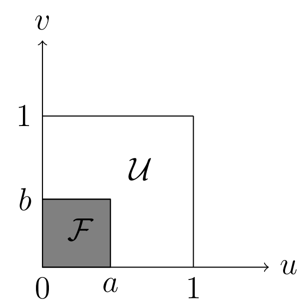


Figure 2.13: When $\mathcal{F} = [0, a] \times [0, b]$ and $\mathcal{U} = [0, 1]^2 \backslash \mathcal{F}$

We have the following assumptions:

1. X, Y *are two random variables follows* F_1 *and* F_2 *distributions where* $F_i, i = 1, 2$ *are* $U[0, 1]$ *distribution;*
2. $p_f := P((X, Y) \in \mathcal{F})$*, hence* $p_u := P((X, Y) \in \mathcal{U})$ *is also known as* $p_u = 1 - p_f$*;*
3. The conditional distribution $(X, Y)|(X, Y) \in \mathcal{F}$ *is known, denote*

$$H(x, y) := F_{(X,Y)|(X,Y)\in\mathcal{F}}(x, y) = P(X \le x, Y \le y|(X, Y) \in \mathcal{F}); \tag{2.65}$$

4. U *is a standard uniformly distributed random variable and is independent of the event* $(X, Y) \in \mathcal{F}$*;*
For some $x \in [0, 1]$*, denote*

$$\begin{aligned} G_1(x) &:= F_{X|(X,Y)\in\mathcal{U}}(x) := P(X \le x|(X, Y) \in \mathcal{U}); \\ G_2(x) &:= F_{Y|(X,Y)\in\mathcal{U}}(x) := P(Y \le x|(X, Y) \in \mathcal{U}). \end{aligned} \tag{2.66}$$

We assume 1,2,3,4, bounds on variance are

$$\mathrm{Var}(\mathbb{I}(X+Y)+(1-\mathbb{I})(G_1^{-1}(U)+G_2^{-1}(1-U))) \leq \mathrm{Var}(X+Y) \leq \mathrm{Var}(\mathbb{I}(X+Y)+(1-\mathbb{I})(G_1^{-1}(U)+G_2^{-1}(U))) \tag{2.67}$$

where $\mathbb{I}$ *is the indicator variable corresponding to event* "$(X, Y) \in \mathcal{F}$" *as*

$$\mathbb{I} := \mathbb{I}_{(X,Y)\in\mathcal{F}}$$

and $G_1, G_2, G_1^{-1}, G_2^{-1}$ *are computed in Lemma 2.3.8.*

Remark 2.3.4. If we change the marginal distributions F_1, F_2 of X and Y, we can still get bounds on variance following the similar technique in Proposition 2.3.3.

Remark 2.3.5. Though G_1 and G_2 (see Equation (2.66)) are defined as distributions conditioned on $(X, Y) \in \mathcal{U}$, G_1 and G_2 do not depend on the copula between X and Y, where $(X, Y) \in \mathcal{U}$ (see Lemma 2.3.8).

Before proving Proposition 2.3.3, we use the following proposition to show that the assumptions are equivalent to the assumptions imposed on Proposition 2.2.5 in Section 2.2.

Proposition 2.3.6. *Given Figure 2.13, under assumptions 1,2 of Proposition 2.3.3, fixing the conditional distribution of* $(X, Y)|(X, Y) \in \mathcal{F}$ *(denoted as* $H(x, y)$*, see Equation* (2.65)*) is equivalent to fixing the copula* C *of* $(X, Y) \in \mathcal{F}$*.*

Proof. " $\Rightarrow$ " When $(x, y) \in \mathcal{F}$, $0 \leq x \leq a, 0 \leq y \leq b$, then

$$H(x, y) = \frac{P(X \leq x, Y \leq y, X \leq a, Y \leq b)}{P((X, Y) \in \mathcal{F})} = \frac{P(X \leq x, Y \leq y)}{P((X, Y) \in \mathcal{F})} = \frac{C(x, y)}{p_f}. \tag{2.68}$$

So $C(x, y) = p_f H(x, y)$, $(x, y) \in \mathcal{F}$.

" $\Leftarrow$ " If X, Y has fixed copula C when $(X,Y) \in \mathcal{F}$, by Equation (2.68),

$$H(x,y) = \begin{cases} \dfrac{C(x,y)}{p_f} & \text{if } x \le a, y \le b \\ \dfrac{C(x,b)}{p_f} & \text{if } x \le a, y > b \\ \dfrac{C(a,y)}{p_f} & \text{if } x > a, y \le b \\ \dfrac{C(a,b)}{p_f} & \text{if } x > a, y > b, \end{cases} \tag{2.69}$$

which means H is fixed. □

Corollary 2.3.7. *Given Figure 2.13, assumptions 1,2,3,4 in Proposition 2.3.3 are equivalent to assumptions in Proposition 2.2.5 on producing the bounds on variance.*

Proof. This follows directly by using Proposition 2.3.6. □

Now we prove Proposition 2.3.3.

Proof. (Proof of Proposition 2.3.3) Let $F := \{(X,Y) \in \mathcal{F}\}, F^c := \{(X,Y) \in \mathcal{U}\}$. By Proposition 2.3.1, take $m = 1$, $v(x) = x^2$, we have

$$E((\mathbb{I}(X+Y) + (1-\mathbb{I})(G_1^{-1}(U) + G_2^{-1}(1-U)))^2) \le E((X+Y)^2) \le$$
$$E((\mathbb{I}(X+Y) + (1-\mathbb{I})(G_1^{-1}(U) + G_2^{-1}(U)))^2).$$

Since $G_1^{-1}(U) \sim G_1 = F_{X|F^c}, G_2^{-1}(U), G_2^{-1}(1-U) \sim G_2 = F_{Y|F^c}$,

$$E(X|F^c) = E(G_1^{-1}(U)), E(Y|F^c) = E(G_2^{-1}(U)) = E(G_2^{-1}(1-U)).$$

So

$$\begin{aligned} E(X+Y) &= E(X+Y|F)p_f + E(X+Y|F^c)(1-p_f) \\ &= E(X+Y|F)p_f + E(G_1^{-1}(U) + G_1^{-1}(U)|F^c)(1-p_f) \\ &= E(X+Y|F)p_f + E(G_1^{-1}(U) + G_1^{-1}(1-U)|F^c)(1-p_f) \\ &= E(\mathbb{I}(X+Y) + (1-\mathbb{I})(G_1^{-1}(U) + G_2^{-1}(U))) \\ &= E(\mathbb{I}(X+Y) + (1-\mathbb{I})(G_1^{-1}(U) + G_2^{-1}(1-U))). \end{aligned}$$

By Equation (2.3), we get inequality (2.67). □

Now we calculate the variance bounds $\mathrm{Var}(\mathbb{I}(X+Y)+(1-\mathbb{I})(G_1^{-1}(U)+G_2^{-1}(1-U)))$ and $\mathrm{Var}(\mathbb{I}(X+Y)+(1-\mathbb{I})(G_1^{-1}(U)+G_2^{-1}(U)))$ in Proposition 2.3.3.

Lemma 2.3.8. *Given Figure 2.13, we assume 1,2,3,4 in 2.3.3. For $x \in [0,1]$,*

$$G_1(x) = \begin{cases} \frac{x-C(x,b)}{1-C(a,b)} & \text{when} \quad x \le a \\ \frac{x-C(a,b)}{1-C(a,b)} & \text{when} \quad x > a \end{cases} \tag{2.70}$$

and

$$G_2(x) = \begin{cases} \frac{x-C(a,x)}{1-C(a,b)} & \text{when} \quad x \le b \\ \frac{x-C(a,b)}{1-C(a,b)} & \text{when} \quad x > b. \end{cases} \tag{2.71}$$

For $y \in [0,1]$,

$$G_1^{-1}(y) = \begin{cases} \inf\{0 \le x \le a : x - C(x,b) = yp_u\} & \text{when} \quad y \le G_1(a) \\ yp_u + p_f & \text{when} \quad y > G_1(a) \end{cases} \tag{2.72}$$

and

$$G_2^{-1}(y) = \begin{cases} \inf\{0 \le y \le b : x - C(a,x) = yp_u\} & \text{when} \quad y \le G_2(b) \\ yp_u + p_f & \text{when} \quad y > G_2(b), \end{cases} \tag{2.73}$$

where $C(x,y) = p_f H(x,y)$. (See Remark 2.3.6.)

Proof. First, $p_f = P(X \le a, Y \le b) = C(F_1(a), F_2(b)) = C(a,b)$. So $p_u = 1 - p_f = 1 - C(a,b)$.
For $x \in [0,1]$,

$$G_1(x) = \frac{P(X \le x, (X,Y) \in \mathcal{U})}{P((X,Y) \in \mathcal{U})} = \begin{cases} \frac{P(X \le x, Y \ge b)}{p_u} =: ① & \text{when} \quad x \le a \\ \frac{P(X \le a, Y \ge b) + P(a < X \le x, 0 \le Y \le 1)}{p_u} =: ② & \text{when} \quad x > a. \end{cases} \tag{2.74}$$

The two cases are shown in Figure 2.14.

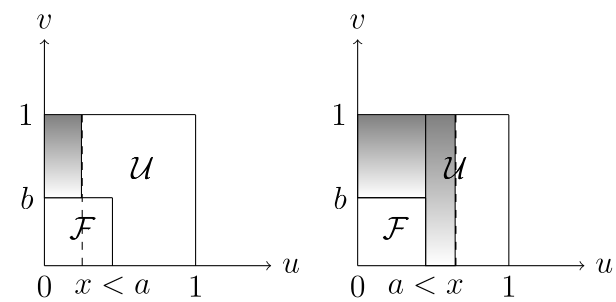


Figure 2.14: Graphs on calculating G_1

When $x \leq a$,

$$\begin{aligned} ① &= \frac{P(X \leq x) - P(X \leq x, Y < b)}{1 - p_f} = \frac{F_1(x) - C(F_1(x), F_2(b))}{1 - p_f} \\ &= \frac{x - C(x, b)}{1 - C(a, b)}. \end{aligned} \tag{2.75}$$

When $x > a$,

$$\begin{aligned} ② &= \frac{P(X \leq a) - P(X \leq a, Y \leq b) + P(a < X \leq x)}{1 - C(a, b)} \\ &= \frac{a - C(a, b) + x - a}{1 - C(a, b)} = \frac{x - C(a, b)}{1 - C(a, b)}. \end{aligned} \tag{2.76}$$

So

$$G_1(x) = \begin{cases} \frac{x - C(x,b)}{1 - C(a,b)} & \text{when} \quad x \leq a \\ \frac{x - C(a,b)}{1 - C(a,b)} & \text{when} \quad x > a. \end{cases} \tag{2.77}$$

On the other hand, for a fixed $x \in [0, 1]$

$$G_2(x) = \frac{P(Y \leq x, (X, Y) \in \mathcal{U})}{P((X, Y) \in \mathcal{U})} = \begin{cases} \frac{P(Y \leq x, X \geq a)}{p_u} =: ③ & \text{when} \quad x \leq b \\ \frac{P(Y \leq b, X \geq a) + P(b < Y \leq x, 0 \leq X \leq 1)}{p_u} =: ④ & \text{when} \quad x > b. \end{cases} \tag{2.78}$$

The two cases are shown in Figure 2.15.

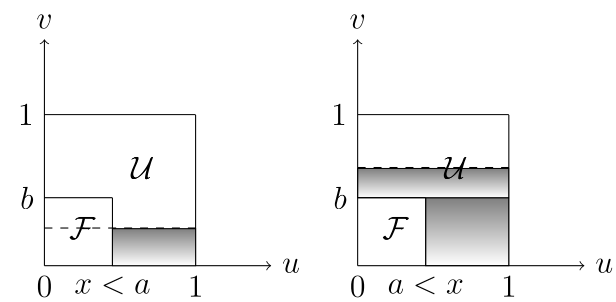


Figure 2.15: Graphs on calculating G_2

When $x \leq b$,

$$\begin{aligned} ③ &= \frac{P(Y \leq x) - P(X \leq a, Y \leq x)}{1 - p_f} = \frac{F_2(x) - C(F_1(a), F_2(x))}{1 - p_f} \\ &= \frac{x - C(a, x)}{1 - C(a, b)}. \end{aligned} \tag{2.79}$$

When $x > b$,

$$\begin{aligned} ④ &= \frac{P(Y \leq b) - P(X \leq a, Y \leq b) + P(b < Y \leq x)}{1 - C(a, b)} \\ &= \frac{b - C(a, b) + x - b}{1 - C(a, b)} = \frac{x - C(a, b)}{1 - C(a, b)}. \end{aligned} \tag{2.80}$$

So

$$G_2(x) = \begin{cases} \frac{x - C(a,x)}{1 - C(a,b)} & \text{when} \quad x \leq b \\ \frac{x - C(a,b)}{1 - C(a,b)} & \text{when} \quad x > b. \end{cases} \tag{2.81}$$

To calculate $G_i^{-1}(U)$ for $i = 1, 2$, first fix $y \in [0, 1]$, take $i = 1$, look at the generalized inverse $G_1^{-1}(y) = \inf\{0 \leq x \leq 1 : G_1(x) = y\}$. Since G_1 is a distribution function, it is non-decreasing in x, and $G_1(a) = \frac{a - C(a,b)}{1 - C(a,b)}$, we can now divide $G_1^{-1}(y)$ into two cases: $y \leq G_1(a)$ and $y > G_1(a)$.

So

$$G_1^{-1}(y) = \begin{cases} \inf\{0 \le x \le a : \frac{x - C(x,b)}{1 - C(a,b)} = y\} & \text{when} \quad y \le G_1(a) \\ \inf\{1 \ge x > a : \frac{x - C(a,b)}{1 - C(a,b)} = y\} & \text{when} \quad y > G_1(a) \end{cases}$$
$$\Longrightarrow G_1^{-1}(y) = \begin{cases} \inf\{0 \le x \le a : x - C(x,b) = yp_u\} & \text{when} \quad y \le G_1(a) \\ yp_u + p_f & \text{when} \quad y > G_1(a). \end{cases} \tag{2.82}$$

By symmetry,

$$G_2^{-1}(y) = \begin{cases} \inf\{0 \le y \le b : x - C(a,x) = yp_u\} & \text{when} \quad y \le G_2(b) \\ yp_u + p_f & \text{when} \quad y > G_2(b). \end{cases} \tag{2.83}$$

□

Then we calculate bounds on variance assuming C is the independence copula.

Lemma 2.3.9. *Given Figure 2.13, we assume 1,2,3,4 in Proposition 2.3.3, take $C(u,v) = uv$ the independence copula, then*

$$\mathrm{Var}(G_1^{-1}(U) + G_2^{-1}(1-U)) = \frac{a^2b^2((4-3b)b + a^2(-3+8b-4b^2) + 2a(2-7b+4b^2))}{12(ab-1)^2} \tag{2.84}$$

and

$$\mathrm{Var}(G_1^{-1}(U) + G_2^{-1}(U)) = \begin{cases} \frac{a^4b + 2a^3(b-1)b + a^2b^3 - 2a\left(b^3+2\right) + 4}{3(a-1)(ab-1)} - \frac{\left(a^2b + ab^2 - 2\right)^2}{(2-2ab)^2} & \text{if} \quad a \le b \\ \frac{ab^4 + 2ab^3(a-1) + a^3b^2 - 2b(a^3+2) + 4}{3(b-1)(ab-1)} - \frac{\left(a^2b + ab^2 - 2\right)^2}{(2-2ab)^2} & \text{if} \quad a \ge b. \end{cases} \tag{2.85}$$

Proof. When $C(u,v) = uv$, then from Lemma 2.3.8,

$$G_1(x) = \begin{cases} \frac{x - xb}{1 - ab} & \text{when} \quad x \le a \\ \frac{x - ab}{1 - ab} & \text{when} \quad x > a \end{cases} \tag{2.86}$$

$$G_2(x) = \begin{cases} \frac{x-ax}{1-ab} & \text{when} \quad x \leq b \\ \frac{x-ab}{1-ab} & \text{when} \quad x > b. \end{cases} \tag{2.87}$$

For $y \in [0,1]$,

$$G_1^{-1}(y) = \begin{cases} \frac{y(1-ab)}{1-b} & \text{when} \quad y \leq G_1(a) \\ y(1-ab)+ab & \text{when} \quad y > G_1(a) \end{cases} \tag{2.88}$$

$$G_2^{-1}(y) = \begin{cases} \frac{y(1-ab)}{1-a} & \text{when} \quad y \leq G_2(b) \\ y(1-ab)+ab & \text{when} \quad y > G_2(b). \end{cases} \tag{2.89}$$

We calculate $\mathrm{Var}(G_1^{-1}(U) + G_2^{-1}(1-U))$ first. Since

$$\mathrm{Var}(G_1^{-1}(U)+G_2^{-1}(1-U)) = E[(G_1^{-1}(U)+G_2^{-1}(1-U))^2] - (E[G_1^{-1}(U)] + E[G_2^{-1}(1-U)])^2 \tag{2.90}$$

and

$$\begin{aligned} E[G_1^{-1}(U)] &= \int_0^1 G_1^{-1}(u) \cdot 1du \\ &= \int_0^{G_1(a)} \frac{u(1-ab)}{1-b} du + \int_{G_1(a)}^1 [u(1-ab)+ab] du. \end{aligned} \tag{2.91}$$

Since $G_1(a) = \frac{a-ab}{1-ab}$,

$$\begin{aligned} \text{Equation} \quad (2.91) &= \int_0^{\frac{a-ab}{1-ab}} \frac{u(1-ab)}{1-b} du + \int_{\frac{a-ab}{1-ab}}^1 [u(1-ab)+ab] du \\ &= \frac{a^2b}{2ab-2} - \frac{a^2}{2ab-2} + \frac{a^2-1}{2ab-2} \\ &= \frac{a^2b-1}{2ab-2}. \end{aligned} \tag{2.92}$$

By Equation (2.89),

$$G_2^{-1}(1-u) = \begin{cases} \frac{(1-u)(1-ab)}{1-a} & \text{when} \quad u \geq 1-G_2(b) \\ (1-u)(1-ab)+ab & \text{when} \quad u < 1-G_2(b). \end{cases} \tag{2.93}$$

So

$$E[G_2^{-1}(1-U)] = \int_0^1 G_2^{-1}(1-u)\cdot 1du$$
$$= \int_0^{1-G_2(b)} [(1-u)(1-ab)+ab]du + \int_{1-G_2(b)}^1 [\frac{(1-u)(1-ab)}{1-a}]du. \tag{2.94}$$

Since $G_2(b) = \frac{b-ab}{1-ab}$,

$$\text{Equation} \quad (2.94) = \int_0^{1-\frac{b-ab}{1-ab}} [(1-u)(1-ab)+ab]du + \int_{1-\frac{b-ab}{1-ab}}^1 [\frac{(1-u)(1-ab)}{1-a}]du$$
$$= \frac{b^2-1}{2ab-2} + \frac{(a-1)b^2}{2ab-2} \tag{2.95}$$
$$= \frac{ab^2-1}{2ab-2},$$

$$E[(G_1^{-1}(U)+G_2^{-1}(1-U))^2] = \int_0^1 [G_1^{-1}(u)+G_2^{-1}(1-u)]^2 du. \tag{2.96}$$

Since

$$G_1(a) = \frac{a-ab}{1-ab} \leq \frac{1-b}{1-ab} = 1-G_2(b), \tag{2.97}$$

then

$$G_1^{-1}(u)+G_2^{-1}(1-u) = \begin{cases} \frac{u(1-ab)}{1-b} + (1-u)(1-ab)+ab & \text{if} \quad u \leq G_1(a) \\ u(1-ab)+ab+(1-u)(1-ab)+ab & \text{if} \quad G_1(a) \leq u \leq 1-G_2(b) \\ u(1-ab)+ab+\frac{(1-u)(1-ab)}{1-a} & \text{if} \quad 1-G_2(b) \leq u \end{cases}$$
$$= \begin{cases} \frac{ab^2u-bu+b-1}{b-1} & \text{if} \quad u \leq G_1(a) \\ 1+ab & \text{if} \quad G_1(a) \leq u \leq 1-G_2(b) \\ \frac{a^2(b-bu)+au-1}{a-1} & \text{if} \quad 1-G_2(b) \leq u. \end{cases} \tag{2.98}$$

So Equation (2.96) =

$$\begin{aligned}
&\int_0^{G_1(a)} (\frac{ab^2u - bu + b - 1}{b-1})^2 du + \int_{G_1(a)}^{1-G_2(b)} (1+ab)^2 du + \int_{1-G_2(b)}^{1} (\frac{a^2(b-bu) + au - 1}{a-1})^2 du \\
&= \frac{a(b-1)\left(a^2b^2 + 3ab + 3\right)}{3ab-3} - \frac{(a-1)(b-1)(ab+1)^2}{ab-1} + \frac{(a-1)b\left(a^2b^2+3ab+3\right)}{3ab-3} \\
&= \frac{a^3(-(b-2))b^2 + a^2b\left(2b^2 - 3b + 3\right) + 3a(b-1)b - 3}{3ab-3}.
\end{aligned} \tag{2.99}$$

So

$$\begin{aligned}
&\mathrm{Var}(G_1^{-1}(U) + G_2^{-1}(1-U)) = \text{Equation (2.96)} - (\text{Equation (2.92)} + \text{Equation (2.95)})^2 \\
=&\frac{a^2b^2((4-3b)b + a^2(-3+8b-4b^2) + 2a(2-7b+4b^2))}{12(ab-1)^2}.
\end{aligned} \tag{2.100}$$

Then we calculate $\mathrm{Var}(G_1^{-1}(U) + G_2^{-1}(U))$.
Since

$$\mathrm{Var}(G_1^{-1}(U) + G_2^{-1}(U)) = E[(G_1^{-1}(U) + G_2^{-1}(U))^2] - (E[G_1^{-1}(U)] + E[G_2^{-1}(U)])^2, \tag{2.101}$$

and by Equation (2.92), we already know

$$E[G_1^{-1}(U)] = \frac{a^2b-1}{2ab-2}. \tag{2.102}$$

By Equation (2.89),

$$\begin{aligned}
E[G_2^{-1}(U)] &= \int_0^1 G_2^{-1}(u) \cdot 1 du \\
&= \int_{G_2(b)}^1 [u(1-ab) + ab] du + \int_0^{G_2(b)} [\frac{u(1-ab)}{1-a}] du.
\end{aligned} \tag{2.103}$$

Since $G_2(b) = \frac{b-ab}{1-ab}$,

$$\begin{aligned}
\text{Equation} \quad (2.103) =& \int_{\frac{b-ab}{1-ab}}^{1} [u(1-ab)+ab]du + \int_0^{\frac{b-ab}{1-ab}} [\frac{u(1-ab)}{1-a}]du \\
=& \frac{b^2-1}{2ab-2} + \frac{(a-1)b^2}{2ab-2} \qquad (2.104) \\
=& \frac{ab^2-1}{2ab-2},
\end{aligned}$$

and

$$E[(G_1^{-1}(U) + G_2^{-1}(U))^2] = \int_0^1 [G_1^{-1}(u) + G_2^{-1}(u)]^2 du. \qquad (2.105)$$

The value of $G_1^{-1}(u) + G_2^{-1}(u)$ depends on whether $a \leq b$ or $a \geq b$.

Case 1: $a \leq b$

As

$$G_1(a) = \frac{a-ab}{1-ab} \leq \frac{b-ab}{1-ab} = G_2(b), \qquad (2.106)$$

then

$$\begin{aligned}
G_1^{-1}(u) + G_2^{-1}(u) &= \begin{cases} \frac{u(1-ab)}{1-b} + \frac{u(1-ab)}{1-a} & \text{if} \quad u \leq G_1(a) \\ u(1-ab) + ab + \frac{u(1-ab)}{1-a} & \text{if} \quad G_1(a) \leq u \leq G_2(b) \\ 2u(1-ab) + 2ab & \text{if} \quad G_2(b) \leq u \end{cases} \\
&= \begin{cases} \frac{u(a+b-2)(ab-1)}{(a-1)(b-1)} & \text{if} \quad u \leq G_1(a) \\ -abu + \frac{u(ab-1)}{a-1} + ab + u & \text{if} \quad G_1(a) \leq u \leq G_2(b) \\ -2abu + 2ab + 2u & \text{if} \quad G_2(b) \leq u. \end{cases}
\end{aligned} \qquad (2.107)$$

So $\int_0^1 [G_1^{-1}(u) + G_2^{-1}(u)]^2 du =$

$$
\begin{aligned}
&\int_0^{G_1(a)} (\frac{u(a+b-2)(ab-1)}{(a-1)(b-1)})^2 du + \int_{G_1(a)}^{G_2(b)} (-abu + \frac{u(ab-1)}{a-1} + ab + u)^2 du \\
&+ \int_{G_2(b)}^1 (-2abu + 2ab + 2u)^2 du \\
&= \frac{a^3(b-1)(a+b-2)^2}{3(a-1)^2(ab-1)} + \frac{3a^3b^2 + 3(a-2)a^3b + (a-2)^2a^3 + ((10-7a)a-4)b^3}{3(a-1)^2(ab-1)} \\
&+ \frac{4\left(b^3-1\right)}{3ab-3} \\
&= \frac{a^4b + 2a^3(b-1)b + a^2b^3 - 2a\left(b^3+2\right) + 4}{3(a-1)(ab-1)}.
\end{aligned}
\tag{2.108}
$$

So

$$
\begin{aligned}
&\mathrm{Var}(G_1^{-1}(U) + G_2^{-1}(U)) = \text{Equation (2.108)} - (\text{Equation (2.92)} + \text{Equation (2.104)})^2 \\
=&\frac{a^4b + 2a^3(b-1)b + a^2b^3 - 2a\left(b^3+2\right) + 4}{3(a-1)(ab-1)} - \frac{\left(a^2b + ab^2 - 2\right)^2}{(2-2ab)^2}.
\end{aligned}
\tag{2.109}
$$

<u>Case 2:</u> $b \le a$

As

$$
G_1(a) = \frac{a-ab}{1-ab} \ge \frac{b-ab}{1-ab} = G_2(b), \tag{2.110}
$$

then

$$
\begin{aligned}
G_1^{-1}(u) + G_2^{-1}(u) &= \begin{cases} \frac{u(1-ab)}{1-b} + \frac{u(1-ab)}{1-a} & \text{if} \quad u \le G_2(b) \\ u(1-ab) + ab + \frac{u(1-ab)}{1-b} & \text{if} \quad G_2(b) \le u \le G_1(a) \\ 2u(1-ab) + 2ab & \text{if} \quad G_1(a) \le u \end{cases} \\
&= \begin{cases} \frac{u(a+b-2)(ab-1)}{(a-1)(b-1)} & \text{if} \quad u \le G_2(b) \\ \frac{(b-2)u(1-ab)}{b-1} + ab & \text{if} \quad G_2(b) \le u \le G_1(a) \\ -2abu + 2ab + 2u & \text{if} \quad G_1(a) \le u. \end{cases}
\end{aligned}
\tag{2.111}
$$

So $\int_0^1 [G_1^{-1}(u) + G_2^{-1}(u)]^2 du =$

$$\int_0^{G_2(b)} (\frac{u(a+b-2)(ab-1)}{(a-1)(b-1)})^2 du + \int_{G_2(b)}^{G_1(a)} (\frac{(b-2)u(1-ab)}{b-1} + ab)^2 du$$
$$+ \int_{G_1(a)}^{1} (-2abu + 2ab + 2u)^2 du$$
$$= \frac{(a-1)b^3(a+b-2)^2}{3(b-1)^2(ab-1)} + \frac{a^3(-7b^2+10b-4) + 3a^2b^3 + 3a(b-2)b^3 + (b-2)^2b^3}{3(b-1)^2(ab-1)} + \frac{4(a^3-1)}{3ab-3}$$
$$= \frac{ab^4 + 2ab^3(a-1) + a^3b^2 - 2b(a^3+2) + 4}{3(b-1)(ab-1)}. \tag{2.112}$$

So

$$\mathrm{Var}(G_1^{-1}(U) + G_2^{-1}(U)) = \text{Equation (2.112)} - (\text{Equation (2.92)} + \text{Equation (2.104)})^2$$
$$= \frac{ab^4 + 2ab^3(a-1) + a^3b^2 - 2b(a^3+2) + 4}{3(b-1)(ab-1)} - \frac{\left(a^2b + ab^2 - 2\right)^2}{(2-2ab)^2}. \tag{2.113}$$

In summary,

$$\mathrm{Var}(G_1^{-1}(U) + G_2^{-1}(U)) = \begin{cases} \frac{a^4b + 2a^3(b-1)b + a^2b^3 - 2a\left(b^3+2\right) + 4}{3(a-1)(ab-1)} - \frac{\left(a^2b+ab^2-2\right)^2}{(2-2ab)^2} & \text{if} \quad a \le b \\ \frac{ab^4 + 2ab^3(a-1) + a^3b^2 - 2b(a^3+2) + 4}{3(b-1)(ab-1)} - \frac{\left(a^2b+ab^2-2\right)^2}{(2-2ab)^2} & \text{if} \quad a \ge b. \end{cases} \tag{2.114}$$

□

Proposition 2.3.10. *(Minimum Variance)*
Given Figure 2.13, assuming 1,2,3,4 in Proposition 2.3.3, if we take $C(u, v) = uv$, *the independence copula, then the lower bound of* $\mathrm{Var}(X + Y)$ *is:*

$$\mathrm{Var}\left(\mathbb{I}(X+Y) + (1-\mathbb{I})(G_1^{-1}(U) + G_2^{-1}(1-U))\right)$$
$$= \frac{a^3b^3}{3} - \frac{2a^3b^2}{3} + \frac{a^3b}{3} - \frac{2a^2b^3}{3} + \frac{3a^2b^2}{2} - a^2b + \frac{ab^3}{3} - ab^2 + ab, \tag{2.115}$$

where $U, G_1, G_2, \mathbb{I}$ *are defined in Proposition 2.3.3.*

Proof. Denote $A_1 := G_1^{-1}(U), A_2 := G_2^{-1}(1-U)$, from Proposition 2.3.3, we know $\text{Var}(X+Y) \geq \text{Var}(\mathbb{I}(X+Y)+(1-\mathbb{I})(A_1+A_2))$. So we only need to prove

$$\text{Var}(\mathbb{I}(X+Y)+(1-\mathbb{I})(A_1+A_2)) \tag{2.116}$$

$$=\frac{a^3b^3}{3}-\frac{2a^3b^2}{3}+\frac{a^3b}{3}-\frac{2a^2b^3}{3}+\frac{3a^2b^2}{2}-a^2b+\frac{ab^3}{3}-ab^2+ab. \tag{2.117}$$

Denote $(\mathbb{I}(X+Y)+(1-\mathbb{I})(A_1+A_2)) := Z$, then by the conditional variance formula,

$$\text{Var}(Z) = E[\text{Var}(Z|\mathbb{I})] + \text{Var}[E(Z|\mathbb{I})].$$

And

$$Z|\mathbb{I} = \begin{cases} X+Y|\mathbb{I}=1 \\ A_1+A_2|\mathbb{I}=0. \end{cases} \tag{2.118}$$

So

$$E[Z|\mathbb{I}=1] = E(X+Y|(X,Y)\in\mathcal{F}). \tag{2.119}$$

To calculate Equation (2.119), we need to know the distribution of $X|(X,Y)\in\mathcal{F}$ and $Y|(X,Y)\in\mathcal{F}$.
For any $x \geq 0$,

$$F_{X|(X,Y)\in\mathcal{F}}(x) = \frac{P(X\leq x,(X,Y)\in\mathcal{F})}{P((X,Y)\in\mathcal{F})} = \begin{cases} \dfrac{P(X\leq x, Y\leq b)}{P(X\leq a, Y\leq b)} & \text{if} \quad x\leq a \\ \dfrac{P(X\leq a, Y\leq b)}{P(X\leq a, Y\leq b)} & \text{if} \quad x> a \end{cases} = \begin{cases} \dfrac{bx}{ab}=\dfrac{x}{a} & \text{if} \quad x\leq a \\ 1 & \text{if} \quad x>a. \end{cases} \tag{2.120}$$

So $X|(X,Y)\in\mathcal{F}$ follows $U[0,a]$. Due to symmetry, $Y|(X,Y)\in\mathcal{F}$ follows $U[0,b]$.
So

$$\text{Equation (2.119)} = E(X|(X,Y)\in\mathcal{F}) + E(Y|(X,Y)\in\mathcal{F}) = \frac{a}{2}+\frac{b}{2}. \tag{2.121}$$

Since when $(X,Y)\in\mathcal{F}$, (X,Y) follows independence copula,

$$\text{Var}[Z|\mathbb{I}=1] = \text{Var}(X+Y|(X,Y)\in\mathcal{F}) = \text{Var}(X|(X,Y)\in\mathcal{F})+\text{Var}(Y|(X,Y)\in\mathcal{F}) = \frac{a^2}{12}+\frac{b^2}{12}.$$

As U is independent of $\mathbb{I}$, then $A_1 = G_1^{-1}(U)$ and $A_2 = G_2^{-1}(1-U)$ are independent of $\mathbb{I}$, by Equations (2.91) and (2.94),

$$E[Z|\mathbb{I}=0] = E(A_1 + A_2|\mathbb{I}=0) = E(A_1 + A_2) = \frac{a^2b + ab^2 - 2}{2(ab-1)},$$

by Lemma 2.4.2,

$$\text{Var}[Z|\mathbb{I}=0] = \text{Var}(A_1 + A_2|\mathbb{I}=0) = \text{Var}(A_1 + A_2) = \text{Equation (2.100)} := \diamondsuit.$$

So

$$\begin{aligned} E[\text{Var}[Z|\mathbb{I}]] &= \frac{a^2+b^2}{12}P(\mathbb{I}=1) + (\diamondsuit)P(\mathbb{I}=0) \\ &= \frac{a^2+b^2}{12}p_f + (\diamondsuit)p_u \\ &= \frac{a^2+b^2}{12}ab + (1-ab)(\diamondsuit) \\ &= \frac{ab\left(4a^3(b-1)^2b - a^2\left(8b^3 - 14b^2 + 4b + 1\right) + 4a(b-1)b^2 - b^2\right)}{12(ab-1)}. \end{aligned} \tag{2.122}$$

And

$$\begin{aligned} E[E[Z|\mathbb{I}]] &= \frac{a+b}{2} \cdot P(\mathbb{I}=1) + \frac{a^2b + ab^2 - 2}{2(ab-1)}P(\mathbb{I}=0) \\ &= \frac{a+b}{2}(ab) + \frac{a^2b + ab^2 - 2}{2(ab-1)}(1-ab) \\ &= 1. \end{aligned} \tag{2.123}$$

So

$$\begin{aligned} \text{Var}(E(Z|\mathbb{I})) &= E[E^2(Z|\mathbb{I})] - E^2[E[Z|\mathbb{I}]] \\ &= (\frac{a+b}{2})^2 \cdot P(\mathbb{I}=1) + (\frac{a^2b + ab^2 - 2}{2(ab-1)})^2 \cdot P(\mathbb{I}=0) - 1^2 \\ &= -\frac{ab(a+b-2)^2}{4ab-4}. \end{aligned} \tag{2.124}$$

We add Equations (2.122) and (2.124) together to get the answer,

$$\mathrm{Var}(Z) = \frac{a^3b^3}{3} - \frac{2a^3b^2}{3} + \frac{a^3b}{3} - \frac{2a^2b^3}{3} + \frac{3a^2b^2}{2} - a^2b + \frac{ab^3}{3} - ab^2 + ab. \tag{2.125}$$

This is the same equation as Equation (2.35). □

Proposition 2.3.11. *(Maximum Variance)*
Given Figure 2.13, assuming 1,2,3,4 in Proposition 2.3.3, if we take $C(u,v) = uv$*, the independence copula, then the upper bound of* $\mathrm{Var}(X+Y)$ *is:*

$$\mathrm{Var}\left(\mathbb{I}(X+Y) + (1-\mathbb{I})(G_1^{-1}(U) + G_2^{-1}(U))\right) = \begin{cases} -\frac{a^3(b-2)b+3a^2b^2-2a\left(b^3+1\right)+2}{6(a-1)} & \text{if} \quad a \le b \\ -\frac{b^3(a-2)a+3a^2b^2-2b\left(a^3+1\right)+2}{6(b-1)} & \text{if} \quad a \ge b, \end{cases} \tag{2.126}$$

where $U, G_1, G_2, \mathbb{I}$ *are defined in Proposition 2.3.3. The two equations above can also be written as*

$$-\frac{(\min(a,b))^3(\max(a,b)-2)\max(a,b) + 3a^2b^2 - 2\min(a,b)((\max(a,b))^3+1)+2}{6(\min(a,b)-1)}. \tag{2.127}$$

Proof. Denote $B_1 := G_1^{-1}(U), B_2 := G_2^{-1}(U)$, from Proposition 2.3.3, we know $\mathrm{Var}(X+Y) \le \mathrm{Var}(\mathbb{I}(X+Y) + (1-\mathbb{I})(B_1+B_2))$. So we only need to prove

$$\mathrm{Var}(\mathbb{I}(X+Y) + (1-\mathbb{I})(B_1+B_2)) \tag{2.128}$$

$$= \begin{cases} -\frac{a^3(b-2)b+3a^2b^2-2a\left(b^3+1\right)+2}{6(a-1)} & \text{if} \quad a \le b \\ -\frac{b^3(a-2)a+3a^2b^2-2b\left(a^3+1\right)+2}{6(b-1)} & \text{if} \quad a \ge b. \end{cases} \tag{2.129}$$

Denote $(\mathbb{I}(X+Y) + (1-\mathbb{I})(B_1+B_2)) := Z$, then by the conditional variance formula,

$$\mathrm{Var}(Z) = E[\mathrm{Var}(Z|\mathbb{I})] + \mathrm{Var}[E(Z|\mathbb{I})].$$

And

$$Z|\mathbb{I} = \begin{cases} X+Y|\mathbb{I} = 1 \\ B_1+B_2|\mathbb{I} = 0. \end{cases} \tag{2.130}$$

From proof in Proposition 2.3.10,

$$E[Z|\mathbb{I}=1] = \frac{a+b}{2} \tag{2.131}$$

and

$$\mathrm{Var}[Z|\mathbb{I}=1] = \frac{a^2+b^2}{12}.$$

As U is independent of $\mathbb{I}$, then $B_1 = G_1^{-1}(U)$ and $B_2 = G_2^{-1}(U)$ are independent of $\mathbb{I}$, by Equations (2.92) and (2.104),

$$E[Z|\mathbb{I}=0] = E(B_1+B_2|\mathbb{I}=0) = E(B_1+B_2) = \frac{a^2b+ab^2-2}{2(ab-1)}.$$

When $a \le b$, by Lemma 2.4.2,

$$\mathrm{Var}[Z|\mathbb{I}=0] = \mathrm{Var}(B_1+B_2|\mathbb{I}=0) = \mathrm{Var}(B_1+B_2) = \text{Equation (2.109)} := \diamondsuit.$$

So

$$\begin{aligned}
E[\mathrm{Var}[Z|\mathbb{I}]] &= \frac{a^2+b^2}{12}P(\mathbb{I}=1) + (\diamondsuit)P(\mathbb{I}=0)\\
&= \frac{a^2+b^2}{12}p_f + (\diamondsuit)p_u\\
&= \frac{a^2+b^2}{12}ab + (1-ab)(\diamondsuit)\\
&= \frac{a^4b\left(-2b^2+4b+3\right)+a^3b\left(-6b^2+8b-19\right)+a^2b\left(4b^3+3b^2-12b+28\right)}{12(a-1)(ab-1)}\\
&- \frac{a\left(7b^3-12b^2+16b+4\right)-4}{12(a-1)(ab-1)}
\end{aligned} \tag{2.132}$$

And from proof in Proposition 2.3.10,

$$\mathrm{Var}(E(Z|\mathbb{I})) = -\frac{ab(a+b-2)^2}{4ab-4}. \tag{2.133}$$

We add Equations (2.132) and (2.133) together to get the answer,

$$\mathrm{Var}(Z) = -\frac{a^3(b-2)b + 3a^2b^2 - 2a\left(b^3+1\right) + 2}{6(a-1)}. \tag{2.134}$$

Similarly, when $a \geq b$,

$$\mathrm{Var}(Z) = -\frac{-2\left(a^3+1\right)b + 3a^2b^2 + (a-2)ab^3 + 2}{6(b-1)}. \tag{2.135}$$

In summary,

$$\mathrm{Var}(\mathbb{I}(X+Y) + (1-\mathbb{I})(B_1+B_2)) = \begin{cases} -\frac{a^3(b-2)b+3a^2b^2-2a\left(b^3+1\right)+2}{6(a-1)} & \text{if} \quad a \leq b \\ -\frac{b^3(a-2)a+3a^2b^2-2b\left(a^3+1\right)+2}{6(b-1)} & \text{if} \quad a \geq b \end{cases} \tag{2.136}$$

□

This is different from Equation (2.59).

Corollary 2.3.12. *(Variance bounds using Bernard and Vanduffel [2014]) Given Figure 2.13, assuming 1,2,3,4 in Proposition 2.3.3, if we take $C(u,v) = uv$, the independence copula, then upper (sharp) and lower bounds (not sharp) of* $\mathrm{Var}(X+Y)$ *are as follows,*

$$\frac{a^3b^3}{3} - \frac{2a^3b^2}{3} + \frac{a^3b}{3} - \frac{2a^2b^3}{3} + \frac{3a^2b^2}{2} - a^2b + \frac{ab^3}{3} - ab^2 + ab \leq \mathrm{Var}(X+Y) \leq \\ -\frac{(\min(a,b))^3(\max(a,b)-2)\max(a,b) + 3a^2b^2 - 2\min(a,b)((\max(a,b))^3+1) + 2}{6(\min(a,b)-1)}. \tag{2.137}$$

Proof. This directly follows from Proposition 2.3.3, 2.3.10 and 2.3.11. □

2.4 Improved Upper and Lower Bounds of $\mathrm{Var}(X+Y)$ Using Convex Order Bounds

From both the theoretical result (see Equations (2.3.10) and (2.3.11)) and the simulation details (see Sections 2.7.2 and 2.7.3 of Appendix 2.7), the lower variance bound derived using Bernard and Vanduffel [2014] is sharp but the upper bound is not (see Corollary 2.3.12). So now we propose another method using convex order bounds and call it improved bounds, which give sharp bounds on variance.

Proposition 2.4.1. *Fix some* $0 < a < 1$ *and* $0 < b < 1$ *such that* $\mathcal{F} = [0, a] \times [0, b]$, $\mathcal{U}_1 = [a, 1] \times [0, b]$, $\mathcal{U}_2 = [a, 1] \times [b, 1]$ *and* $\mathcal{U}_3 = [0, a] \times [b, 1]$. *We have the following assumptions:*
1. X, Y *are two random variables following* $U[0, 1]$*;*
2. (X, Y) *has copula* C *on area* $\mathcal{F}$*;*
3. U *is* $U[0, 1]$ *independent of events* "$(X, Y) \in \mathcal{F}, (X, Y) \in \mathcal{U}_i, i = 1, 2, 3$".
For some $x \in [0, 1]$, *denote*

$$G_{2i-1}(x) := F_{X|(X,Y)\in\mathcal{U}_i}(x), G_{2i}(x) := F_{Y|(X,Y)\in\mathcal{U}_i}(x), i = 1, 2, 3,$$

then bounds on variance are

$$\mathrm{Var}(\mathbb{I}_{\mathcal{F}}(X+Y) + \sum_{i=1}^{3} \mathbb{I}_{\mathcal{U}_i}(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(1-U)) \leq \mathrm{Var}(X+Y) \leq \mathrm{Var}(\mathbb{I}_{\mathcal{F}}(X+Y) + \sum_{i=1}^{3} \mathbb{I}_{\mathcal{U}_i}(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(U)) \quad (2.138)$$

Proof. The proof is similar with the proof in Proposition 2.3 just by replacing $m = 3$. □

Assume we know how the masses are distributed on area $[0, 1]^2\backslash\mathcal{F}$, then we calculate these improved bounds under special case.

Proposition 2.4.2. *(Minimum Variance) Given Figure 2.16 and assumptions in Proposition 2.4.1, if we take* C *as the independence copula, then*

$$\mathrm{Var}\left(\mathbb{I}_{\mathcal{F}}(X+Y) + \sum_{i=1}^{3} \mathbb{I}_{\mathcal{U}_i}(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(1-U))\right) = \frac{a^3b^3}{3} - \frac{2a^3b^2}{3} + \frac{a^3b}{3} - \frac{2a^2b^3}{3} + \frac{3a^2b^2}{2} - a^2b + \frac{ab^3}{3} - ab^2 + ab, \quad (2.139)$$

which is the same as Equation (2.35), *where* $U \sim U[0, 1]$.

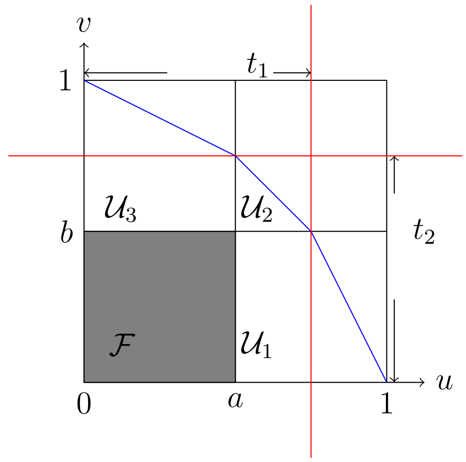


Figure 2.16: When $\mathcal{F} = [0, a] \times [0, b]$

Proof.

$$G_1(x) = F_{X|(X,Y)\in\mathcal{U}_1}(x) = \frac{P(X \le x, t_1 \le X, Y \le b)}{P(a \le X \le, Y \le b)} = \frac{x - t_1}{b - ab} \text{ if } a \le x \le t_1. \tag{2.140}$$

Since $G_1(t_1)$ must be 1, we get $t_1 = 1 - b + ab$. So $G_1^{-1}(u) = u(b - ab) + 1 - b + ab$. Similarly,

$$G_2(x) = F_{Y|(X,Y)\in\mathcal{U}_1}(x) = \frac{P(Y \le x, a \le X, Y \le b)}{P(a \le X, Y \le b)} = \frac{x - ax}{b - ab} = \frac{x}{b} \text{ if } x \le b, \tag{2.141}$$

so $G_2^{-1}(u) = ub$.

$$G_3(x) = F_{X|(X,Y)\in\mathcal{U}_2}(x) = \frac{P(X \le x, a \le X, Y \ge t_2)}{P(a \le X, Y \ge b)} = \frac{x - a}{1 - a - b + ab}, \tag{2.142}$$

so $G_3^{-1}(u) = u(1 - a - b + ab) + a$.

$$G_4(x) = F_{Y|(X,Y)\in\mathcal{U}_2}(x) = \frac{P(a \le X \le t_1, b \le Y \le x)}{P(a \le X, b \le Y)} = \frac{x - b}{1 - a - b + ab} \text{ when } x \le t_2, \tag{2.143}$$

so $G_4^{-1}(u) = u(1-a-b+ab)+b$. When $x = t_2, G_4(t_2) = 1$, so $t_2 = 1-a+ab$.

$$G_5(x) = F_{X|(X,Y)\in\mathcal{U}_3}(x) = \frac{P(X \le x, X \le a, b \le Y)}{P(X \le a, b \le Y)} = \frac{x-bx}{a-ab} = \frac{x}{a}, \tag{2.144}$$

so $G_5^{-1}(u) = au$.

$$G_6(x) = F_{Y|(X,Y)\in\mathcal{U}_3}(x) = \frac{P(t_2 \le Y \le x, X \le a)}{P(X \le a, b \le Y)} = \frac{x-(1-a+ab)}{a-ab}, \tag{2.145}$$

so $G_6^{-1}(u) = u(a-ab)+1-a+ab$.
For convenience, denote $A_{2i-1} := G_{2i-1}^{-1}(u), A_{2i} := G_{2i}^{-1}(u), i = 1,2,3$.
Define another random variable T such that,

$$T = \begin{cases} 1 \text{ if } \mathbb{I}_{\mathcal{F}} = 1 \\ 2 \text{ if } \mathbb{I}_{\mathcal{U}_1} = 1 \\ 3 \text{ if } \mathbb{I}_{\mathcal{U}_2} = 1 \\ 4 \text{ if } \mathbb{I}_{\mathcal{U}_3} = 1. \end{cases} \tag{2.146}$$

Denote $Z := \mathbb{I}_{\mathcal{F}}(X+Y) + \mathbb{I}_{\mathcal{U}_1}(G_1^{-1}(U) + G_2^{-1}(1-U)) + \mathbb{I}_{\mathcal{U}_2}(G_3^{-1}(U) + G_4^{-1}(1-U)) + \mathbb{I}_{\mathcal{U}_3}(G_5^{-1}(U) + G_6^{-1}(1-U))$. Then $\text{Var}(Z) = E(\text{Var}(Z|T)) + \text{Var}(E(Z|T))$.
And the distributions of $(X,Y)|(X,Y) \in \mathcal{F}$ are computed in Proposition 2.3.10, thus

$$\begin{aligned} E(\text{Var}(Z|T)) &= \sum_{i=1}^{4} P(T=i)\text{Var}(Z|T=i) \\ &= ab\text{Var}(X+Y|(X,Y)\in\mathcal{F}) + (b-ab)\text{Var}(A_1+A_2|(X,Y)\in\mathcal{U}_1) \\ &+ (1-a-b+ab)\text{Var}(A_3+A_4|(X,Y)\in\mathcal{U}_2) + (a-ab)\text{Var}(A_5+A_6|(X,Y)\in\mathcal{U}_3) \\ &= -\frac{1}{6}a^3b^3 + \frac{a^3b^2}{12} + \frac{a^3b}{12} + \frac{a^2b^3}{12} + \frac{ab^3}{12}, \end{aligned} \tag{2.147}$$

$$\begin{aligned} \text{Var}(E(Z|T)) &= \sum_{i=1}^{4} P(T=i)E^2(Z|T=i) - \left(\sum_{i=1}^{4} P(T=i)E(Z|T=i)\right)^2 \\ &= \frac{a^3b^3}{2} - \frac{3a^3b^2}{4} + \frac{a^3b}{4} - \frac{3a^2b^3}{4} + \frac{3a^2b^2}{2} - a^2b + \frac{ab^3}{4} - ab^2 + ab. \end{aligned} \tag{2.148}$$

So

$$\mathrm{Var}(Z) = \frac{a^3b^3}{3} - \frac{2a^3b^2}{3} + \frac{a^3b}{3} - \frac{2a^2b^3}{3} + \frac{3a^2b^2}{2} - a^2b + \frac{ab^3}{3} - ab^2 + ab.$$

This result is the same equation with lower bound of $\mathrm{Var}(X+Y)$ when using the improved Fréchet lower bound (Equation (2.35)). □

Proposition 2.4.3. *(Maximum Variance) Given Figure 2.17 and assumptions in Proposition 2.4.1, take C as the independence copula, then*

$$\mathrm{Var}(\mathbb{I}_{\mathcal{F}}(X+Y)+\sum_{i=1}^{3}\mathbb{I}_{\mathcal{U}_i}(G_{2i-1}^{-1}(U)+G_{2i}^{-1}(U)) = \frac{2a^3b^3}{3}-\frac{4a^3b^2}{3}+\frac{2a^3b}{3}-\frac{4a^2b^3}{3}+\frac{5a^2b^2}{2}-a^2b+\frac{2ab^3}{3}-ab^2+\frac{1}{3},$$

which is the same as Equation (2.59)*, where* $U \sim U[0,1]$.

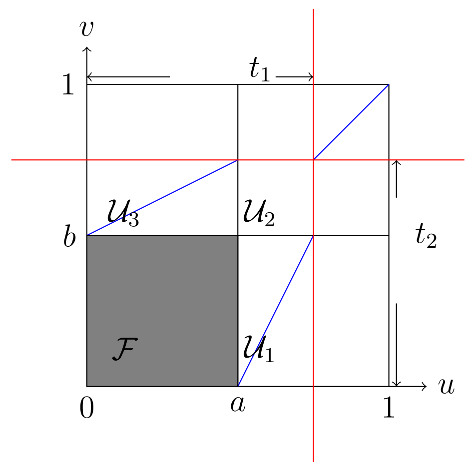


Figure 2.17: When $\mathcal{F} = [0,a] \times [0,b]$

Proof. Based on Figure 2.17, note here G_i is different from the G_i in the proof of Proposition 2.4.2.

$$G_1(x) = F_{X|(X,Y)\in\mathcal{U}_1}(x) = \frac{P(X \le x, a \le X \le t_1, Y \le b)}{P(a \le X, Y \le b)} = \frac{x-a}{b-ab} \text{ if } a \le x \le t_1. \quad (2.149)$$

Since $G_1(t_1)$ must be 1, we get $t_1 = a+b-ab$. So $G_1^{-1}(u) = u(b-ab)+a$.

Similarly,

$$G_2(x) = F_{Y|(X,Y)\in\mathcal{U}_1}(x) = \frac{P(Y \le x, a \le X \le t_1, Y \le b)}{P(a \le X, Y \le b)} = \frac{x - ax}{b - ab} = \frac{x}{b} \text{ if } x \le b, \quad (2.150)$$

so $G_2^{-1}(u) = ub$.

$$G_3(x) = F_{X|(X,Y)\in\mathcal{U}_2}(x) = \frac{P(X \le x, t_1 \le X, Y \ge t_2)}{P(a \le X, b \le Y)} = \frac{x - a - b + ab}{1 - a - b + ab}, \quad (2.151)$$

so $G_3^{-1}(u) = u(1 - a - b + ab) + a + b - ab$.

$$G_4(x) = F_{Y|(X,Y)\in\mathcal{U}_2}(x) = \frac{x - a - b + ab}{1 - a - b + ab}, \quad (2.152)$$

so $G_4^{-1}(u) = u(1 - a - b + ab) + a + b - ab$.

$$G_5(x) = F_{X|(X,Y)\in\mathcal{U}_3}(x) = \frac{P(X \le x, X \le a, b \le Y \le t_2)}{P(X \le a, b \le Y)} = \frac{x - bx}{a - ab} = \frac{x}{a}, \quad (2.153)$$

so $G_5^{-1}(u) = au$.

$$G_6(x) = F_{Y|(X,Y)\in\mathcal{U}_3}(x) = \frac{x - b}{a - ab}, \quad (2.154)$$

so $G_6^{-1}(u) = u(a - ab) + b$.
For convenience, denote $A_i := G_i^{-1}(u), i = 1, ..., 6$.
Define another random variable T such that,

$$T = \begin{cases} 1 \text{ if } \mathbb{I}_{\mathcal{F}} = 1 \\ 2 \text{ if } \mathbb{I}_{\mathcal{U}_1} = 1 \\ 3 \text{ if } \mathbb{I}_{\mathcal{U}_2} = 1 \\ 4 \text{ if } \mathbb{I}_{\mathcal{U}_3} = 1. \end{cases} \quad (2.155)$$

Denote $Z := \mathbb{I}_{\mathcal{F}}(X+Y) + \mathbb{I}_{\mathcal{U}_1}(G_1^{-1}(U) + G_2^{-1}(U)) + \mathbb{I}_{\mathcal{U}_2}(G_3^{-1}(U) + G_4^{-1}(U)) + \mathbb{I}_{\mathcal{U}_3}(G_5^{-1}(U) + G_6^{-1}(U))$. Then $\text{Var}(Z) = E(\text{Var}(Z|T)) + \text{Var}(E(Z|T))$.

And

$$E(\text{Var}(Z|T)) = \sum_{i=1}^{4} P(T=i)\text{Var}(Z|T=i)$$
$$= ab\text{Var}(X+Y|(X,Y)\in\mathcal{F}) + (b-ab)\text{Var}(A_1+A_2|(X,Y)\in\mathcal{U}_1)$$
$$+ (1-a-b+ab)\text{Var}(A_3+A_4|(X,Y)\in\mathcal{U}_2) + (a-ab)\text{Var}(A_5+A_6|(X,Y)\in\mathcal{U}_3)$$
$$= \frac{a^3b^3}{6} - \frac{7a^3b^2}{12} + \frac{5a^3b}{12} - \frac{7a^2b^3}{12} + 3a^2b^2 - 3a^2b + a^2 + \frac{5ab^3}{12} - 3ab^2 + 3ab - a + b^2 - b + \frac{1}{3}, \tag{2.156}$$

$$\text{Var}(E(Z|T)) = \sum_{i=1}^{4} P(T=i)E^2(Z|T=i) - E^2(Z)$$
$$= \frac{a^3b^3}{2} - \frac{3a^3b^2}{4} + \frac{a^3b}{4} - \frac{3a^2b^3}{4} - \frac{a^2b^2}{2} + 2a^2b - a^2 + \frac{ab^3}{4} + 2ab^2 - 3ab + a - b^2 + b. \tag{2.157}$$

So

$$\text{Var}(Z) = \frac{2a^3b^3}{3} - \frac{4a^3b^2}{3} + \frac{2a^3b}{3} - \frac{4a^2b^3}{3} + \frac{5a^2b^2}{2} - a^2b + \frac{2ab^3}{3} - ab^2 + \frac{1}{3}.$$

This result is the same equation with upper bound of $\text{Var}(X+Y)$ when using the improved Fréchet upper bound (Equation (2.59)). □

Corollary 2.4.4. *(Improved variance bounds) Given Figure 2.13, assuming 1,2,3 in Proposition 2.4.1, if we take C as the independence copula, then upper and lower bounds of $\text{Var}(X+Y)$ are as follows,*

$$\frac{a^3b^3}{3} - \frac{2a^3b^2}{3} + \frac{a^3b}{3} - \frac{2a^2b^3}{3} + \frac{3a^2b^2}{2} - a^2b + \frac{ab^3}{3} - ab^2 + ab$$
$$\leq \text{Var}(X+Y) \leq$$
$$\frac{2a^3b^3}{3} - \frac{4a^3b^2}{3} + \frac{2a^3b}{3} - \frac{4a^2b^3}{3} + \frac{5a^2b^2}{2} - a^2b + \frac{2ab^3}{3} - ab^2 + \frac{1}{3}.$$

Proof. This directly follows from Propositions 2.4.1, 2.4.2 and 2.4.3. □

2.5 Relation between Improved Bounds and Improved Fréchet Bounds

Conjecture 2.5.1. *Assume (X, Y) has known copula on area $\mathcal{F} \subset [0,1]^2$ and unknown dependence structure on area $[0,1]^2 \backslash \mathcal{F}$, then improved bounds on variance using convex order bounds (see Proposition 2.3.1) are sharp if and only if improved Fréchet bounds are sharp (i.e., $A^{\mathcal{F},C}$ and $B^{\mathcal{F},C}$ are copulas.)*

Using Conjecture 2.5.1 and 2.2.7, we give an example to illustrate how to produce the improved upper bound using convex order bounds.

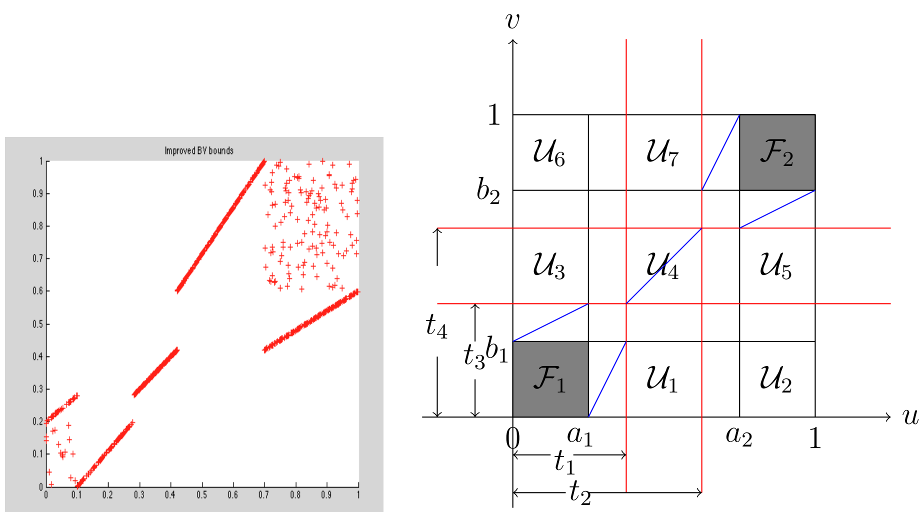


Figure 2.18: Comparisons of simulation and sketch of improved bounds using convex order, $a_1 = 0.1, b_1 = 0.2, a_2 = 0.7, b_2 = 0.6$.

Remark 2.5.2. One difficult part of using the improved bounds to get variance bounds is that the right way to distribute the probability masses on the unknown area $\mathcal{U}$. One possible solution includes treating the problem as a linear programming problem. By Corollary 2.1.2, maximizing or minimizing $\text{Var}(X+Y)$ is equivalent to maximizing or minimizing $E[XY]$. Fix some large number n, we approximate X and Y by discrete distributions as follows: $X = x_i := \frac{i}{n}, 1 \leq i \leq n$ where $p_i := P(X = x_i) = \frac{1}{n}$; $Y = y_i := \frac{i}{n}, 1 \leq i \leq n$ where

$q_i := P(Y = y_i) = \frac{1}{n}$. Denote $p_{i,j} := P(X = x_i, Y = y_j)$. Then $E[XY] = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{i,j}x_i y_j = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{i,j}\frac{ij}{n^2}$. So we $\max_{p_{i,j}} \sum_{i=1}^{n}\sum_{j=1}^{n} p_{i,j}\frac{ij}{n^2}$ or $\min_{p_{i,j}} \sum_{i=1}^{n}\sum_{j=1}^{n} p_{i,j}\frac{ij}{n^2}$ subject to

$$\begin{cases} 0 \le p_{i,j} \le 1 \\ \sum_{i=1}^{n}\sum_{j=1}^{n} p_{i,j} = 1 \\ \forall i, \sum_{j=1}^{n} p_{i,j} = p_i = \frac{1}{n} \\ \forall j, \sum_{i=1}^{n} p_{i,j} = p_j = \frac{1}{n} \\ p_{i,j} = p_i q_j = \frac{1}{n^2} \text{ when } (x_i, y_j) \in \mathcal{F}. \end{cases}$$

Example 2.5.3. Given $\mathcal{F}_1 = [0, a_1] \times [0, b_1], \mathcal{F}_2 = [a_2, 1] \times [b_2, 1]$ where $a_1 < a_2, b_1 < b_2, a_2 b_2 > a_1 + b_1 - a_1 b_1$ and Figure 2.18, assume (X, Y) has independence copula on area $\mathcal{F}_1, \mathcal{F}_2$. To produce the improved bounds on variance, we first plot the simulation of the copula using Tankov's method, then split the $[0, 1]^2$ rectangles based on the simulation. Denote $G_{2i-1}(x) = F_{X|(X,Y)\in\mathcal{U}_i}(x), G_{2i}(x) = F_{Y|(X,Y)\in\mathcal{U}_i}(x), i = 1, 2, ..., 7$, then improved upper bound on variance is defined as

$$\mathrm{Var}\left(\mathbb{I}_{\mathcal{F}_1}(X+Y) + \mathbb{I}_{\mathcal{F}_2}(X+Y) + \sum_{i=1}^{7} \mathbb{I}_{\mathcal{U}_i}(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(U))\right)$$

where U is $U[0, 1]$.
By Conjecture 2.2.7, when $a_2 b_2 > a_1 + b_1 - a_1 b_1$ holds, $A^{\mathcal{F},C}$ is a copula. So by Conjecture 2.5.1, improved bounds are sharp and the plot is well-defined based on the condition $a_2 b_2 > a_1 + b_1 - a_1 b_1$.
Then we show how to get $G_i(x)$.

$$G_1(x) = F_{X|(X,Y)\in\mathcal{U}_1}(x) = \frac{P(X \le x, a_1 \le X \le t_1, Y \le b_1)}{P(a_1 \le X \le a_2, Y \le b_1)} = \frac{x - a_1}{b_1 - a_1 b_1}$$

for $a_1 \le x \le t_1$.
By $G_1(t_1) = 1$, we get $t_1 = a_1 + b_1 - a_1 b_1$. So $G_1^{-1}(u) = u(b_1 - a_1 b_1) + a_1$.

$$G_2(x) = F_{Y|(X,Y)\in\mathcal{U}_1}(x) = \frac{x}{b_1},$$

so $G_2^{-1}(u) = ub_1$.
Based on the plot of Figure 2.18, $G_3(x) = G_4(x) = G_{11}(x) = G_{12}(x) = 0$.

$$G_5(x) = F_{X|(X,Y)\in\mathcal{U}_3}(x) = \frac{P(X \le x, X \le a_1, b_1 \le Y \le t_3)}{P(X \le a_1, b_1 \le Y \le b_2)} = \frac{x}{a_1},$$

so $G_5^{-1}(u) = ua_1$.

$$G_6(x) = F_{Y|(X,Y)\in\mathcal{U}_3}(x) = \frac{x - b_1}{a_1 - a_1 b_1},$$

so $G_6^{-1}(u) = u(a_1 - a_1b_1) + b_1$. By $G_6(t_3) = 1$, we get $t_3 = a_1 + b_1 - a_1b_1$. Notice $t_3 = t_1$.

$$G_7(x) = F_{X|(X,Y)\in\mathcal{U}_4}(x) = \frac{P(t_1 \le X \le x, t_3 \le Y \le t_4)}{P(a_1 \le X \le a_2, b_1 \le Y \le b_2)} = \frac{x - t_1}{t_2 - t_1},$$

$$G_8(x) = F_{Y|(X,Y)\in\mathcal{U}_4}(x) = \frac{x - t_3}{t_2 - t_1}.$$

By $G_8(t_4) = 1$, we get $t_4 - t_3 = t_2 - t_1$. So $t_2 = t_4$.
We know (X, Y) are independent when $(X, Y) \in \mathcal{F}_2$,

$$P(X \ge a_2, Y \ge b_2) = P(X \ge a_2)P(Y \ge b_2) = 1 - a_2 - b_2 + a_2b_2,$$

$$G_9(x) = F_{X|(X,Y)\in\mathcal{U}_5}(x) = \frac{P(a_2 \le X \le x, t_4 \le Y \le b_2)}{P(a_2 \le X, b_1 \le Y \le b_2)} = \frac{x - a_2}{1 - a_2},$$

$$G_{10}(x) = F_{Y|(X,Y)\in\mathcal{U}_5}(x) = \frac{x - t_4}{b_2 - a_2b_2}.$$

By $G_{10}(b_2) = 1$, we get $t_4 = a_2b_2$.
So $G_7^{-1}(u) = u(t_2 - t_1) + t_1 = u(a_2b_2 - a_1 - b_1 + a_1b_1) + a_1 + b_1 - a_1b_1, G_8^{-1}(u) = u(a_2b_2 - a_1 - b_1 + a_1b_1) + a_1 + b_1 - a_1b_1$, $G_9^{-1}(u) = u(1 - a_2) + a_2, G_{10}^{-1}(u) = u(b_2 - a_2b_2) + a_2b_2$. From Figure 2.18, we need $t_3 < t_4$, which is $a_2b_2 > a_1 + b_1 - a_1b_1$, this is given in the proposition.

$$G_{13}(x) = F_{X|(X,Y)\in\mathcal{U}_7}(x) = \frac{P(t_2 \le X \le x, b_2 \le Y)}{P(b_2 \le X \le a_2, b_2 \le Y)} = \frac{x - a_2b_2}{a_2 - a_2b_2},$$

$$G_{14}(x) = F_{Y|(X,Y)\in\mathcal{U}_7}(x) = \frac{P(b_2 \le Y \le x, t_2 \le X \le a_2)}{a_2 - a_2b_2} = \frac{x - b_2}{1 - b_2}.$$

Then $G_{13}^{-1}(u) = u(a_2 - a_2b_2) + a_2b_2$, $G_{14}^{-1}(u) = u(1 - b_2) + b_2$.
For convenience, denote $A_i := G_i^{-1}(u)$ for $i = 1, 2, ...14$ except $3, 4, 11, 12$. Define another

random variable T such that

$$T = \begin{cases} 1 \text{ if } \mathbb{I}_{\mathcal{F}_1} = 1 \\ 2 \text{ if } \mathbb{I}_{\mathcal{F}_2} = 1 \\ 3 \text{ if } \mathbb{I}_{\mathcal{U}_1} = 1 \\ 4 \text{ if } \mathbb{I}_{\mathcal{U}_3} = 1 \\ 5 \text{ if } \mathbb{I}_{\mathcal{U}_4} = 1 \\ 6 \text{ if } \mathbb{I}_{\mathcal{U}_5} = 1 \\ 7 \text{ if } \mathbb{I}_{\mathcal{U}_7} = 1. \end{cases}$$

Denote $Z := \mathbb{I}_{\mathcal{F}_1}(X+Y) + \mathbb{I}_{\mathcal{F}_2}(X+Y) + \sum_{i=1}^{7} \mathbb{I}_{\mathcal{U}_i}(G_{2i-1}^{-1}(U) + G_{2i}^{-1}(U))$. Then $\text{Var}(Z) = E(\text{Var}(Z|T)) + \text{Var}(E(Z|T))$. $\text{Var}(Z)$ can be computed. The procedure is similar if we want improved lower bound, just replace $G_{2i}^{-1}(U)$ with $G_{2i}^{-1}(1-U)$.

2.6 Conclusion of Chapter 2 and Future Work

In this chapter, for any random variables X and Y, we give upper and lower bounds of $\text{Var}(X+Y)$ under the condition that X and Y are independent when $(X,Y) \in [0,a] \times [0,b], a, b \in (0,1)$. Three methods are used: 1. the use of copula bounds from Tankov [2011] (see Section 2.2), 2. the use of bounds from Bernard and Vanduffel [2014] (see Section 2.3) 3. the use of improved bounds (see Section 2.4). Method 1 and 3 gives the same result and they are sharp bounds, while method 2 does not give sharp upper bound. The computation of improved bounds is much simpler than the use of copula bounds from Tankov [2011]. We give an example in Section 2.5.3 on how to calculate it.

There are many open questions remaining in this chapter:

1. Several Conjectures 2.2.7, 2.2.8, 2.5.1 are not proved. Conjectures 2.2.7 and 2.2.8 give weaker sufficient conditions than in Bernard et al. [2013a] for $A^{\mathcal{F},C}, B^{\mathcal{F},C}$ to be copulas.

2. We illustrate with very special examples, assuming (X,Y) have the independence copula when $(X,Y) \in [0,a] \times [0,b]$. More complex examples can be done.

3. We do not find necessary conditions for improved Fréchet bounds to be copulas. If Conjecture 2.5.1 is proved, improved bounds can be investigated as a way to give the necessary condition.

4. The lower variance bound using Bernard and Vanduffel [2014] is sharp while the upper bound is not. So what is the sufficient condition for bound using Bernard and Vanduffel [2014] to be sharp?

5. In Example 2.5.3, we see that the right way to split the rectangle and distribute the masses in Figure 2.18 is a key step to get the improved bounds. What will be a valid split for the bounds to be sharp? When the improved Fréchet bounds are not sharp, does there still exist a split that gets sharp bounds?

6. In Figure 2.11, the simulation plot in Panel A gives a copula while the simulation plot in Panel B does not. This looks quite like a shuffle of the copula (see Mikusinski et al. [1992], Durante et al. [2009], Durante and Fernández-Sánchez [2010], Durante and Sánchez [2012], Trutschnig and Fernández Sánchez [2013], Ruankong et al. [2013] for detail on shuffles) but the slopes may not be $+1, -1$. Can we find sufficient conditions to ensure that it is a copula?

7. We only study bounds on variance, all the methods in Sections 2.2, 2.3, 2.4 can be used to study bounds on other convex risk measures.

2.7 Appendix to Chapter 2

2.7.1 Checking Equations (2.115) and (2.134)

Since our computation in Section 2.3 is quite long, we verify that our Equations (2.115) and (2.134) are right with the following 4 steps:

- Step 1
  Simulate n pairs of uniform distributed random variables (X, Y) which have copula $B^{\mathcal{F},C}$. Fix some a, b, check $G_1(x)$ and $G_2(x)$ empirically and plot them against the theoretical result: Equations (2.86) and (2.87). To approximate $G_1(x)$ empirically, as it is defined as (2.74), if we want to approximate $P((X,Y) \in \mathcal{U})$, since $P((X,Y) \in \mathcal{U}) = P(X > a \quad or \quad Y > b)$,

$$\frac{\sum_{i=1}^{n} \mathbb{I}_{\{X_i > a \quad or \quad Y_i > b\}}}{n} \longrightarrow P(X > a \quad or \quad Y > b) \quad \text{when} \quad n \to \infty, \tag{2.158}$$

by the law of large numbers. Figures 2.19, 2.20, 2.21, 2.22 are comparisons of the empirical G_1, G_2 and theoretical G_1, G_2 when $n = 1000$.

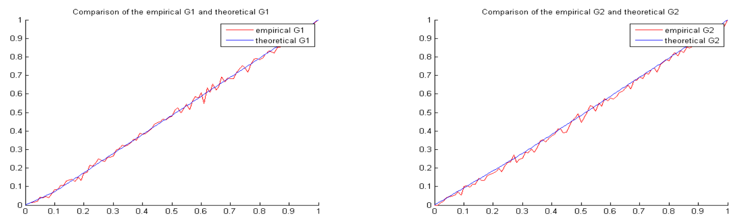


Figure 2.19: Comparison of empirical G_1, G_2 and theoretical G_1, G_2 when $a = 0.1, b = 0.3$.

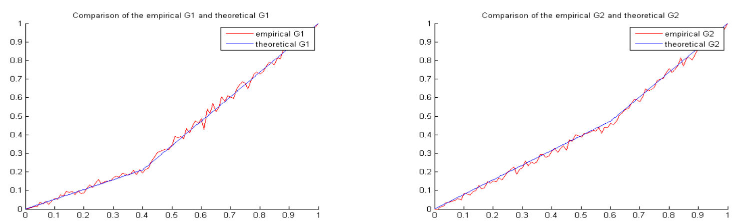


Figure 2.20: Comparison of empirical G_1, G_2 and theoretical G_1, G_2 when $a = 0.4, b = 0.6$.

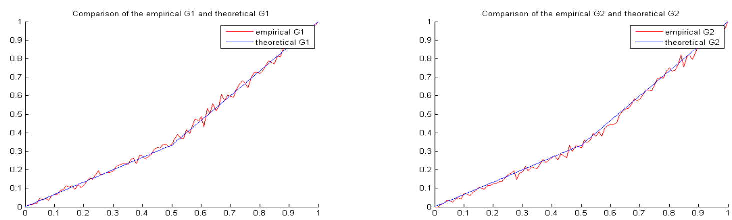


Figure 2.21: Comparison of empirical G_1, G_2 and theoretical G_1, G_2 when $a = 0.5, b = 0.5$.

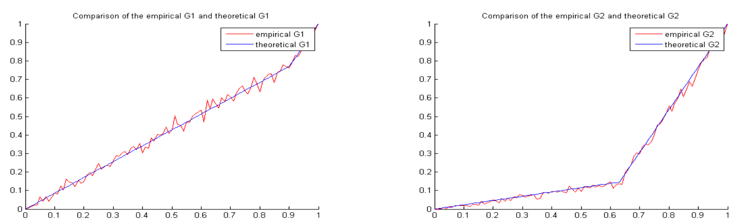


Figure 2.22: Comparison of empirical G_1, G_2 and theoretical G_1, G_2 when $a = 0.9, b = 0.63$.

Plots in Figure 2.19, 2.20, 2.21, 2.22 look right in terms of Equations (2.86) and (2.87).

- Step 2

  To check G_1^{-1} and G_2^{-1}, we fix some a, b and take in some $y = G_1(x), G_2(x)$ to Equations (2.88), (2.89). Since G_1, G_2 are strictly increasing, the right result should give:

$$G_1^{-1}(G_1(x)) = x \qquad \text{and} \qquad G_2^{-1}(G_2(x)) = x. \tag{2.159}$$

  So we pick the same 4 pairs of a, b as above and plot $(x, G_1^{-1}(G_1(x)))$, $(x, G_2^{-1}(G_2(x)))$ against $y = x$. All the four pairs give same graphs as in Figure 2.23,

  So Equations (2.88) and (2.89) are right.

- Step 3

  To check Equation (2.100), we simulate n uniformly distributed numbers u and cal-

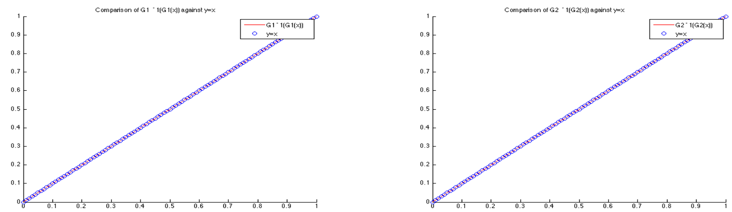


Figure 2.23: Comparison of G_1^{-1}, G_2^{-1} and $y = x$.

culate $\mathrm{Var}(G_1^{-1}(u) + G_2^{-1}(1-u))$, the result is shown in Table 2.3. Here $n = 1000$, the theoretical result refers to Equation (2.100).

a	b	empirical result	theoretical result	error
0.1	0.3	0.0001	0.0001	0.0000
0.4	0.6	0. 0064	0.0064	0.0000
0.5	0.5	0.0062	0.0069	0.0007
0.9	0.63	0.0313	0.0308	0.0005

Table 2.3: Error between the empirical result and the theoretical result of $\mathrm{Var}(G_1^{-1}(U) + G_2^{-1}(U))$

In a similar way, we check Equation (2.114) in Table 2.4.

a	b	empirical result	theoretical result	error
0.1	0.3	0.3269	0.3224	0.0046
0.4	0.6	0.3030	0.3014	0.0016
0.5	0.5	0.2992	0.3056	0.0063
0.9	0.63	0.2549	0.2614	0.0064

Table 2.4: Error between the empirical result and the theoretical result of $\mathrm{Var}(G_1^{-1}(U) + G_2^{-1}(1-U))$

- Step 4
  To check Equation (2.115), which is Equation (2.116), we take n pairs of (X, Y) in step 1 and n uniformly distributed random variables u to get

$$\mathrm{Var}(\mathbb{I}(X+Y)+(1-\mathbb{I})(G_1^{-1}(u)+G_2^{-1}(1-u)),$$

the result is shown in Table 2.5. Here $n = 1000$, the theoretical result refers to Equation (2.115).

a	b	empirical result	theoretical result	error
0.1	0.3	0.0113	0.0201	0.0089
0.4	0.6	0.0897	0.0942	0.0046
0.5	0.5	0.0987	0.0990	0.0002
0.9	0.63	0.1458	0.1427	0.0031

Table 2.5: Error between the empirical result and the theoretical result of $\mathrm{Var}(\mathbb{I}(X+Y)+(1-\mathbb{I})(G_1^{-1}(U)+G_2^{-1}(1-U))$

Then we check Equation (2.128), $\mathrm{Var}(\mathbb{I}(X+Y)+(1-\mathbb{I})(B_1+B_2))$ in Table 2.6, the theoretical result refers to Equation (2.136). Equation (2.136) looks good, however,

a	b	empirical result	theoretical result	error
0.1	0.3	0.3358	0.3327	0.0031
0.4	0.6	0.3227	0.3184	0.0043
0.5	0.5	0.3195	0.3229	0.0034
0.9	0.63	0.2399	0.2425	0.0026

Table 2.6: Error between the empirical result and the theoretical result of $\mathrm{Var}(\mathbb{I}(X+Y)+(1-\mathbb{I})(B_1+B_2))$

this does not coincide with Equation (2.59) (which calculated using improved Fréchet upper bound). Table 2.7 is a comparison between Equations (2.136) and (2.59), which refer to theoretical result of $\mathrm{Var}(\mathbb{I}(X+Y)+(1-\mathbb{I})(B_1+B_2))$ and upper variance bound calculated using improved Fréchet upper bound.

a	b	(2.136)	(2.59)	error
0.1	0.3	0.3327	0.3251	0.0076
0.4	0.6	0.3184	0.2529	0.0655
0.5	0.5	0.3229	0.2500	0.0729
0.9	0.63	0.2425	0.1914	0.0511

Table 2.7: Comparison between Equations (2.136) and (2.59)

Table 2.8 contains both empirical results of Equations (2.136) and (2.59), denoted as empirical1 and empirical2 respectively, $n = 1000$.

a	b	empirical1	empirical2	error
0.1	0.3	0.3263	0.3335	0.0072
0.4	0.6	0.3145	0.2407	0.0738
0.5	0.5	0.3282	0.2563	0.0720
0.9	0.63	0.2462	0.1836	0.0626

Table 2.8: Empirical results of Equations (2.136) and (2.59) when $n = 1000$.

Table 2.9 contains the above comparison when $n = 5000$.

a	b	empirical1	empirical2	error
0.1	0.3	0.3290	0.3200	0.0090
0.4	0.6	0.3120	0.2584	0.0537
0.5	0.5	0.3271	0.2442	0.0830
0.9	0.63	0.2470	0.1896	0.0574

Table 2.9: Empirical results of Equations (2.136) and (2.59) when $n = 5000$

2.7.2 Checking Bounds Using Bernard and Vanduffel [2014] by Deriving Copula Directly

This section is to check the copula of upper variance bound using Bernard and Vanduffel [2014] is indeed the same with the copula (X, Y) in Equation (2.126).

Proposition 2.7.1. *If we start with assumptions 1,2,3,4 in Proposition 2.3.3 and take C as the independence copula, then the copula of (X,Y) in the upper bound of* $\mathrm{Var}(X+Y)$, $\mathrm{Var}(X+Y) = \mathrm{Var}(\mathbb{I}(X+Y) + (1-\mathbb{I})(G_1^{-1}(U) + G_2^{-1}(U)))$ *is*

$$C^{Bv}(x,y) := min(x,a)min(y,b) + min(G_1(x), G_2(y))(1-ab), x, y \in [0,1]$$

Proof. For $x, y \in [0,1]$, V_1, V_2 are independent $U[0,1]$ which are independent of $G_1^{-1}(U), G_2^{-1}(U)$

$$\begin{aligned}
P(X \le x, Y \le y) =& P(X \le x, Y \le y, \mathbb{I}=1) + P(X \le x, Y \le y, \mathbb{I}=0) \\
=& P(V_1 \le x, V_2 \le y, V_1 \le a, V_2 \le b) + P(G_1^{-1}(U) \le x, G_2^{-1}(U) \le y, V_1 \ge a, V_2 \le b) \\
& + P(G_1^{-1}(U) \le x, G_2^{-1}(U) \le y, V_1 \ge a, V_2 \ge b) \\
& + P(G_1^{-1}(U) \le x, G_2^{-1}(U) \le y, V_1 \le a, V_2 \le b) \\
=& P(V_1 \le \min(x,a), V_2 \le \min(y,b)) \\
& + P(G_1^{-1}(U) \le x, G_2^{-1}(U) \le y)(P(V_1 \ge a, V_2 \le b) \\
& + (P(V_1 \ge a, V_2 \ge b) + (P(V_1 \le a, V_2 \ge b)) \\
=& \min(x,a)\min(y,b) + \min(G_1(x), G_2(y))(1-ab)
\end{aligned}$$

□

The comparisons of the simulation of C^{BV} and simulation of the copula of Equation (2.126) are in Figures 2.24 and 2.25.

2.7.3 Another Approach for Comparing (X,Y) Simulated from Bounds Using Bernard and Vanduffel [2014] and Tankov [2011]

First, we start with assumptions in Proposition 2.3.3 and take C as the independence copula. We simulate (X,Y) from upper bounds in Equation (2.67) as follows: we simulate n uniformly distributed random variables U. If $U \le p_f = ab$, simulate (X,Y) following

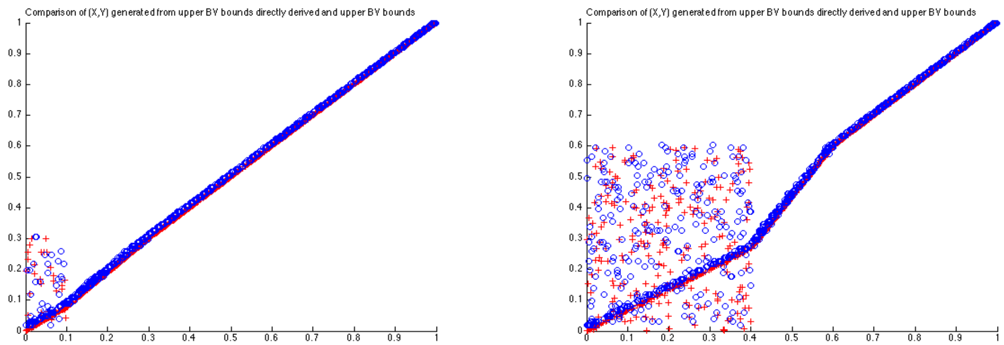


Figure 2.24: Comparisons of copula of the upper variance bound using Bernard and Vanduffel [2014] derived directly and by simulation, where a=0.1, b=0.3 (left), a=0.4, b=0.6 (right).

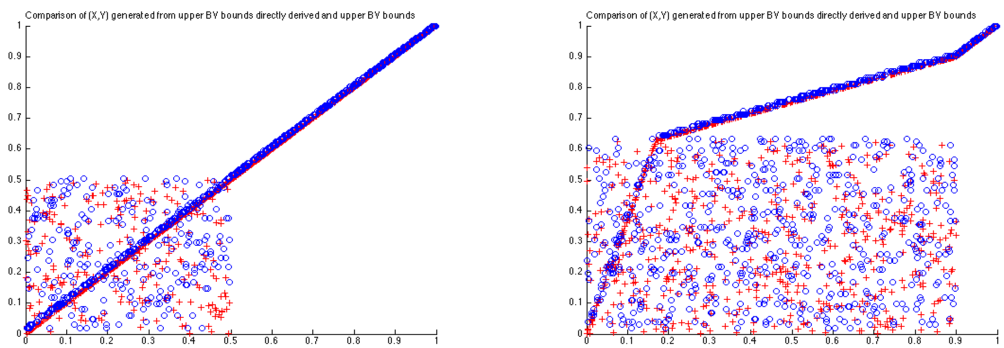


Figure 2.25: Comparisons of copula of the upper variance bound using Bernard and Vanduffel [2014] derived directly and by simulation, where a=0.5, b=0.5 (left), a=0.9, b=0.63 (right).

the given copula (independence copula here). Denote

$$\begin{aligned} H_1(x) :=& P(X \le x | (X,Y) \in \mathcal{F}) = \frac{P(X \le x, X \le a, Y \le b)}{P((X,Y) \in \mathcal{F})} \\ =& \begin{cases} \frac{P(X \le x, Y \le b)}{p_f} & \text{if } x \le a \\ \frac{P(X \le a, Y \le b)}{p_f} & \text{if } x > a \end{cases} \\ =& \begin{cases} \frac{bx}{ab} = \frac{x}{a} & \text{if } x \le a \\ \frac{ab}{ab} = 1 & \text{if } x > a. \end{cases} \end{aligned}$$

Similarly,

$$H_2(y) = \begin{cases} \frac{y}{b} & \text{if } y \leq b \\ 1 & \text{if } y > b. \end{cases}$$

We simulate independent $V_1, V_2 \sim U[0,1]$, let $X = H_1^{-1}(V_1), Y = H_2^{-1}(V_2)$; if $U > p_f = ab$, we simulate $(X,Y) = (G_1^{-1}(U), G_2^{-1}(U))$. Then plot (X,Y) on the graph.
Second, we plot the copula of (X,Y) from upper bound derived using improved Fréchet bound, which is Equation (2.37).
Similarly, we plot (X,Y) from lower bounds with $n = 1000$. In Figures 2.26, 2.27, 2.28, 2.29, the left figure is the lower bound and the right is the upper bound. For the upper bound, (X,Y) do not coincide.

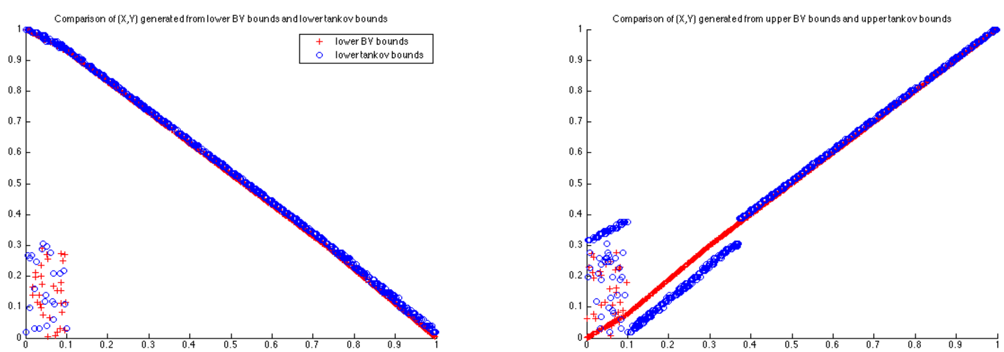


Figure 2.26: Comparison of (X,Y) generated from bounds using Bernard and Vanduffel [2014] and improved Fréchet bounds, where a=0.1, b=0.3.

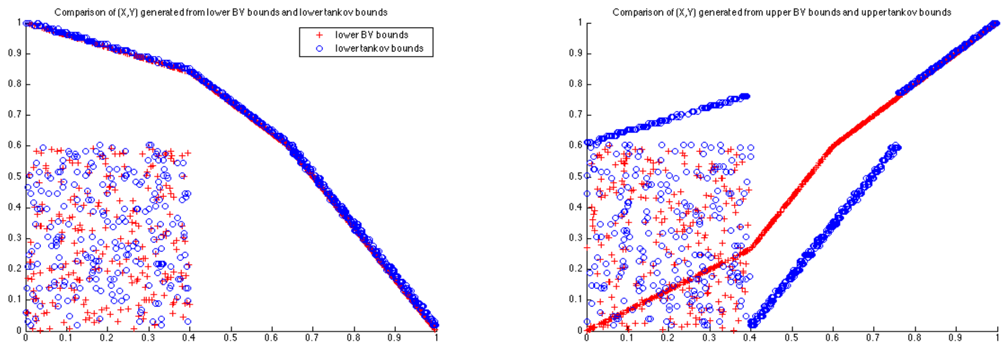


Figure 2.27: Comparison of (X, Y) generated from bounds using Bernard and Vanduffel [2014] and improved Fréchet bounds, where a=0.4, b=0.6.

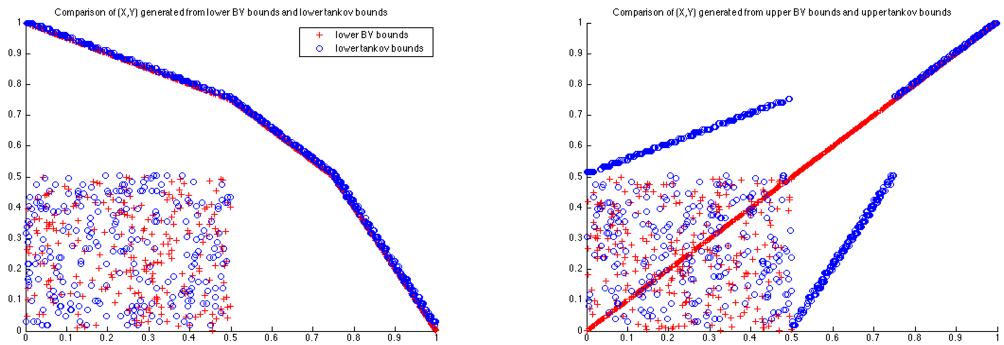


Figure 2.28: Comparison of (X, Y) generated from bounds using Bernard and Vanduffel [2014] and improved Fréchet bounds, where a=0.5, b=0.5.

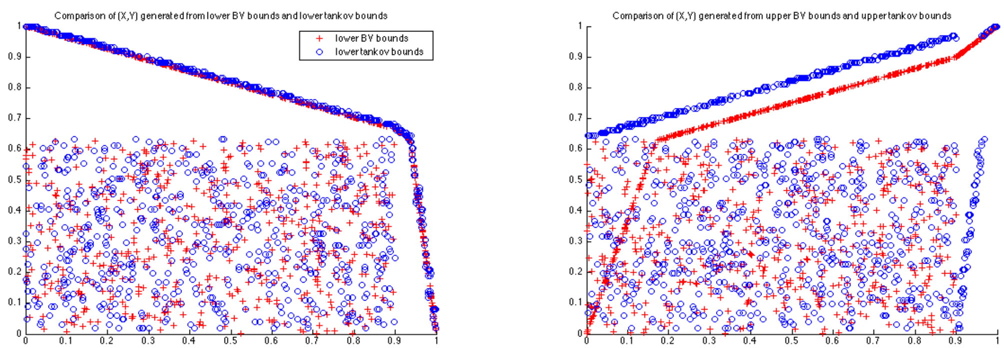
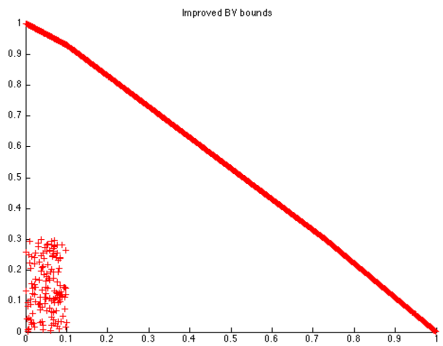


Figure 2.29: Comparison of (X, Y) generated from bounds using Bernard and Vanduffel [2014] and improved Fréchet bounds, where a=0.9, b=0.63.

2.7.4 Simulations of Improved Bounds Using Convex Order

Figures 2.30, 2.31 and 2.32, 2.33 include the simulations of lower and upper improved bounds from Section 2.5 under different pairs of a, b. They coincide with simulations of variance bounds derived using improved Fréchet bounds.



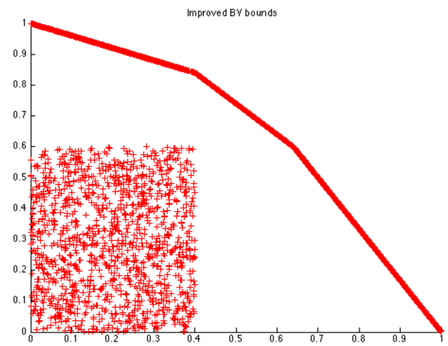
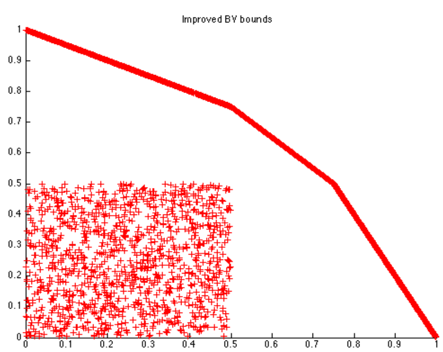


Figure 2.30: Simulation of lower improved bounds, a=0.1, b=0.3 (left), a=0.4, b=0.6 (right).



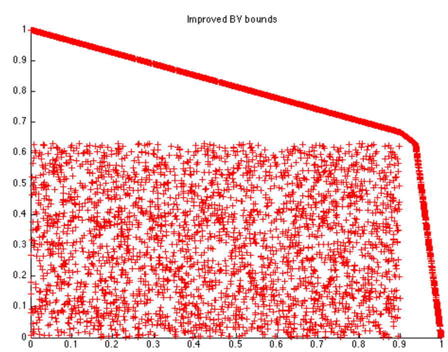


Figure 2.31: Simulation of lower improved bounds, a=0.5, b=0.5 (left), a=0.9, b=0.63 (right).

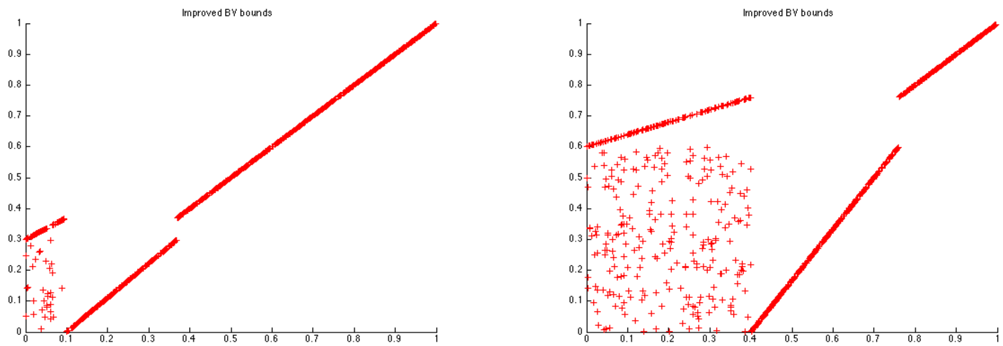


Figure 2.32: Simulation of upper improved bounds, a=0.1, b=0.3 (left), a=0.4, b=0.6 (right).

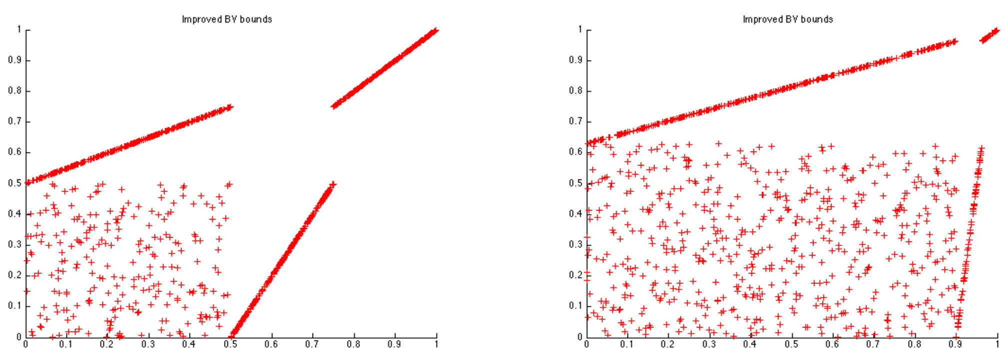


Figure 2.33: Simulation of upper improved bounds, a=0.5, b=0.5 (left), a=0.9, b=0.63 (right).

Chapter 3

Bounds on Variance with Background Risk

This chapter is organized as follows. In Section 3.1, we give the covariance matrix for $\mathrm{Var}(\sum_{i=1}^{n} X_i)$ to reach its minimum where each X_i is normal distributed. Sections 3.1.1 and 3.1.2 deal with cases $n = 3$ and 4 respectively. We apply our result in Section 3.2 by deriving sharp bounds for the variance of the sum $S = \sum_{i=1}^{n} X_i$ when $n = 3$ and 4. Every X_i is normal distributed and we have additional information on a given risk factor Z such that the distribution of (X_i, Z) is given. See Bernard et al. [2014] for more details. In Section 3.3, rearrangement algorithm from Puccetti and Rüschendorf [2012] is introduced to approximate the minimum of $\mathrm{Var}(X_1 + ... + X_n)$ with the existence of a background risk Z. Two examples on Pareto risks are presented. Section 3.4 gives a short conclusion and future research directions of Chapter 3.

3.1 Dependence among Normal Variables Such That $\mathrm{Var}(\sum_{i=1}^{d} X_i) = 0$

Given random variables $X_i \sim F_i, 1 \leq i \leq d$, what is the covariance matrix for $\mathrm{Var}(\sum_{i=1}^{d} X_i)$ reaches minimum? In this section, this question is answered when $d = 3, 4$ with assumptions that F_i is normal distribution $N(\mu_i, \sigma_i^2)$ where $\sigma_i > 0$. We introduce several useful definitions and propositions first, more details can be found in Horn and Johnson [2012].

Definition 3.1.1. A real symmetric $d \times d$ matrix M is called *positive-semidefinite* if

$$x^T M x \geq 0$$

for all x in $\mathbb{R}^d$.

Definition 3.1.2. A matrix M is *diagonally dominant* if

$$|a_{ii}| \geq \Sigma_{j \neq i} |a_{ij}|, \forall i,$$

where a_{ij} denotes the entry in the ith row and jth column.

Lemma 3.1.3. *A real symmetric diagonally dominant matrix M with non-negative diagonal entries is positive-semidefinite.*

Proposition 3.1.4. *If $X_i \sim N(\mu_i, \sigma_i^2), i = 1, ..., d, d \geq 3$, the sufficient condition for a matrix M to be a covariance matrix for $(X_i)_{1 \leq i \leq d}$ is:*
(i) M is a real symmetric, positive-semidefinite matrix with diagonal element $m_{i,i} = \sigma_i^2$;
(ii) each element of M satisfies $-1 \leq \frac{m_{i,j}}{\sigma_i \sigma_j} \leq 1$.

Remark 3.1.5. Condition (ii) in Proposition 3.1.4 is not necessary. We can get condition (ii) by using the positive-semidefinite in condition (i). For $\forall i, j$, by Definition 3.1.1, take $x = [0...0...\frac{1}{\sigma_i}....\frac{1}{\sigma_j}...0...0]^T$. Since $x^T M x \geq 0$, we get $m_{i,j} \geq -1$. Similarly, take $x = [0...0... -\frac{1}{\sigma_i}....\frac{1}{\sigma_j}...0...0]^T$ to get $m_{i,j} \leq 1$. For convenience, we leave condition (ii) here.

3.1.1 Case of 3 Normal Distributed Random Variables

We start with a special case.

Proposition 3.1.6. *Assume that $X_1, X_2, X_3 \sim N(0, 1)$, the covariance matrix for* $\mathrm{Var}(X_1 + X_2 + X_3) = 0$ *is*

$$\begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}$$

Proof. By Lemma 3.1.3, this matrix is positive-semidefinite and it clearly satisfies conditions (i), (ii) in Proposition 3.1.4. Then $\mathrm{Var}(X_1 + X_2 + X_3) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \mathrm{Var}(X_3) + 2\mathrm{Cov}(X_1, X_2) + 2\mathrm{Cov}(X_1, X_3) + 2\mathrm{Cov}(X_2, X_3) = 1 + 1 + 1 - 2 \cdot \frac{1}{2} - 2 \cdot \frac{1}{2} - 2 \cdot \frac{1}{2} = 0.$ □

Proposition 3.1.7 is the generalization of Proposition 3.1.6.

Proposition 3.1.7. *Assume random variables* $X_i \sim N(\mu_i, \sigma_i^2)$*,* $i = 1, 2, 3$*,*
(i) when $\max\limits_{1\le i\le 3} \sigma_i \le \dfrac{1}{2}\Sigma_{i=1}^3 \sigma_i$*, then the covariance matrix for* $\mathrm{Var}(X_1 + X_2 + X_3) = 0$ *is*

$$\begin{bmatrix} \sigma_1^2 & \frac{\sigma_3^2-\sigma_1^2-\sigma_2^2}{2} & \frac{\sigma_2^2-\sigma_1^2-\sigma_3^2}{2} \\ \frac{\sigma_3^2-\sigma_1^2-\sigma_2^2}{2} & \sigma_2^2 & \frac{\sigma_1^2-\sigma_2^2-\sigma_3^2}{2} \\ \frac{\sigma_2^2-\sigma_1^2-\sigma_3^2}{2} & \frac{\sigma_1^2-\sigma_2^2-\sigma_3^2}{2} & \sigma_3^2 \end{bmatrix}$$

(ii) when $\max\limits_{1\le i\le 3} \sigma_i > \dfrac{1}{2}\Sigma_{i=1}^3 \sigma_i$*, WLOG, assume* $\max\limits_{1\le i\le 3} \sigma_i = \sigma_1$*, if the covariance matrix is*

$$\begin{bmatrix} \sigma_1^2 & -\sigma_1\sigma_2 & -\sigma_1\sigma_3 \\ -\sigma_1\sigma_2 & \sigma_2^2 & \sigma_2\sigma_3 \\ -\sigma_1\sigma_3 & \sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix}$$

then $\mathrm{Var}(X_1 + X_2 + X_3)$ *reaches its minimum:* $(\sigma_1 - \frac{1}{2}\Sigma_{i=1}^3 \sigma_i)^2$.

Proof. (i) When $\max\limits_{1\le i\le n} \sigma_i \le \dfrac{1}{2}\Sigma_{i=1}^n \sigma_i$, X_1, X_2, X_3 are jointly mixable (see Theorem 1.2.10). Then $X_1 + X_2 + X_3 = C$ a.s. for some $C \in \mathbb{R}$.
So

$$\begin{aligned} X_1 =& C - X_2 - X_3 \qquad a.s. \\ \mathrm{Var}(X_1) =& \mathrm{Var}(C - X_2 - X_3) \\ \sigma_1^2 =& \sigma_2^2 + 2\mathrm{Cov}(X_2, X_3) + \sigma_3^2 \\ \mathrm{Cov}(X_2, X_3) =& \frac{\sigma_1^2 - \sigma_2^2 - \sigma_3^2}{2}. \end{aligned}$$

Similarly, $\mathrm{Cov}(X_1, X_2) = \frac{\sigma_3^2-\sigma_1^2-\sigma_2^2}{2}$, $\mathrm{Cov}(X_1, X_3) = \frac{\sigma_2^2-\sigma_1^2-\sigma_3^2}{2}$. We do not need to check this is a covariance matrix since this is the only possible solution.
(ii)
This proof follows from the proof in Wang et al. [2013].
When $\sigma_1 > \sigma_2 + \sigma_3$,

$$\mathrm{Var}(X_1 + X_2 + X_3) \ge (\sigma_1 - \sqrt{\mathrm{Var}(X_2 + X_3)})^2 \ge (\sigma_1 - (\sigma_2 + \sigma_3))^2$$

$\mathrm{Var}(X_1+X_2+X_3)$ reaches $(\sigma_1-(\sigma_2+\sigma_3))^2$ when taking $X_1=\sigma_1 Z+\mu_1$, $X_i=-\sigma_i Z+\mu_i, i=2,3$ where $Z\sim N(0,1)$. □

3.1.2 Case of 4 Normal Distributed Random Variables

The following is a generalization of Proposition 3.1.7 under case (i) in dimension $d=4$.

Proposition 3.1.8. *Assume* $X_i\sim N(\mu_i,\sigma_i^2)$, $\max_{1\le i\le 4}\sigma_i\le\frac{1}{2}\Sigma_{i=1}^4\sigma_i$, *the covariance matrix for* $\mathrm{Var}(\Sigma_{i=1}^4 X_i)=0$ *is*

$$\begin{bmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \rho_{1,3}\sigma_1\sigma_3 & -\sigma_1(\sigma_1+\rho_{1,2}\sigma_2+\rho_{1,3}\sigma_3) \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & -\sigma_1\sigma_2(A+\rho_{1,2})-\rho_{1,3}\sigma_1\sigma_3 & \sigma_1\sigma_3(B+\rho_{1,3}) \\ \rho_{1,3}\sigma_1\sigma_3 & -\sigma_1\sigma_2(A+\rho_{1,2})-\rho_{1,3}\sigma_1\sigma_3 & \sigma_3^2 & \sigma_1\sigma_2(C+\rho_{1,2}) \\ -\sigma_1(\sigma_1+\rho_{1,2}\sigma_2+\rho_{1,3}\sigma_3) & \sigma_1\sigma_3(B+\rho_{1,3}) & \sigma_1\sigma_2(C+\rho_{1,2}) & \sigma_4^2 \end{bmatrix} \tag{3.1}$$

where

$$A=\frac{\sigma_1^2+\sigma_2^2+\sigma_3^2-\sigma_4^2}{2\sigma_1\sigma_2}, B=\frac{\sigma_1^2-\sigma_2^2+\sigma_3^2-\sigma_4^2}{2\sigma_1\sigma_3}, C=\frac{\sigma_1^2+\sigma_2^2-\sigma_3^2-\sigma_4^2}{2\sigma_1\sigma_2} \tag{3.2}$$

and where $\rho_{1,2},\rho_{1,3}$ *need to satisfy the following constraints:*
(i) When $\sigma_2+\sigma_3\ge\sigma_1+\sigma_4$,

$$\begin{pmatrix} \textit{Case 1} \\ \textit{Left panel} \\ \textit{of Figure 3.1} \end{pmatrix} \begin{cases} A_1(\rho_{1,2})\le\rho_{1,3}\le A_6, & \textit{if } A_7\le\rho_{1,2}\le\rho^{A_1,A_5} \\ A_5\le\rho_{1,3}\le A_6, & \textit{if } \rho^{A_1,A_5}\le\rho_{1,2}\le\rho^{A_2,A_6} \\ A_5\le\rho_{1,3}\le A_2(\rho_{1,2}), & \textit{if } \rho^{A_2,A_6}\le\rho_{1,2}\le A_8. \end{cases} \tag{3.3}$$

When $\sigma_2+\sigma_3\le\sigma_1+\sigma_4$,

$$\begin{pmatrix} \textit{Case 2} \\ \textit{Right panel} \\ \textit{of Figure 3.1} \end{pmatrix} \begin{cases} A_3(\rho_{1,2})\le\rho_{1,3}\le A_6, & \textit{if } -1<\rho_{1,2}\le\rho^{A_3,-1} \\ -1\le\rho_{1,3}\le A_6, & \textit{if } \rho^{A_3,-1}\le\rho_{1,2}\le\rho^{A_2,A_6} \\ -1\le\rho_{1,3}\le A_2(\rho_{1,2}), & \textit{if } \rho^{A_2,A_6}\le\rho_{1,2}\le A_8, \end{cases} \tag{3.4}$$

where

$$
\begin{aligned}
&A_1(\rho_{1,2}) = \frac{-\sigma_1 - \rho_{1,2}\sigma_2 - \sigma_4}{\sigma_3},\ A_2(\rho_{1,2}) = \frac{-\sigma_1 - \rho_{1,2}\sigma_2 + \sigma_4}{\sigma_3} \\
&A_3(\rho_{1,2}) = -(A + \rho_{1,2})\frac{\sigma_2}{\sigma_3} - \frac{\sigma_2}{\sigma_1},\ A_5 = -B - \frac{\sigma_2\sigma_4}{\sigma_1\sigma_3} \\
&A_6 = -B + \frac{\sigma_2\sigma_4}{\sigma_1\sigma_3},\ A_7 = -\frac{\sigma_3\sigma_4}{\sigma_1\sigma_2} - C \\
&A_8 = \frac{\sigma_3\sigma_4}{\sigma_1\sigma_2} - C,\ \rho^{A_1,A_5} = \frac{(\sigma_2 - \sigma_1)\sigma_4}{\sigma_1\sigma_2} - D \\
&\rho^{A_2,A_6} = \frac{(\sigma_1 - \sigma_2)\sigma_4}{\sigma_1\sigma_2} - D,\ \rho^{A_3,-1} = \frac{(\sigma_1 - \sigma_2)\sigma_3}{\sigma_1\sigma_2} - A
\end{aligned}
\tag{3.5}
$$

where A, B, C are defined in Equation (3.2) *and* $D = \frac{\sigma_1^2+\sigma_2^2-\sigma_3^2+\sigma_4^2}{2\sigma_1\sigma_2}$.

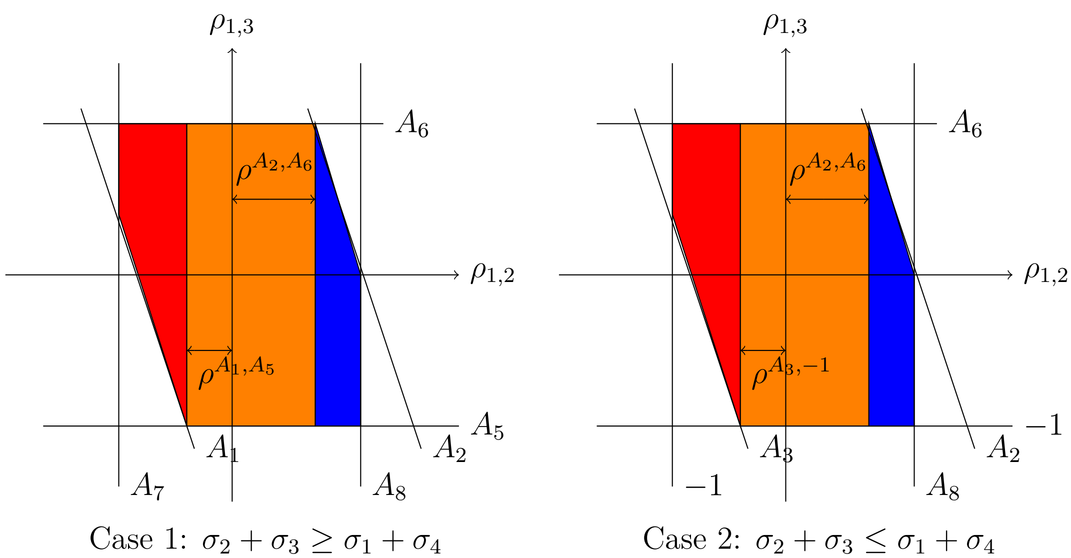


Figure 3.1: Illustrations of how to choose $\rho_{1,2}, \rho_{1,3}$

(ii) The sufficient condition for the covariance matrix to be positive-semidefinite, is $L_1 > 0$ *where*

$$
L_1 = (\rho_{1,2}^2 - 1)(\rho_{1,3}^2 - 1) - \left(\rho_{1,2}(\rho_{1,3} + \sigma_1\sigma_2) + \rho_{1,3}\sigma_1\sigma_3 + \sigma_1\sigma_2 A\right)^2. \tag{3.6}
$$

Proof. WLOG, we arrange

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \sigma_4. \tag{3.7}$$

Denote $\rho_{i,j} := \mathrm{Corr}(X_i, X_j)$, the covariance matrix is symmetric, so it is

$$M := \begin{bmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \rho_{1,3}\sigma_1\sigma_3 & \rho_{1,4}\sigma_1\sigma_4 \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{2,3}\sigma_2\sigma_3 & \rho_{2,4}\sigma_2\sigma_4 \\ \rho_{1,3}\sigma_1\sigma_3 & \rho_{2,3}\sigma_2\sigma_3 & \sigma_3^2 & \rho_{3,4}\sigma_3\sigma_4 \\ \rho_{1,4}\sigma_1\sigma_4 & \rho_{2,4}\sigma_2\sigma_4 & \rho_{3,4}\sigma_3\sigma_4 & \sigma_4^2 \end{bmatrix} \tag{3.8}$$

Since $\mathrm{Var}(\Sigma_{i=1}^4 X_i) = \mathbf{1}^T M \mathbf{1}$, let $M\mathbf{1} = \mathbf{0}$, we need

$$\begin{cases} \sigma_1 + \rho_{1,2}\sigma_2 + \rho_{1,3}\sigma_3 + \rho_{1,4}\sigma_4 = 0 \\ \rho_{1,2}\sigma_1 + \sigma_2 + \rho_{2,3}\sigma_3 + \rho_{2,4}\sigma_4 = 0 \\ \rho_{1,3}\sigma_1 + \rho_{2,3}\sigma_2 + \sigma_3 + \rho_{3,4}\sigma_4 = 0 \\ \rho_{1,4}\sigma_1 + \rho_{2,4}\sigma_2 + \rho_{3,4}\sigma_3 + \sigma_4 = 0. \end{cases} \tag{3.9}$$

It implies in particular that

$$\begin{cases} \mathrm{Var}(X_1) = \mathrm{Var}(X_2 + X_3 + X_4) \\ \mathrm{Var}(X_2) = \mathrm{Var}(X_1 + X_3 + X_4) \\ \mathrm{Var}(X_3) = \mathrm{Var}(X_1 + X_2 + X_4) \\ \mathrm{Var}(X_4) = \mathrm{Var}(X_1 + X_2 + X_3) \end{cases} \tag{3.10}$$

and that

$$\begin{cases} \mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_3 + X_4) \\ \mathrm{Var}(X_1 + X_3) = \mathrm{Var}(X_2 + X_4) \\ \mathrm{Var}(X_1 + X_4) = \mathrm{Var}(X_2 + X_3). \end{cases} \tag{3.11}$$

So from Equation (3.9),

$$\begin{cases} \rho_{1,4} = -\frac{\sigma_1+\rho_{1,2}\sigma_2+\rho_{1,3}\sigma_3}{\sigma_4} \\ \rho_{2,3} = -\frac{\sigma_1^2+2\rho_{1,2}\sigma_1\sigma_2+\sigma_2^2+2\rho_{1,3}\sigma_1\sigma_3+\sigma_3^2-\sigma_4^2}{2\sigma_2\sigma_3} \\ \rho_{2,4} = \frac{\sigma_1^2-\sigma_2^2+2\rho_{1,3}\sigma_1\sigma_3+\sigma_3^2-\sigma_4^2}{2\sigma_2\sigma_4} \\ \rho_{3,4} = \frac{\sigma_1^2+2\rho_{1,2}\sigma_1\sigma_2+\sigma_2^2-\sigma_3^2-\sigma_4^2}{2\sigma_3\sigma_4}. \end{cases} \tag{3.12}$$

Then we get the covariance matrix M in terms of $\rho_{1,2}$ and $\rho_{1,3}$. To simplify writing, let

$$A = \frac{\sigma_1^2+\sigma_2^2+\sigma_3^2-\sigma_4^2}{2\sigma_1\sigma_2}, B = \frac{\sigma_1^2-\sigma_2^2+\sigma_3^2-\sigma_4^2}{2\sigma_1\sigma_3}, C = \frac{\sigma_1^2+\sigma_2^2-\sigma_3^2-\sigma_4^2}{2\sigma_1\sigma_2}. \tag{3.13}$$

So $M =$

$$\begin{bmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \rho_{1,3}\sigma_1\sigma_3 & -\sigma_1(\sigma_1+\rho_{1,2}\sigma_2+\rho_{1,3}\sigma_3) \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & -\sigma_1\sigma_2(A+\rho_{1,2})-\rho_{1,3}\sigma_1\sigma_3 & \sigma_1\sigma_3(B+\rho_{1,3}) \\ \rho_{1,3}\sigma_1\sigma_3 & -\sigma_1\sigma_2(A+\rho_{1,2})-\rho_{1,3}\sigma_1\sigma_3 & \sigma_3^2 & \sigma_1\sigma_2(C+\rho_{1,2}) \\ -\sigma_1(\sigma_1+\rho_{1,2}\sigma_2+\rho_{1,3}\sigma_3) & \sigma_1\sigma_3(B+\rho_{1,3}) & \sigma_1\sigma_2(C+\rho_{1,2}) & \sigma_4^2 \end{bmatrix} \tag{3.14}$$

Now we only need to determine $\rho_{1,2}$ and $\rho_{1,3}$ and the matrix M (3.8) should be positive-semidefinite.
Fix $\rho_{1,2} = \rho$ and assume $-1 < \rho < 1$, since $\rho_{1,4}, \rho_{2,3}, \rho_{2,4}, \rho_{3,4}$ are functions in terms of $\rho_{1,2}$ and $\rho_{1,3}$, we only need to get a range of $\rho_{1,3}$ in terms of ρ to investigate the existence of all the parameters.
If $-1 \le \rho_{1,4} \le 1$, by Equation (3.12),

$$\underbrace{\frac{-\sigma_1-\rho\sigma_2-\sigma_4}{\sigma_3}}_{:=A_1} \le \rho_{1,3} \le \underbrace{\frac{-\sigma_1-\rho\sigma_2+\sigma_4}{\sigma_3}}_{:=A_2}. \tag{3.15}$$

Similarly, by $-1 \le \rho_{2,3} \le 1$ and Equation (3.12),

$$\underbrace{\frac{-2\rho\sigma_1\sigma_2-\sigma_1^2-\sigma_2^2-\sigma_3^2+\sigma_4^2-2\sigma_2\sigma_3}{2\sigma_1\sigma_3}}_{:=A_3} \le \rho_{1,3} \le \underbrace{\frac{-2\rho\sigma_1\sigma_2-\sigma_1^2-\sigma_2^2+2\sigma_2\sigma_3-\sigma_3^2+\sigma_4^2}{2\sigma_1\sigma_3}}_{:=A_4}. \tag{3.16}$$

To simplify writing, we write $A_3 = -(A+\rho_{1,2})\frac{\sigma_2}{\sigma_3} - \frac{\sigma_2}{\sigma_1}, A_4 = -(A+\rho_{1,2})\frac{\sigma_2}{\sigma_3} + \frac{\sigma_2}{\sigma_1}$ where A is defined in Equation (3.13).
Similarly, by $-1 \le \rho_{2,4} \le 1$ and Equation (3.12),

$$\underbrace{\frac{-\sigma_1^2+\sigma_2^2-2\sigma_2\sigma_4-\sigma_3^2+\sigma_4^2}{2\sigma_1\sigma_3}}_{:=A_5} \le \rho_{1,3} \le \underbrace{\frac{-\sigma_1^2+\sigma_2^2+2\sigma_2\sigma_4-\sigma_3^2+\sigma_4^2}{2\sigma_1\sigma_3}}_{:=A_6}. \tag{3.17}$$

Also, we write $A_5 = -B - \frac{\sigma_2\sigma_4}{\sigma_1\sigma_3}, A_6 = -B + \frac{\sigma_2\sigma_4}{\sigma_1\sigma_3}$ where B is defined in Equation (3.13).
Similarly, by $-1 \le \rho_{3,4} \le 1$ and Equation (3.12),

$$\underbrace{\frac{-\sigma_1^2-\sigma_2^2+\sigma_3^2-2\sigma_3\sigma_4+\sigma_4^2}{2\sigma_1\sigma_2}}_{:=A_7} \le \rho \le \underbrace{\frac{-\sigma_1^2-\sigma_2^2+\sigma_3^2+2\sigma_3\sigma_4+\sigma_4^2}{2\sigma_1\sigma_2}}_{:=A_8}. \tag{3.18}$$

Also, we write $A_7 = -\frac{\sigma_3\sigma_4}{\sigma_1\sigma_2} - C, A_8 = \frac{\sigma_3\sigma_4}{\sigma_1\sigma_2} - C$ where C is defined in Equation (3.13).
We first compare A_5, A_6 with $-1, 1$.

Case 1: $\sigma_2+\sigma_3 \ge \sigma_1+\sigma_4$

$$A_5 - (-1) = \frac{-(\sigma_1-\sigma_3)^2+(\sigma_2-\sigma_4)^2}{2\sigma_1\sigma_3} \ge 0, \tag{3.19}$$

so $A_5 \ge -1$.

$$A_6 - 1 = \frac{-(\sigma_1+\sigma_3)^2+(\sigma_2+\sigma_4)^2}{2\sigma_1\sigma_3}, \tag{3.20}$$

so $A_6 \le 1$.
As $A_1 = -\frac{\sigma_2}{\sigma_3}\rho - \frac{\sigma_1+\sigma_4}{\sigma_3}, A_3 = -\frac{\sigma_2}{\sigma_3}\rho + \frac{\sigma_4^2-\sigma_1^2-\sigma_2^2-\sigma_3^2-2\sigma_2\sigma_3}{2\sigma_1\sigma_3}$, A_1, A_2, A_3, A_5 are parallel lines.

$$A_1 - A_3 = \frac{(\sigma_2+\sigma_3)^2-(\sigma_1+\sigma_4)^2}{2\sigma_1\sigma_3} \ge 0, \tag{3.21}$$

so $A_1 \ge A_3$.

$$A_4 - A_2 = \frac{(\sigma_1-\sigma_4)^2-(\sigma_2-\sigma_3)^2}{2\sigma_1\sigma_3} \ge 0, \tag{3.22}$$

so $A_4 \ge A_2$.

Now we compare $A_7, A_8, -1, 1$.

$$A_7 - (-1) = \frac{-(\sigma_1 - \sigma_2)^2 + (\sigma_3 - \sigma_4)^2}{2\sigma_1\sigma_2} \geq 0, \tag{3.23}$$

so $A_7 \geq -1$.
Since $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \sigma_4$,

$$A_8 - 1 = \frac{-(\sigma_1 + \sigma_2)^2 + (\sigma_3 + \sigma_4)^2}{2\sigma_1\sigma_2} \leq 0, \tag{3.24}$$

so $A_8 \leq 1$.
When $A_5 = A_1$,

$$\rho = \frac{-\sigma_1^2 - 2\sigma_1\sigma_4 - \sigma_2^2 + 2\sigma_2\sigma_4 + \sigma_3^2 - \sigma_4^2}{2\sigma_1\sigma_2} := \rho^{A_1,A_5}.$$

Now define $D := \frac{\sigma_1^2+\sigma_2^2-\sigma_3^2+\sigma_4^2}{2\sigma_1\sigma_2}$, then $\rho^{A_1,A_5} = \frac{(\sigma_2-\sigma_1)\sigma_4}{\sigma_1\sigma_2} - D$.

$$\rho^{A_1,A_5} - A_7 = \frac{\sigma_4(\sigma_2 - \sigma_1 + \sigma_3 - \sigma_4)}{\sigma_1\sigma_2} \geq 0,$$

so $\rho^{A_1,A_5} \geq A_7$.
When $A_1 = A_6$,

$$\rho = \frac{-\sigma_1^2 - 2\sigma_1\sigma_4 - \sigma_2^2 - 2\sigma_2\sigma_4 + \sigma_3^2 - \sigma_4^2}{2\sigma_1\sigma_2} := \rho^{A_1,A_6},$$

$$\rho^{A_1,A_6} - A_7 = -\frac{\sigma_4(\sigma_1 + \sigma_2 - \sigma_3 + \sigma_4)}{\sigma_1\sigma_2} \leq 0,$$

so $\rho^{A_1,A_6} \leq A_7$.
And $\rho^{A_2,A_6} - \rho^{A_1,A_5} = 2\sigma_4(\frac{1}{\sigma_2} - \frac{1}{\sigma_1}) \geq 0$. When $A_2 = A_5$,

$$\rho = \frac{-\sigma_1^2 + 2\sigma_1\sigma_4 - \sigma_2^2 + 2\sigma_2\sigma_4 + \sigma_3^2 - \sigma_4^2}{2\sigma_1\sigma_2} := \rho^{A_2,A_5},$$

$$A_8 - \rho^{A_2,A_5} = \frac{\sigma_4(\sigma_3 + \sigma_4 - \sigma_1 - \sigma_2)}{\sigma_1\sigma_2} \leq 0,$$

so $A_8 \leq \rho^{A_2,A_5}$.

When $A_2 = A_6$,

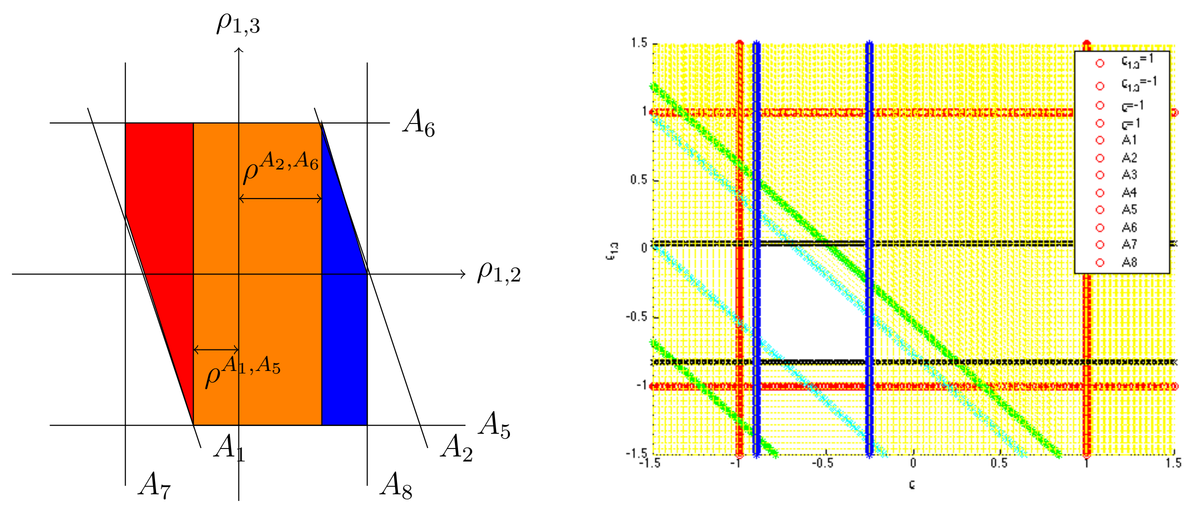
$$\rho = \frac{-\sigma_1^2 + 2\sigma_1\sigma_4 - \sigma_2^2 - 2\sigma_2\sigma_4 + \sigma_3^2 - \sigma_4^2}{2\sigma_1\sigma_2} := \rho^{A_2,A_6}.$$

Also, write $\rho^{A_2,A_6} = \frac{(\sigma_1-\sigma_2)\sigma_4}{\sigma_1\sigma_2} - D$.

$$A_8 - \rho^{A_2,A_6} = \frac{\sigma_4(\sigma_2+\sigma_3+\sigma_4-\sigma_1)}{\sigma_1\sigma_2} \geq 0,$$

so $A_8 \geq \rho^{A_2,A_6}$.

Now we get an area for choosing $\rho, \rho_{1,3}$ except checking $L_1 > 0$. Panel A in Figure 3.2 is a sketch of the shape of the area, the shaded area is the acceptance area for $\rho, \rho_{1,3}$ (How the line intercept x, y coordinate is not clear). Panel B in Figure 3.2 is a simulation plot using MATLAB with $\sigma_1 = 8, \sigma_2 = 7.5, \sigma_3 = 6.5, \sigma_4 = 3$ where the four red lines denote $+1, -1$, cyan lines denote A_1, A_2, green lines denote A_3, A_4, black lines denote A_5, A_6, blue lines denote A_7, A_8. The white area is the acceptance area for $(\rho, \rho_{1,3})$ which has the same shape as in Panel A of Figure 3.2.



Panel A Panel B

Figure 3.2: Panel A: Sketch of the area under case 1.
Panel B: Simulation of the area under case 1

Case 2: $\sigma_2 + \sigma_3 \leq \sigma_1 + \sigma_4$
By Equations (3.19), (3.20), (3.21), (3.22), (3.23), (3.24), we get

$$A_5 \leq -1, A_6 \leq 1, A_1 \leq A_3, A_4 \geq A_2, A_7 \leq -1, A_8 \leq 1.$$

Next we need to decide how the lines intersect with each other to see the shape of the area.
When $A_2 = A_6$,

$$\rho = \frac{-\sigma_1^2 + 2\sigma_1\sigma_4 - \sigma_2^2 - 2\sigma_2\sigma_4 + \sigma_3^2 - \sigma_4^2}{2\sigma_1\sigma_2} := \rho^{A_2,A_6}.$$

When $-1 = A_3$,

$$\rho = \frac{-\sigma_1^2 + 2\sigma_1\sigma_3 - \sigma_2^2 - 2\sigma_2\sigma_3 - \sigma_3^2 + \sigma_4^2}{2\sigma_1\sigma_2} := \rho^{A_3,-1}.$$

Also, write $\rho^{A_3,-1} = \frac{(\sigma_1-\sigma_2)\sigma_3}{\sigma_1\sigma_2} - A$.

$$\rho^{A_2,A_6} - \rho^{A_3,-1} = \frac{(\sigma_3 - \sigma_4)(-\sigma_1 + \sigma_2 + \sigma_3 + \sigma_4)}{\sigma_1\sigma_2} \geq 0,$$

so $\rho^{A_2,A_6} \geq \rho^{A_3,-1}$.
When $A_3 = A_6$,

$$\rho = \frac{-\sigma_2 - \sigma_3 - \sigma_4}{\sigma_1} := \rho^{A_3,A_6},$$

$$-1 - \rho^{A_3,A_6} = \frac{\sigma_2 + \sigma_3 + \sigma_4 - \sigma_1}{\sigma_1} \geq 0,$$

so $-1 \geq \rho^{A_3,A_6}$.

$$\begin{aligned}\rho^{A_3,-1} - (-1) &= -\frac{\sigma_1^2 - 2\sigma_1(\sigma_2 + \sigma_3) + \sigma_2^2 + 2\sigma_2\sigma_3 + \sigma_3^2 - \sigma_4^2}{2\sigma_1\sigma_2} \\ &= \frac{(\sigma_2 + \sigma_3 + \sigma_4 - \sigma_1)((\sigma_1 + \sigma_4) - (\sigma_2 + \sigma_3))}{2\sigma_1\sigma_2} \geq 0,\end{aligned} \tag{3.25}$$

so $\rho^{A_3,-1} \geq -1$.
And

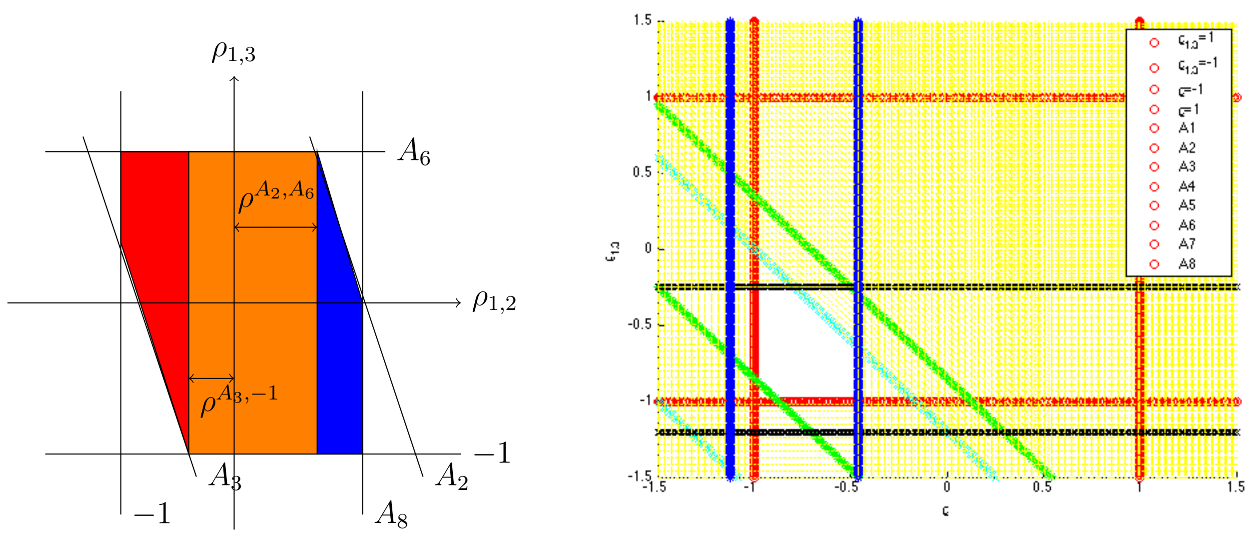
$$A_8 - \rho^{A_2,A_6} = \frac{\sigma_4(\sigma_2 + \sigma_3 + \sigma_4 - \sigma_1)}{\sigma_1\sigma_2} \geq 0,$$

so $A_8 \geq \rho^{A_2,A_6}$.
When $A_2 = -1$,

$$\rho = \frac{\sigma_3 + \sigma_4 - \sigma_1}{\sigma_2} := \rho^{A_2,-1},$$

$$\rho^{A_2,-1} - A_8 = \frac{(\sigma_1 + \sigma_2 - \sigma_3 - \sigma_4)(\sigma_2 + \sigma_3 + \sigma_4 - \sigma_1)}{2\sigma_1\sigma_2} \geq 0,$$

so $\rho^{A_2,-1} \geq A_8$.
Similarly, we get a sketch of the area in Panel A of Figure 3.3. Panel B of Figure 3.3 is a simulation plot using Matlab with $\sigma_1 = 10, \sigma_2 = 6, \sigma_3 = 5, \sigma_4 = 4$ where the four red lines denote $+1, -1$, cyan lines denote A_1, A_2, green line denotes A_3, A_4, black lines denote A_5, A_6, blue lines denote A_7, A_8. The white area is the acceptance area for $(\rho, \rho_{1,3})$ which has the same shape as in Panel A of Figure 3.3.



Panel A Panel B

Figure 3.3: Panel A: Sketch of the area under case 2
Panel B: Simulation of the area under case 2

In summary, every colored equation in the following corresponds to the colored area in

Panel A of Figures 3.2 and 3.3.

$$\text{When } \sigma_2+\sigma_3 \geq \sigma_1+\sigma_4, \begin{cases} A_1(\rho) \leq \rho_{1,3} \leq A_6, & \text{if } A_7 \leq \rho \leq \rho^{A_1,A_5} \\ A_5 \leq \rho_{1,3} \leq A_6, & \text{if } \rho^{A_1,A_5} \leq \rho \leq \rho^{A_2,A_6} \\ A_5 \leq \rho_{1,3} \leq A_2(\rho), & \text{if } \rho^{A_2,A_6} \leq \rho \leq A_8, \end{cases} \tag{3.26}$$

$$\text{when } \sigma_2+\sigma_3 \leq \sigma_1+\sigma_4, \begin{cases} A_3(\rho) \leq \rho_{1,3} \leq A_6, & \text{if } -1 \leq \rho \leq \rho^{A_3,-1} \\ -1 \leq \rho_{1,3} \leq A_6, & \text{if } \rho^{A_3,-1} \leq \rho \leq \rho^{A_2,A_6} \\ -1 \leq \rho_{1,3} \leq A_2(\rho), & \text{if } \rho^{A_2,A_6} \leq \rho \leq A_8. \end{cases} \tag{3.27}$$

An area works for both cases is,

$$\begin{cases} \max_{\rho}(A_1, A_3) \leq \rho_{1,3} \leq A_6 & \text{if } \max(A_7,-1) \leq \rho \leq \min(\rho^{A_1,A_5}, \rho^{A_3,-1}) \\ \max(-1, A_5) \leq \rho_{1,3} \leq A_6 & \text{if } \max(\rho^{A_1,A_5}, \rho^{A_3,-1}) \leq \rho \leq \rho^{A_2,A_6} \\ \max(-1, A_5) \leq \rho_{1,3} \leq A_2(\rho) & \text{if } \rho^{A_2,A_6} \leq \rho \leq A_8. \end{cases} \tag{3.28}$$

To ensure the positive-semidefinite of M, we use Cholesky decomposition here: every real symmetric, positive-semidefinite matrix M can be written as $M = LL^T$ where L is a lower triangular matrix with real and non-negative diagonal entries, and L^T denotes the conjugate transpose of L (see Golub and Van Loan [2012]). If we can find such matrix L, then the matrix M is positive-semidefinite. Note L is not unique when M is positive-semidefinite matrix but not positive definite. Let $L_{i,j}$ denotes the entries of L. By the Cholesky-Banachiewicz and Cholesky-Crout algorithms (see Golub and Van Loan [2012]) and by $M = LL^T$,

$$\begin{aligned} \begin{bmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \rho_{1,3}\sigma_1\sigma_3 & \rho_{1,4}\sigma_1\sigma_4 \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{2,3}\sigma_2\sigma_3 & \rho_{2,4}\sigma_2\sigma_4 \\ \rho_{1,3}\sigma_1\sigma_3 & \rho_{2,3}\sigma_2\sigma_3 & \sigma_3^2 & \rho_{3,4}\sigma_3\sigma_4 \\ \rho_{1,4}\sigma_1\sigma_4 & \rho_{2,4}\sigma_2\sigma_4 & \rho_{3,4}\sigma_3\sigma_4 & \sigma_4^2 \end{bmatrix} &= \begin{bmatrix} L_{1,1} & 0 & 0 & 0 \\ L_{2,1} & L_{2,2} & 0 & 0 \\ L_{3,1} & L_{3,2} & L_{3,3} & 0 \\ L_{4,1} & L_{4,2} & L_{4,3} & L_{4,4} \end{bmatrix} \begin{bmatrix} L_{1,1} & L_{2,1} & L_{3,1} & L_{4,1} \\ 0 & L_{2,2} & L_{3,2} & L_{4,2} \\ 0 & 0 & L_{3,3} & L_{4,3} \\ 0 & 0 & 0 & L_{4,4} \end{bmatrix} \\ &= \begin{bmatrix} L_{1,1}^2 & L_{1,1}L_{2,1} & L_{1,1}L_{3,1} & L_{1,1}L_{4,1} \\ 0 & L_{2,1}^2+L_{2,2}^2 & L_{2,1}L_{3,1}+L_{2,2}L_{3,2} & L_{2,1}L_{4,1}+L_{2,2}L_{4,2} \\ 0 & 0 & L_{3,1}^2+L_{3,2}^2+L_{3,3}^2 & L_{3,1}L_{4,1}+L_{3,2}L_{4,2}+L_{3,3}L_{4,3} \\ 0 & 0 & 0 & L_{4,1}^2+L_{4,2}^2+L_{4,3}^2+L_{4,4}^2 \end{bmatrix} \end{aligned} \tag{3.29}$$

Then

$$L_{j,j} = \sqrt{M_{j,j} - \Sigma_{k=1}^{j-1} L_{j,k}^2},$$
$$L_{i,j} = \frac{1}{L_{j,j}}(M_{i,j} - \Sigma_{k=1}^{j-1} L_{i,k} L_{j,k}) \text{ for } i > j, \tag{3.30}$$

where $M_{i,j}$ denotes the entry of M. We need $L_{j,j} \neq 0$ to make this algorithm work.

$$\begin{aligned}
L_{1,1} &= \sigma_1 > 0, L_{2,1} = \rho_{1,2}\sigma_2, L_{3,1} = \rho_{1,3}\sigma_3, L_{4,1} = \rho_{1,4}\sigma_4 = -\sigma_1 - \rho_{1,2}\sigma_2 - \rho_{1,3}\sigma_3, \\
L_{2,2} &= \sqrt{\sigma_2^2 - \rho_{1,2}^2\sigma_2^2}, (L_{2,2} > 0 \Rightarrow \rho_{1,2} \neq \pm 1), \\
L_{3,2} &= \frac{1}{L_{2,2}}(M_{3,2} - L_{3,1}L_{2,1}) = \frac{1}{\sqrt{1-\rho_{1,2}^2}}(\rho_{2,3}\sigma_3 - \rho_{1,3}\rho_{1,2}\sigma_3) \\
&= -\frac{2\rho_{1,2}\rho_{1,3}\sigma_2\sigma_3 + 2\rho_{1,2}\sigma_1\sigma_2 + 2\rho_{1,3}\sigma_1\sigma_3 + \sigma_1^2 + \sigma_2^2 + \sigma_3^2 - \sigma_4^2}{2\sigma_2\sqrt{1-\rho_{1,2}^2}}, \\
L_{4,2} &= \frac{1}{L_{2,2}}(M_{4,2} - L_{4,1}L_{2,1}) \\
&= \frac{1}{\sqrt{1-\rho_{1,2}^2}}(\rho_{2,4}\sigma_4 - \rho_{1,2}\rho_{1,4}\sigma_4) \\
&= \frac{\left(2\rho_{1,2}^2 - 1\right)\sigma_2^2 + 2\sigma_1(\rho_{1,2}\sigma_2 + \rho_{1,3}\sigma_3) + 2\rho_{1,2}\rho_{1,3}\sigma_2\sigma_3 + \sigma_1^2 + \sigma_3^2 - \sigma_4^2}{2\sigma_2\sqrt{1-\rho_{1,2}^2}}, \\
L_{3,3} &= \sqrt{\sigma_3^2 - L_{3,2}^2 - L_{3,1}^2}, \\
&\quad \text{(we need } \sigma_3^2 - L_{3,2}^2 - L_{3,1}^2 > 0 \text{ here)} \\
L_{4,3} &= \frac{1}{L_{3,3}}(M_{4,3} - L_{4,1}L_{3,1} - L_{4,2}L_{3,2}), \\
L_{4,4} &= \sqrt{M_{4,4} - L_{4,1}^2 - L_{4,2}^2 - L_{4,3}^2} = 0.
\end{aligned} \tag{3.31}$$

If we pick $\rho_{1,2} \in (-1, 1)$, then the only condition for M to be positive-semidefinite is $\sigma_3^2 - L_{3,2}^2 - L_{3,1}^2 > 0$,

$$\sigma_3^2 - L_{3,2}^2 - L_{3,1}^2 = \frac{\sigma_3^2}{1-\rho_{1,2}^2}\underbrace{[(1-\rho_{1,2}^2)(1-\rho_{1,3}^2) - (\rho_{2,3} - \rho_{1,2}\rho_{1,3})^2]}_{L_1} > 0, \tag{3.32}$$

which is:

$$L_1 = (\rho_{1,2}^2 - 1)(\rho_{1,3}^2 - 1) - (\rho_{1,2}(\rho_{1,3} + \sigma_1\sigma_2) + \rho_{1,3}\sigma_1\sigma_3 + \sigma_1\sigma_2 A)^2 > 0. \tag{3.33}$$

□

The following is an example on how to use Proposition 3.1.8.

Example 3.1.9. When $X_i \sim N(\mu_i, \sigma_i^2)$, where $\mu_i \in \mathbb{R}$ and $\sigma_1 = 10, \sigma_2 = 6, \sigma_3 = 5, \sigma_4 = 4$, the covariance matrix for $\text{Var}(\Sigma_{i=1}^4 X_i) = 0$ is

$$\begin{bmatrix} 100 & -\frac{95}{2} & -\frac{25}{2} & -40 \\ -\frac{95}{2} & 36 & -\frac{25}{2} & 24 \\ -\frac{25}{2} & -\frac{25}{2} & 25 & 0 \\ -40 & 24 & 0 & 16 \end{bmatrix} \tag{3.34}$$

Proof. By Proposition 3.1.8, Equations (3.2) and (3.5), $A = \frac{29}{24}, B = \frac{73}{100}, C = \frac{19}{24}, D = \frac{127}{120}$. Then by Equation (3.28), we can choose

$$\rho_{1,2} = \rho^{A_2,A_6} = \frac{(\sigma_1 - \sigma_2)\sigma_4}{\sigma_1\sigma_2} - D = -\frac{19}{24},$$

$$\rho_{1,3} = A_6 = -B + \frac{\sigma_2\sigma_4}{\sigma_1\sigma_3} = -\frac{1}{4},$$

which satisfy constraint (ii) in Proposition 3.1.4 (In addition, $\rho_{1,2} = \rho^{A_2,A_6}, \rho_{1,3} = A_6$ always satisfy constraint (ii)). Then we check constraint (i) using Equation (3.33), $L_1 = \frac{1811}{5184}$ is indeed positive. So by the matrix in Equation (3.1), $A = \frac{29}{24}, B = \frac{73}{100}, C = \frac{19}{24}$, we get the covariance matrix. □

Conjecture 3.1.10. *In Proposition 3.1.8, we get a set of constraints for $\rho_{1,2}, \rho_{1,3}$ under 2 cases ($\sigma_2 + \sigma_3 \leq \sigma_1 + \sigma_4$ and $\sigma_2 + \sigma_3 \geq \sigma_1 + \sigma_4$) respectively. An area works for both cases is:*

$$\begin{cases} \max\limits_{\rho_{1,2}}(A_1, A_3) \leq \rho_{1,3} \leq A_6 & \textit{if } \max(A_7, -1) \leq \rho_{1,2} \leq \min(\rho^{A_1,A_5}, \rho^{A_3,-1}) \\ \max(-1, A_5) \leq \rho_{1,3} \leq A_6 & \textit{if } \max(\rho^{A_1,A_5}, \rho^{A_3,-1}) \leq \rho_{1,2} \leq \rho^{A_2,A_6} \\ \max(-1, A_5) \leq \rho_{1,3} \leq A_2(\rho_{1,2}) & \textit{if } \rho^{A_2,A_6} \leq \rho_{1,2} \leq A_8. \end{cases} \tag{3.35}$$

There are some constants (in terms of σ_i) satisfy Equation (3.35) *such as* $\rho_{1,2} = \rho^{A_2,A_6}$, $\rho_{1,3} = A_6$. *However,* $L_1 > 0$ *needs to be checked.*

For $n \geq 5$, a similar method can be used but there is a larger number of constraints and the feasible set for the correlation coefficients is in dimension ≥ 2.

3.2 Bounds on Variance with Background Risk

3.2.1 Case of 3 Normal Distributed Random Variables

Both Sections 3.2.1 and 3.2.2 can be seen as an application of Section 3.1.

Proposition 3.2.1. *Assume* X_1, X_2, X_3 *follow* $N(0,1)$ *marginals,* Z *is a background risk and* (X_i, Z) *has* $\mathrm{BVN}(\rho_i)$, $i = 2, 3$. *The maximum and minimum variance* $\mathrm{Var}(X_1 + X_2 + X_3)$ *when* $Z = X_1$ *are* $3 + 2(\rho_2 + \rho_3) + 2\rho_2\rho_3 - 2\sqrt{(1-\rho_2^2)(1-\rho_3^2)}$ *and* $3 + 2\rho_2 + 2\rho_3 + 2\sqrt{(1-\rho_2^2)(1-\rho_3^2)} + 2\rho_2\rho_3$ *respectively.*

Proof. In general,

$$\begin{aligned}&\mathrm{Var}(X_1 + X_2 + X_3)\\&= \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \mathrm{Var}(X_3) + 2\mathrm{Cov}(X_1, X_2) + 2\mathrm{Cov}(X_1, X_3) + 2\mathrm{Cov}(X_2, X_3)\\&= 3 + 2\rho_2 + 2\rho_3 + 2\mathrm{Cov}(X_2, X_3).\end{aligned} \tag{3.36}$$

And $2\mathrm{Cov}(X_2, X_3) = \mathrm{Var}(X_2 + X_3) - \mathrm{Var}(X_2) - \mathrm{Var}(X_3) = \mathrm{Var}(X_2 + X_3) - 2$.

$$\begin{aligned}\mathrm{Var}(X_2 + X_3) = E((X_2 + X_3)^2) - E(X_2 + X_3)^2 = E((X_2 + X_3)^2) = E(E((X_2 + X_3)^2|Z))\\ = E(E((X_2|Z + X_3|Z)^2)).\end{aligned} \tag{3.37}$$

To get the maximum and minimum variance, we just need to maximize and minimize $E((X_2|Z + X_3|Z)^2)$.
Let U_2, U_3 follow $U[0,1]$ independently of Z, and $F_{2|Z}, F_{3|Z}$ be the conditional distributions

of $X_2|Z, X_3|Z$. Then $X_2|Z + X_3|Z = F^{-1}_{2|Z}(U_2) + F^{-1}_{3|Z}(U_3)$, so by Theorem 1.2.11,

$$F^{-1}_{2|Z}(U) + F^{-1}_{3|Z}(1-U) \leq_{cx} X_2|Z + X_3|Z \leq_{cx} F^{-1}_{2|Z}(U) + F^{-1}_{3|Z}(U)$$

where $U \sim U[0,1]$ independent of Z. Then

$$E((F^{-1}_{2|Z}(U) + F^{-1}_{3|Z}(U))^2) \geq E((X_2|Z + X_3|Z)^2) \geq E((F^{-1}_{2|Z}(U) + F^{-1}_{3|Z}(1-U))^2).$$

Since $Z = X_1$ and $(X_i, Z) \sim \text{BVN}(\rho_i)$, $X_i|X_1 = x_1 \sim N(\rho_i x_1, 1-\rho_i^2), i = 2, 3$.
We calculate the minimum variance first. Denote $Y_2 := F^{-1}_{2|X_1=x_1}(U), Y_3 := F^{-1}_{3|X_1=x_1}(1-U)$, then

$$Y_i \sim N(\rho_i x_1, 1-\rho_i^2),$$
$$\frac{Y_i - \rho_i x_1}{\sqrt{1-\rho_i^2}} \sim N(0,1),$$

since Y_2, Y_3 move countermonotonically and let Φ denote the distribution function of $N(0,1)$,

$$\frac{Y_2 - \rho_2 x_1}{\sqrt{1-\rho_2^2}} = \Phi^{-1}(U), \frac{Y_3 - \rho_3 x_1}{\sqrt{1-\rho_3^2}} = \Phi^{-1}(1-U).$$

So

$$Y_2 + Y_3 = \sqrt{1-\rho_2^2}\Phi^{-1}(U) + \rho_2 x_1 + \sqrt{1-\rho_3^2}\Phi^{-1}(1-U) + \rho_3 x_1.$$

Since $\Phi^{-1}(U)$ and $\Phi^{-1}(1-U)$ are symmetric to 0,

$$\frac{Y_2 - \rho_2 x_1}{\sqrt{1-\rho_2^2}} = -\frac{Y_3 - \rho_3 x_1}{\sqrt{1-\rho_3^2}}$$

$$\begin{aligned}
\implies Y_3 &= -\frac{\sqrt{1-\rho_3^2}}{\sqrt{1-\rho_2^2}}(Y_2 - \rho_2 x_1) + \rho_3 x_1 \\
&= -\frac{\sqrt{1-\rho_3^2}}{\sqrt{1-\rho_2^2}}(\sqrt{1-\rho_2^2}\Phi(U) + \rho_2 x_1 - \rho_2 x_1) + \rho_3 x_1 \\
&= -\sqrt{1-\rho_3^2}\Phi^{-1}(U) + \rho_3 x_1.
\end{aligned}$$

So

$$F^{-1}_{2|X_1}(U) + F^{-1}_{3|X_1}(1-U) = (\sqrt{1-\rho_2^2} - \sqrt{1-\rho_3^2})\Phi^{-1}(U) + (\rho_2+\rho_3)X_1.$$

Since U, X_1 are independent and $\Phi^{-1}(U), X_1 \sim N(0,1)$,

$$\begin{aligned}
&\mathrm{Var}(F^{-1}_{2|Z}(U) + F^{-1}_{3|Z}(1-U)) \\
=&\mathrm{Var}((\sqrt{1-\rho_2^2} - \sqrt{1-\rho_3^2})\Phi^{-1}(U) + (\rho_2+\rho_3)X_1) \\
=&\mathrm{Var}((\sqrt{1-\rho_2^2} - \sqrt{1-\rho_3^2})\Phi^{-1}(U)) + 2\mathrm{Cov}((\sqrt{1-\rho_2^2} - \sqrt{1-\rho_3^2})\Phi^{-1}(U), (\rho_2+\rho_3)X_1) \\
&\quad + \mathrm{Var}((\rho_2+\rho_3)X_1) \\
=&(\sqrt{1-\rho_2^2} - \sqrt{1-\rho_3^2})^2 + 0 + (\rho_2+\rho_3)^2 \\
=&2(1+\rho_2\rho_3 - \sqrt{(1-\rho_2^2)(1-\rho_3^2)}), \\
&E(F^{-1}_{2|Z}(U) + F^{-1}_{3|Z}(1-U)) = E((\sqrt{1-\rho_2^2} - \sqrt{1-\rho_3^2})\Phi^{-1}(U) + (\rho_2+\rho_3)X_1) = 0.
\end{aligned}$$

Then

$$\begin{aligned}
E((F^{-1}_{2|Z}(U) + F^{-1}_{3|Z}(1-U))^2) =&\mathrm{Var}(F^{-1}_{2|Z}(U) + F^{-1}_{3|Z}(1-U)) + E^2(F^{-1}_{2|Z}(U) + F^{-1}_{3|Z}(1-U)) \\
=&(2(1+\rho_2\rho_3 - \sqrt{(1-\rho_2^2)(1-\rho_3^2)})).
\end{aligned}$$

So by Equations (3.36), (3.37), the minimum of $\mathrm{Var}(X_1+X_2+X_3)$ is

$$\begin{aligned}
&3 + 2\rho_2 + 2\rho_3 + E(2(1+\rho_2\rho_3 - \sqrt{(1-\rho_2^2)(1-\rho_3^2)})) - 2 \\
=&3 + 2(\rho_2+\rho_3) + 2\rho_2\rho_3 - 2\sqrt{(1-\rho_2^2)(1-\rho_3^2)}.
\end{aligned} \tag{3.38}$$

Then we calculate the maximum variance, denote $Y_2 := F^{-1}_{2|X_1=x_1}(U), Y_3 := F^{-1}_{3|X_1=x_1}(U)$, similar to above, $\dfrac{Y_i - \rho_i x_1}{\sqrt{1-\rho_i^2}} \sim N(0,1)$, let $\dfrac{Y_i - \rho_i x_1}{\sqrt{1-\rho_i^2}} = \Phi^{-1}(U)$ so that Y_2, Y_3 move comonotonically.

So

$$F^{-1}_{2|X_1}(U) + F^{-1}_{3|X_1}(U) = Y_2 + Y_3 = (\sqrt{1-\rho_2^2} + \sqrt{1-\rho_3^2})\Phi^{-1}(U) + (\rho_2+\rho_3)x_1.$$

Since U, X_1 are independent and $\Phi^{-1}(U), X_1 \sim N(0,1)$,

$$\begin{aligned}E((F^{-1}_{2|Z}(U)+F^{-1}_{3|Z}(U))^2) =&\mathrm{Var}(F^{-1}_{2|Z}(U)+F^{-1}_{3|Z}(U))+E^2(F^{-1}_{2|Z}(U)+F^{-1}_{3|Z}(U))\\ =&\mathrm{Var}((\sqrt{1-\rho_2^2}+\sqrt{1-\rho_3^2})\Phi^{-1}(U)+(\rho_2+\rho_3)X_1)+0^2\\ =&(\sqrt{1-\rho_2^2}+\sqrt{1-\rho_3^2})^2+(\rho_2+\rho_3)^2\\ =&2+2\sqrt{(1-\rho_2^2)(1-\rho_3^2)}+2\rho_2\rho_3.\end{aligned} \tag{3.39}$$

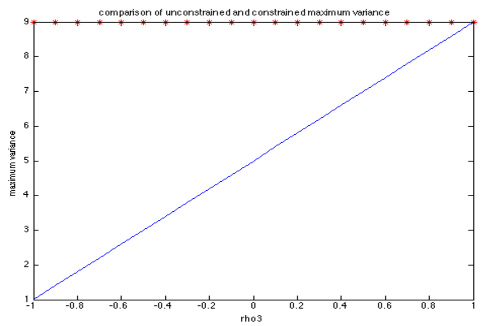
So by Equations (3.36), (3.37), the maximum of $\mathrm{Var}(X_1+X_2+X_3)$ is

$$3+2\rho_2+2\rho_3+2\sqrt{(1-\rho_2^2)(1-\rho_3^2)}+2\rho_2\rho_3.$$

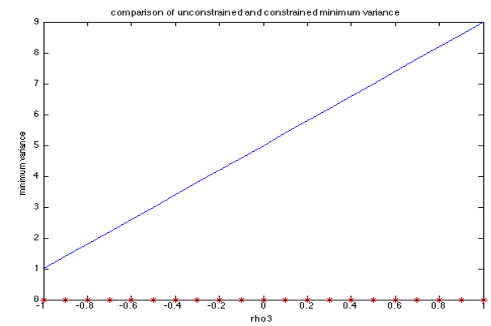
□

We now illustrate Proposition 3.2.1 with a comparison of unconstrained and constrained maximum (Panel A, C, E) and minimum (Panel B, D, F) variance under different assumptions on ρ_2. In Figure 3.4, the lines with $*$ and solid lines indicate respectively the maximum unconstrained variance on Panel A, C, E and the minimum unconstrained variance on Panel B, D, F.

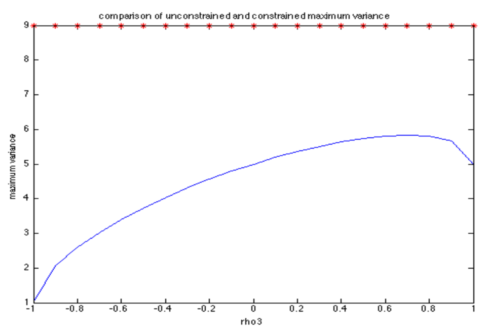
When X_1, X_2, X_3 are comonotonic, $F_1(X_1) = F_2(X_2) = F_3(X_3) = U$, then $\mathrm{Var}(X_1 + X_2+X_3)$ gets the maximum variance $\mathrm{Var}(F_1^{-1}(U)+F_2^{-1}(U)+F_3^{-1}(U)) = \mathrm{Var}(3\Phi^{-1}(U)) = 9$. The minimum of $\mathrm{Var}(X_1+X_2+X_3)$ is 0 since X_1, X_2, X_3 are jointly mixable (see Definition 1.2.9). The two cases correspond to the lines with $*$.



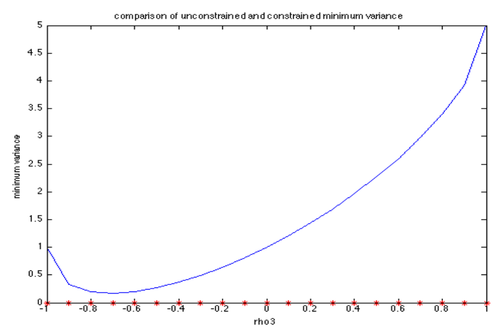
Panel A: $\rho_2 = 1,\ \rho_3 \in [-1, 1]$



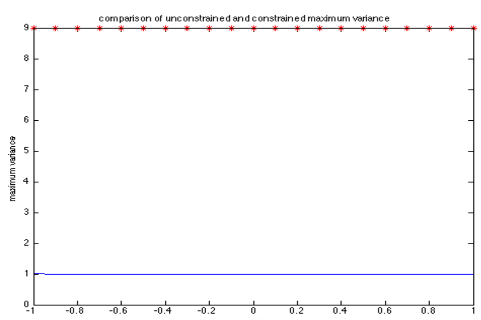
Panel B: $\rho_2 = 1,\ \rho_3 \in [-1, 1]$



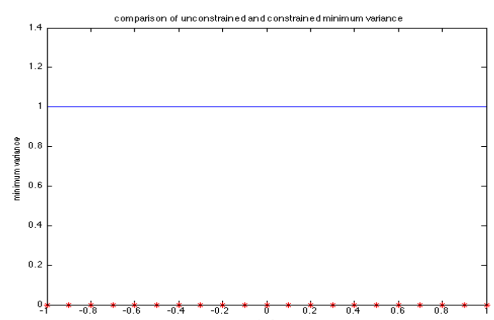
Panel C: $\rho_2 = 0,\ \rho_3 \in [-1, 1]$



Panel D: $\rho_2 = 0,\ \rho_3 \in [-1, 1]$



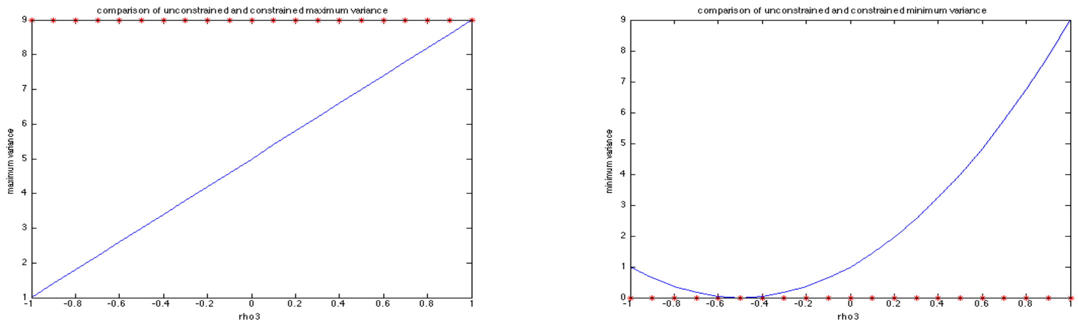
Panel E: $\rho_2 = -1,\ \rho_3 \in [-1, 1]$



Panel F: $\rho_2 = -1,\ \rho_3 \in [-1, 1]$

Figure 3.4: Comparisons of unconstrained variance and constrained variance when $\rho_2 = 1$ with ρ_3 varies from -1 to 1.

We apply Proposition 3.1.6 under the assumption that $\rho_2 = \rho_3$. As can be seen from the Panel B of Figure 3.5, $\mathrm{Var}(X_1 + X_2 + X_3)$ reaches minimum 0 when $\rho_2 = \rho_3 = -\frac{1}{2}$, which is $\mathrm{Cov}(X_1, X_2) = \mathrm{Cov}(X_1, X_3) = -\frac{1}{2}$.



Panel A: $\rho_2 = \rho_3 \in [-1, 1]$ Panel B: $\rho_2 = \rho_3 \in [-1, 1]$

Figure 3.5: Comparison of unconstrained variance and constrained variance in proposition 3.2.1 when $\rho_2 = \rho_3$

3.2.2 Case of 4 Normal Distributed Random Variables

The following proposition is a generalization of Proposition 3.2.1 in dimension $d = 4$.

Proposition 3.2.2. *Assume X_1, X_2, X_3, X_4 follow $N(0,1)$ marginals, Z is a background risk and (X_i, Z) has* BVN(ρ_i)*, $i = 1, 2, 3$. The maximum and minimum variance* $\mathrm{Var}(X_1 + X_2 + X_3 + X_4)$ *when $Z = X_4$ are*

$$4 + 2\rho_1 + 2\rho_2 + 2\rho_3 + 2\sqrt{(1-\rho_1^2)(1-\rho_2^2)} + 2\sqrt{(1-\rho_1^2)(1-\rho_3^2)} + 2\sqrt{(1-\rho_2^2)(1-\rho_3^2)}$$
$$+ 2\rho_1\rho_2 + 2\rho_1\rho_3 + 2\rho_2\rho_3$$

and

$$\begin{cases} (\rho_1 + \rho_2 + \rho_3 + 1)^2 & \text{if } \max\limits_{1 \le i \le 3} \sqrt{1-\rho_i^2} \le \frac{1}{2}\Sigma_{i=1}^3 \sqrt{1-\rho_i^2} \\ (\max\limits_{1 \le i \le 3} \sqrt{1-\rho_i^2} - \frac{1}{2}\Sigma_{i=1}^3 \sqrt{1-\rho_i^2})^2 & \text{if } \max\limits_{1 \le i \le 3} \sqrt{1-\rho_i^2} > \frac{1}{2}\Sigma_{i=1}^3 \sqrt{1-\rho_i^2}. \end{cases}$$

Proof. In general,

$$\begin{aligned}
&\mathrm{Var}(X_1 + X_2 + X_3 + X_4) \\
&= \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \mathrm{Var}(X_3) + \mathrm{Var}(X_4) + 2\mathrm{Cov}(X_1, X_2) + 2\mathrm{Cov}(X_1, X_3) + 2\mathrm{Cov}(X_1, X_4) \\
&+ 2\mathrm{Cov}(X_2, X_3) + 2\mathrm{Cov}(X_2, X_4) + 2\mathrm{Cov}(X_3, X_4) \\
&= 4 + 2\rho_1 + 2\rho_2 + 2\rho_3 + 2\mathrm{Cov}(X_1, X_2) + 2\mathrm{Cov}(X_1, X_3) + 2\mathrm{Cov}(X_2, X_3).
\end{aligned} \tag{3.40}$$

And

$$2\mathrm{Cov}(X_1, X_2) + 2\mathrm{Cov}(X_1, X_3) + 2\mathrm{Cov}(X_2, X_3) = \mathrm{Var}(X_1 + X_2) + \mathrm{Var}(X_1 + X_3) + \mathrm{Var}(X_2 + X_3) - 6,$$

then

$$\begin{aligned}
&\mathrm{Var}(X_1 + X_2) + \mathrm{Var}(X_1 + X_3) + \mathrm{Var}(X_2 + X_3) = E((X_1 + X_2)^2) + E((X_1 + X_3)^2) + E((X_2 + X_3)^2) \\
&= E(E((X_1 + X_2)^2 + (X_1 + X_3)^2 + (X_2 + X_3)^2 | Z)) \\
&= E(E((X_1|Z + X_2|Z)^2 + (X_1|Z + X_3|Z)^2 + (X_2|Z + X_3|Z)^2)).
\end{aligned} \tag{3.41}$$

To get the maximum and minimum variance, we just need to maximize and minimize $E((X_1|Z + X_2|Z)^2 + (X_1|Z + X_3|Z)^2 + (X_2|Z + X_3|Z)^2)$.
For the upper bound, the argument is similar with the proof of Proposition 3.2.1, we have

$$\begin{aligned}
&E((F_{1|Z}^{-1}(U) + F_{2|Z}^{-1}(U))^2) + E((F_{1|Z}^{-1}(U) + F_{3|Z}^{-1}(U))^2) + E((F_{2|Z}^{-1}(U) + F_{3|Z}^{-1}(U))^2) \\
\geq &E((X_1|Z + X_2|Z)^2) + E((X_1|Z + X_3|Z)^2) + E((X_2|Z + X_3|Z)^2)
\end{aligned}$$

where $U \sim U[0, 1]$ independent of Z, $F_{1|Z}, F_{2|Z}, F_{3|Z}$ are the conditional distributions of $X_1|Z, X_2|Z, X_3|Z$.
Since $Z = X_4 \sim N(0, 1)$ and $(X_i, Z) \sim \mathrm{BVN}(\rho_i)$, $X_i|X_4 = x_4 \sim N(\rho_i x_4, 1 - \rho_i^2), i = 1, 2, 3,$ U and X_4 are independent. WLOG, by Equation (3.39),

$$E((F_{1|Z}^{-1}(U) + F_{2|Z}^{-1}(U))^2) = 2 + 2\sqrt{(1 - \rho_1^2)(1 - \rho_2^2)} + 2\rho_1\rho_2,$$

$$E((F_{1|Z}^{-1}(U) + F_{3|Z}^{-1}(U))^2) = 2 + 2\sqrt{(1 - \rho_1^2)(1 - \rho_3^2)} + 2\rho_1\rho_3,$$

$$E((F^{-1}_{2|Z}(U) + F^{-1}_{3|Z}(U))^2) = 2 + 2\sqrt{(1-\rho_2^2)(1-\rho_3^2)} + 2\rho_2\rho_3.$$

So by Equations (3.40), (3.41), the maximum of $\mathrm{Var}(X_1 + X_2 + X_3 + X_4)$ is

$$\begin{aligned}
&4 + 2\rho_1 + 2\rho_2 + 2\rho_3 + 2 + 2\sqrt{(1-\rho_1^2)(1-\rho_2^2)} + 2\rho_1\rho_2 \\
&+ 2 + 2\sqrt{(1-\rho_1^2)(1-\rho_3^2)} + 2\rho_1\rho_3 + 2 + 2\sqrt{(1-\rho_2^2)(1-\rho_3^2)} + 2\rho_2\rho_3 - 6 \\
&= 4 + 2\rho_1 + 2\rho_2 + 2\rho_3 + 2\sqrt{(1-\rho_1^2)(1-\rho_2^2)} + 2\sqrt{(1-\rho_1^2)(1-\rho_3^2)} + 2\sqrt{(1-\rho_2^2)(1-\rho_3^2)} \\
&+ 2\rho_1\rho_2 + 2\rho_1\rho_3 + 2\rho_2\rho_3.
\end{aligned}$$

To get the minimum variance,

$$\mathrm{Var}(X_1+X_2+X_3+X_4) = E[E[(X_1+X_2+X_3+X_4)^2|X_4]] = \int_{-\infty}^{\infty} E((X_1+X_2+X_3+X_4)^2|X_4 = x_4)dF_4(x_4),$$

fixing X_4 as a constant x_4, we only need to minimize

$$E((X_1+X_2+X_3+X_4)^2|X_4 = x_4) = E((X_1|X_4 = x_4+X_2|X_4 = x_4+X_3|X_4 = x_4+X_4|X_4 = x_4)^2). \tag{3.42}$$

Since we know $X_i|X_4 = x_4 \sim N(\rho_i x_4, 1-\rho_i^2), i = 1, 2, 3$, denote

$$X_i|X_4 = x_4 := \rho_i x_4 + \sqrt{1-\rho_i^2}W_i,$$

where $W_i \sim N(0, 1)$.
Then,

$$\text{Equation (3.42)} = (\rho_1 + \rho_2 + \rho_3 + 1)^2 x_4^2 + E[(\sqrt{1-\rho_1^2}W_1 + \sqrt{1-\rho_2^2}W_2 + \sqrt{1-\rho_3^2}W_3)^2].$$

Now we minimize $E[(\sqrt{1-\rho_1^2}W_1 + \sqrt{1-\rho_2^2}W_2 + \sqrt{1-\rho_3^2}W_3)^2]$ using Proposition 3.1.7.

<u>Case 1:</u> When $\max_{1\le i\le 3} \sqrt{1-\rho_i^2} \le \frac{1}{2}\Sigma_{i=1}^3\sqrt{1-\rho_i^2}$.
$\min\{E[(\sqrt{1-\rho_1^2}W_1 + \sqrt{1-\rho_2^2}W_2 + \sqrt{1-\rho_3^2}W_3)^2]\} = 0$. So,

$$\begin{aligned}
\min\{\mathrm{Var}(X_1 + X_2 + X_3 + X_4)\} &= \int_{-\infty}^{\infty} (\rho_1 + \rho_2 + \rho_3 + 1)^2 x_4^2 dF_4(x_4) \\
&= (\rho_1 + \rho_2 + \rho_3 + 1)^2 E(X_4^2) = (\rho_1 + \rho_2 + \rho_3 + 1)^2
\end{aligned}$$

Case 2: When $\max_{1 \le i \le 3} \sqrt{1-\rho_i^2} > \frac{1}{2}\Sigma_{i=1}^3 \sqrt{1-\rho_i^2}$.
$\min\{E[(\sqrt{1-\rho_1^2}W_1 + \sqrt{1-\rho_2^2}W_2 + \sqrt{1-\rho_3^2}W_3)^2]\} = (\max_{1 \le i \le 3} \sqrt{1-\rho_i^2} - \frac{1}{2}\Sigma_{i=1}^3 \sqrt{1-\rho_i^2})^2$.
So

$$
\begin{aligned}
&\min\{\mathrm{Var}(X_1 + X_2 + X_3 + X_4)\} \\
=&\int_{-\infty}^{\infty} ((\rho_1 + \rho_2 + \rho_3 + 1)^2 x_4^2 + (\max_{1 \le i \le 3} \sqrt{1-\rho_i^2} - \frac{1}{2}\Sigma_{i=1}^3 \sqrt{1-\rho_i^2})^2) dF_4(x_4) \\
=&(\rho_1 + \rho_2 + \rho_3 + 1)^2 + (\max_{1 \le i \le 3} \sqrt{1-\rho_i^2} - \frac{1}{2}\Sigma_{i=1}^3 \sqrt{1-\rho_i^2})^2.
\end{aligned}
$$

□

3.3 Case of Non-normal Distributed Random Variables

3.3.1 Rearrangement Algorithm

Rearrangement algorithm (RA) is introduced in Puccetti and Rüschendorf [2012], here is a description on how to use it. Given a $n \times d$ matrix A as

$$
\begin{bmatrix}
a_{11} & a_{12} & \dots & a_{1d} \\
. & . & . & . \\
. & . & . & . \\
a_{n1} & a_{n2} & \dots & a_{nd}
\end{bmatrix} \tag{3.43}
$$

we rearrange the entries within each column to obtain row sums with minimal variance. (variance of $(\Sigma_{1 \le i \le d} a_{1i}, \Sigma_{1 \le i \le d} a_{2i}, ..., \Sigma_{1 \le i \le d} a_{ni})$ is the smallest.)

Step 1. For jth column, $2 \le j \le d$, order the elements oppositely to the sum of the other columns $\Sigma_{1 \le i \le d, i \ne j} a_{ij}$.

Step 2. Redo step 1 for each jth column, $1 \le j \le d$.

When there are d risks $X_i, 1 \le i \le d$ with a background risk Z which takes k values in $\{z_1, ..., z_k\}$, we can still use rearrangement algorithm to approximate the minimum of

$\text{Var}(\sum_{i=1}^d X_i)$ since

$$\text{Var}(\sum_{i=1}^d X_i) = \underbrace{E\left[\text{Var}(\sum_{i=1}^d X_i|Z)\right]}_{①} + \underbrace{\text{Var}\left[E(\sum_{i=1}^d X_i|Z)\right]}_{②} \tag{3.44}$$

Then

$$① = \sum_{j=1}^k P(Z = z_j)\text{Var}\left(\sum_{i=1}^d (X_i|Z = z_j)\right). \tag{3.45}$$

For each $Z = z_j$, we have submatrix $A_j =$

$$\begin{bmatrix} a_{11} & a_{12} & ... & a_{1d} & a_{1,d+1} \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{n1} & a_{n2} & ... & a_{nd} & a_{n+1,d+1} \end{bmatrix} \tag{3.46}$$

where $a_{i,d+1} = z_j$. If we know $X_i|Z = z_j$ with some cumulative probability function $F_{i,j}$, then for some large number n, we generate $F^{-1}(\frac{t}{n+1}), t = 1, ..., n$ as $a_{i,t}$ in the submatrix A_j. Then we approximate the minimum of $\text{Var}\left(\sum_{i=1}^d (X_i|Z = z_j)\right)$ by using RA on the submatrix A_j. Now we evaluate ②,

$$\begin{aligned} ② &= E\left(E^2(\sum_{i=1}^d X_i|Z)\right) - E^2\left(E(\sum_{i=1}^d X_i|Z)\right) \\ &= E\left(E^2(\sum_{i=1}^d X_i|Z)\right) - E^2(\sum_{i=1}^d X_i) \\ &= \sum_{j=1}^k P(Z = z_j)\left(E(\sum_{i=1}^d X_i|Z = z_j)\right)^2 - \left(\sum_{i=1}^d E(X_i)\right)^2 \\ &= \sum_{j=1}^k P(Z = z_j)\left(\sum_{i=1}^d E(X_i|Z = z_j)\right)^2 - \left(\sum_{i=1}^d E(X_i)\right)^2. \end{aligned} \tag{3.47}$$

If we know the distribution function of each $X_i|Z = z_j, i = 1, ..., d, j = 1, ..., k$ and the marginal distribution of every $X_i, i = 1, ..., d$, then ② can be computed.

3.3.2 Application with 2 Pareto Risks

We start with an example. Consider two random variables X_1, X_2 and a background risk Z such that $P(Z=3)=P(Z=4)=\frac{1}{2}$. Suppose $X_i|Z=z \sim Pareto(iz), z=3$ or $4, i=$ 1 or 2, then the minimum variance $\mathrm{Var}(X_1+X_2)$ is

$$\frac{\Gamma(\frac{2}{3})\Gamma(\frac{5}{6})}{\Gamma(\frac{3}{2})}+\frac{\Gamma(\frac{7}{8})\Gamma(\frac{3}{4})}{\Gamma(\frac{13}{8})}-\frac{490669}{176400}\approx 0.4325 \tag{3.48}$$

where Γ is the Gamma function.

Proof. (Proof of Formula (3.48))

$$\mathrm{Var}(X_1+X_2)=\underbrace{E\left[\mathrm{Var}(X_1+X_2|Z)\right]}_{①}+\underbrace{\mathrm{Var}\left[E(X_1+X_2|Z)\right]}_{②}, \tag{3.49}$$

$$\begin{aligned} ① &= P(Z=3)\mathrm{Var}(X_1|Z=3+X_2|Z=3)+P(Z=4)\mathrm{Var}(X_1|Z=4+X_2|Z=4) \\ &= \frac{1}{2}\mathrm{Var}(X_1|Z=3+X_2|Z=3)+\frac{1}{2}\mathrm{Var}(X_1|Z=4+X_2|Z=4). \end{aligned} \tag{3.50}$$

We consider the two terms separately. First when $X_1|Z=3 \sim Pareto(3)$ and $X_2|Z=3 \sim Pareto(6)$. Denote $(X_1|Z=3)=\tilde{X}_1, (X_2|Z=3)=\tilde{X}_2$, then

$$E(\tilde{X}_1)=\frac{1}{2}, E(\tilde{X}_2)=\frac{1}{5}, \mathrm{Var}(\tilde{X}_1)=\frac{3}{4}, \mathrm{Var}(\tilde{X}_2)=\frac{3}{50},$$

$$\begin{aligned} &\mathrm{Var}(\tilde{X}_1+\tilde{X}_2) \\ =&\mathrm{Var}(\tilde{X}_1)+2\mathrm{Cov}(\tilde{X}_1,\tilde{X}_2)+\mathrm{Var}(\tilde{X}_2) \\ =&\frac{3}{4}+2\left(E\left(\tilde{X}_1\tilde{X}_2\right)-\frac{1}{2}\cdot\frac{1}{5}\right)+\frac{3}{50} \\ =&\frac{61}{100}+2E\left(\tilde{X}_1\tilde{X}_2\right). \end{aligned} \tag{3.51}$$

By Lemma 2.1.1,

$$\begin{aligned} E[\tilde{X}_1\tilde{X}_2] &= \int_0^\infty\int_0^\infty P(\tilde{X}_1\geq u, \tilde{X}_2\geq v)dudv \\ &= \int_0^\infty\int_0^\infty \left(1-P(\tilde{X}_1\leq u)-P(\tilde{X}_2\leq v)+P(\tilde{X}_1\leq u,\tilde{X}_2\leq v)\right)dudv. \end{aligned} \tag{3.52}$$

To get the minimum of $E[\tilde{X_1}\tilde{X_2}]$, let $\tilde{X_1}, \tilde{X_2}$ be antimonotonic, thus $\tilde{X_1} = F^{-1}_{\tilde{X_1}}(U) = F^{-1}_{X_1|Z=3}(U), \tilde{X_2} = F^{-1}_{\tilde{X_2}}(1-U) = F^{-1}_{X_2|Z=3}(1-U)$ where $U \sim U[0,1]$ is independent of Z. If $X \sim F$ where F is the distribution function of a Pareto(θ) distribution ($\theta > 1$), then $F(x) = 1-(1+x)^{-\theta}$. Let $F(X) = U$ for some $U \sim U[0,1]$, then $X = F^{-1}(U) = (1-U)^{-\frac{1}{\theta}}-1$. Then $\tilde{X_1} = F^{-1}_{\tilde{X_1}}(U) = (1-U)^{-\frac{1}{3}}-1, \tilde{X_2} = F^{-1}_{\tilde{X_2}}(1-U) = U^{-\frac{1}{6}}-1$, so the minimum of $E[\tilde{X_1}\tilde{X_2}]$ is:

$$E\left[((1-U)^{-\frac{1}{3}}-1)(U^{-\frac{1}{6}}-1)\right] = E\left[(1-U)^{-\frac{1}{3}}U^{-\frac{1}{6}} - (1-U)^{-\frac{1}{3}} - U^{-\frac{1}{6}} + 1\right]. \tag{3.53}$$

And

$$\begin{aligned} E\left[(1-U)^{-\frac{1}{3}}U^{-\frac{1}{6}}\right] &= \int_0^1 u^{-\frac{1}{3}}(1-u)^{-\frac{1}{6}}du \\ &= \int_0^1 \frac{\Gamma(\frac{2}{3}+\frac{5}{6})}{\Gamma(\frac{2}{3})\Gamma(\frac{5}{6})}u^{\frac{2}{3}-1}(1-u)^{\frac{5}{6}-1}du \cdot \frac{\Gamma(\frac{2}{3})\Gamma(\frac{5}{6})}{\Gamma(\frac{2}{3}+\frac{5}{6})} = \frac{\Gamma(\frac{2}{3})\Gamma(\frac{5}{6})}{\Gamma(\frac{3}{2})} \end{aligned} \tag{3.54}$$

because the integral term is equal to 1 as it is the density of BETA($\frac{2}{3}, \frac{5}{6}$). Furthermore, $E\left[(1-U)^{-\frac{1}{3}}\right] = \int_0^1 (1-u)^{-\frac{1}{3}}du = \frac{3}{2}$ and $E[U^{-\frac{1}{6}}] = \int_0^1 u^{-\frac{1}{6}}du = \frac{6}{5}$. Therefore Equation (3.53) can be simplified to $\frac{\Gamma(\frac{2}{3})\Gamma(\frac{5}{6})}{\Gamma(\frac{3}{2})} - \frac{17}{10}$.
So

$$\min\left(\mathrm{Var}((X_1|Z=3)+(X_2|Z=3))\right) = \frac{2\Gamma(\frac{2}{3})\Gamma(\frac{5}{6})}{\Gamma(\frac{3}{2})} - \frac{279}{100}. \tag{3.55}$$

Second, when $X_1|Z=4 \sim Pareto(4)$ and $X_2|Z=4 \sim Pareto(8)$, then we find that

$$E(X_1|Z=4) = \frac{1}{3}, E(X_2|Z=4) = \frac{1}{7}, \mathrm{Var}(X_1|Z=4) = \frac{2}{9}, \mathrm{Var}(X_2|Z=4) = \frac{4}{147}.$$

Similarly, we get

$$\min\left(\mathrm{Var}((X_1|Z=4)+(X_2|Z=4))\right) = \frac{2\Gamma(\frac{7}{8})\Gamma(\frac{3}{4})}{\Gamma(\frac{13}{8})} - \frac{1234}{441}. \tag{3.56}$$

So by Equations (3.50), (3.55), (3.56), we obtain

$$\text{the minimum of equation ①} = \frac{\Gamma(\frac{2}{3})\Gamma(\frac{5}{6})}{\Gamma(\frac{3}{2})} + \frac{\Gamma(\frac{7}{8})\Gamma(\frac{3}{4})}{\Gamma(\frac{13}{8})} - \frac{246439}{88200}. \tag{3.57}$$

By Equation (3.47),

$$\begin{aligned}
② &= P(Z=3)\left(\sum_{i=1}^{2} E(X_i|Z=3)\right)^2 + P(Z=4)\left(\sum_{i=1}^{2} E(X_i|Z=4)\right)^2 - E^2\left(\sum_{i=1}^{2} E(X_i|Z)\right)\\
&= \frac{1}{2}\cdot\left(\frac{1}{2}+\frac{1}{5}\right)^2 + \frac{1}{2}\cdot\left(\frac{1}{3}+\frac{1}{7}\right)^2 - \left(E\left[E(X_1|Z)\right] + E\left[E(X_2|Z)\right]\right)^2\\
&= \frac{31609}{88200} - \left(E\left(\frac{1}{Z-1}\right) + E\left(\frac{1}{2Z-1}\right)\right)^2\\
&= \frac{31609}{88200} - \left(P(Z=3)\cdot\frac{1}{3-1} + P(Z=4)\cdot\frac{1}{4-1} + P(Z=3)\cdot\frac{1}{6-1} + P(Z=4)\cdot\frac{1}{8-1}\right)^2\\
&= \frac{2209}{176400}.
\end{aligned} \tag{3.58}$$

Then by Equations (3.57), (3.58),

$$\min\left(\mathrm{Var}(X_1+X_2)\right) = ① + ② = \frac{\Gamma(\frac{2}{3})\Gamma(\frac{5}{6})}{\Gamma(\frac{3}{2})} + \frac{\Gamma(\frac{7}{8})\Gamma(\frac{3}{4})}{\Gamma(\frac{13}{8})} - \frac{490669}{176400} \approx 0.4325, \tag{3.59}$$

and Equation (3.48) is proved. □

In general, it is not possible to compute these bounds explicitly. It is thus helpful to approximate them. To do so, we use the RA developed by Puccetti and Rüschendorf [2012]. We can then compare this result with the approximation from RA. Following the procedures described in Section 3.3.1, here $d=2, k=2, z_1=3, z_2=4$, the discretization size $n=10^8$, we conduct RA on each submatrix A_j,

$$\min(\mathrm{Var}((X_1|Z=3)+(X_2|Z=3))) = 0.6542,$$

$$\min(\mathrm{Var}(((X_1|Z=4)+(X_2|Z=4)))) = 0.1803,$$

so

$$\min(\mathrm{Var}(X_1+X_2)) = \frac{1}{2}\cdot 0.6542 + \frac{1}{2}\cdot 0.1803 + \frac{2209}{176400} = 0.4298, \tag{3.60}$$

which is very closed to the theoretical result in Equation (3.59).

Now we treat Pareto distribution as discrete distributions using discretization to get an approximation of the minimum variance, which is

$$
\begin{aligned}
&\min\left(\mathrm{Var}((X_1|Z=3)+(X_2|Z=3))\right)\\
=&\mathrm{Var}(F^{-1}_{X_1|Z=3}(U)+F^{-1}_{X_2|Z=3}(U))\\
\approx&\sum_{i=1}^{n}\left(F^{-1}_{X_1|Z=3}(\frac{i}{n+1})+F^{-1}_{X_2|Z=3}(1-\frac{i}{n+1})\right)^2\cdot\frac{1}{n}-\left(\left(\sum_{i=1}^{n}F^{-1}_{X_1|Z=3}(\frac{i}{n+1})+F^{-1}_{X_2|Z=3}(1-\frac{i}{n+1})\right)\cdot\frac{1}{n}\right)^2
\end{aligned}
$$

Take $n=10^6$,

$$\min\left(\mathrm{Var}((X_1|Z=3)+(X_2|Z=3))\right)\approx 0.879287-0.471526=0.6353,$$

similarly,

$$\min\left(\mathrm{Var}((X_1|Z=4)+(X_2|Z=4))\right)\approx 0.358293-0.128374=0.1790,$$

then

$$\min(\mathrm{Var}(X_1+X_2))\approx 0.40715,$$

which is slightly smaller than our result from RA (see Equation (3.60)).

3.3.3 Application with 3 Pareto Risks

There are random variables X_1, X_2, X_3 with background risk Z such that $P(Z=3)=P(Z=4)=\frac{1}{2}$. Suppose $X_i|Z=z\sim \mathit{Pareto}(iz), z=3$ or 4, then the minimum of $\mathrm{Var}(X_1+X_2+X_3)$ is 0.3803 using RA approximation.

Proof. Following the procedures described in Section 3.3.1, here $d=3, k=2, z_1=3, z_2=4$, the discretization size $n=10^6$, we conduct RA on each submatrix A_j,

$$\min(\mathrm{Var}((X_1|Z=3)+(X_2|Z=3)+(X_3|Z=3)))=0.5748,$$

$$\min(\mathrm{Var}(((X_1|Z=4)+(X_2|Z=4)+(X_3|Z=4))))=0.1526,$$

so

$$\min\left(E\left[\mathrm{Var}(X_1+X_2+X_3|Z)\right]\right)=\frac{1}{2}\cdot 0.5748+\frac{1}{2}\cdot 0.1526=0.3637.$$

Since $X_1|Z=3 \sim Pareto(3), X_2|Z=3 \sim Pareto(6), X_3|Z=3 \sim Pareto(9)$,

$$E(X_1|Z=3) = \frac{1}{2}, E(X_2|Z=3) = \frac{1}{5}, E(X_3|Z=3) = \frac{1}{8}.$$

Since $X_1|Z=4 \sim Pareto(4), X_2|Z=4 \sim Pareto(8), X_3|Z=4 \sim Pareto(12)$,

$$E(X_1|Z=4) = \frac{1}{3}, E(X_2|Z=4) = \frac{1}{7}, E(X_3|Z=4) = \frac{1}{11}.$$

$$\begin{aligned}
&\operatorname{Var}\left(E\left[X_1+X_2+X_3|Z\right]\right) \\
=&P(Z=3)\left(\sum_{i=1}^{3} E(X_i|Z=3)\right)^2 + P(Z=4)\left(\sum_{i=1}^{3} E(X_i|Z=4)\right)^2 - E^2\left(\sum_{i=1}^{3} E(X_i|Z)\right) \\
=&\frac{85567729}{170755200} - \left(E\left(\frac{1}{Z-1}\right) + E\left(\frac{1}{2Z-1}\right) + E\left(\frac{1}{3Z-1}\right)\right)^2 \\
\approx&0.016628.
\end{aligned} \tag{3.61}$$

So

$$\min\left(\operatorname{Var}(X_1+X_2+X_3)\right) = 0.3637 + 0.016628 = 0.3803.$$

□

3.4 Conclusion of Chapter 3 and Future Work

In this chapter, we give characterizations of the covariance matrix of normal distributed $(X_1, ..., X_n)$ so that $\operatorname{Var}(X_1 + ... + X_n)$ reaches minimum. An application has been done on deriving upper and lower bounds of $\operatorname{Var}(X_1 + ...X_n)$ with knowing the distribution of (X_i, Z) where Z is a background risk. $(X_i)_{1 \le i \le n}$ are normal distributed here. Only case $n = 3, 4$ are studied here.

There are some open questions remaining in this chapter:

1. Conjecture 3.1.10 is not proved.

2. All of our result can be generalized to dimension $n \ge 5$.

3. We only focus on normal distributed random variables. For non-normal ones, the RA can be used to get an approximation of the minimum variance. The compatible covariance

matrix for sum of non-normal random variables to reach minimum variance can be studied in the future. See Chaganty and Joe [2006] for related work on correlation matrix of Bernoulli random variables.

4. Our result can be applied to many areas in risk management, such as the implied correlation problem in option pricing.

Chapter 4

A New Multivariate Dependence Measure

This chapter is organized as follows, Section 4.1 is a summary of the existing bivariate dependence measures. Section 4.2 reviews some existing multivariate dependence measures. We introduce our new multivariate dependence measure ϱ in Section 4.3 with its properties (see Section 4.3.1) and estimation method (see Section 4.3.2). Section 4.4 gives a short conclusion.

4.1 Existing Bivariate Dependence Measures

This section is a review of some dependence measures of bivariate data, which can mostly be found in Nelsen [2007] and Joe [1997]. Nelsen [2007] makes a difference between measures of concordance and measures of dependence, but we ignore this subtle difference.

To measure dependence between two variables, we can use Kendall's τ, Spearman's ρ, Gini's γ or Blomqvist's β defined as follows. All (X, Y) below are continuous random variables.

Definition 4.1.1. (Kendall's τ for a vector (X, Y) of continuous random variables with joint distribution H)

$$\tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0],$$

where (X_1, Y_1) and (X_2, Y_2) are independent and identically distributed random vectors, each with joint distribution function H.

$$\tau_{X,Y} = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1,$$

where C denotes the copula of (X, Y).

Definition 4.1.2. (Spearman's ρ for vector (X, Y) with copula C)

$$\rho_{X,Y} = \rho(F_1(X_1), F_2(X_2)),$$

where ρ is the linear correlation defined as $\rho(X, Y) = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}}$.

$$\rho_{X,Y} = 12\int_0^1\int_0^1 C(u,v)dudv - 3,$$

where C denote the copula of (X, Y).

Before the following definition, we recall copulas M and W defined in Theorem 1.2.8.

Definition 4.1.3. (Gini's γ for vector (X, Y) with copula C)

$$\gamma_{X,Y} = 2\int_0^1\int_0^1 (|u+v-1| - |u-v|)dC(u,v)$$

According to Nelsen [2007], Gini's γ "measures a concordance relationship or distance between C and monotone dependence (i.e, comonotonic and countermonotonic defined in Section 1.2.3), as represented by the copulas M and W". In addition, there are measures of association based on the "distance" (such as L_p, l_p-distance) between C and M,W, (see Chapter 5.3.2 of Nelsen [2007], Conti [1993], Gideon and Hollister [1987].

Definition 4.1.4. (Blomqvist β)

$$\beta_{X,Y} = P[(X-\tilde{x})(Y-\tilde{y}) > 0] - P[(X-\tilde{x})(Y-\tilde{y}) < 0],$$

where $\tilde{x}$ and $\tilde{y}$ are medians of X and Y.

$$\beta_{X,Y} = 4C\left(\frac{1}{2}, \frac{1}{2}\right) - 1,$$

where C is the copula of (X, Y).

According to Schmid et al. [2010], Blomqvist β "can be interpreted as a normalized difference between the copula C and the independence copula at $(\frac{1}{2}, \frac{1}{2})$". Some other measures of association between two variates include Spearman's foot-rule, denoted as ϕ.

Definition 4.1.5. (Spearman's foot-rule where C is the copula of X and Y)

$$\phi_{X,Y} = 1 - 3 \int_0^1 \int_0^1 |u - v| dC(u, v)$$

By Genest et al. [2010], it is "an alternative to the correlation in the pairs $(R_1, S_1), ..., (R_n, S_n)$ of ranks associated with a random sample $(X_1, Y_1), ...(X_n, Y_n)$ from some continuous bivariate distribution $H(x, y) = P(X \leq x, Y \leq y)$".

Nelsen [2007] defines a kind of measure of association as measure of dependence on page 208, including:

Definition 4.1.6. (Schweizer and Wolff's σ)

$$\sigma_{X,Y} = 12 \int_0^1 \int_0^1 |C(u, v) - uv| dudv$$

In addition, according to Schweizer and Wolff [1981] and Nelsen [2007], the following "normalized measure of distance between the surfaces $z = C(u, v)$ and $z = uv$" yields a symmetric nonparametric measure of dependence.

Definition 4.1.7. For any p, $1 \leq p < \infty$,

$$\left(k_p \int_0^1 \int_0^1 |C(u, v) - uv|^p dudv \right)^{1/p},$$

where k_p is a constant chosen so that the above quantity is 1 when $C = W$ or M.

Nelsen [2007] gives many other examples such as $(90 \int_0^1 \int_0^1 |C(u, v) - uv|^2 dudv)^{1/2}$, $4 \sup_{u,v \in [0,1]} |C(u, v) - uv|$, but we can not present all of them.

Tail dependence refers to measuring the likelihood of observing simultaneous small or large values.

Definition 4.1.8. (Upper tail dependence λ_U and lower tail dependence λ_L)

$$\lambda_U = \lim_{t \to 1^-} P[Y > G^{-1}(t) | X > F^{-1}(t)],$$

$$\lambda_L = \lim_{t \to 0^+} P[Y \le G^{-1}(t) | X \le F^{-1}(t)],$$

where X and Y are continuous random variables with distribution functions F and G, respectively.

We end this section with the informational measure of dependence from Linfoot [1957], although the list of dependence measures is far from being exhaustive.

Definition 4.1.9. (Informational coefficient of correlation r_1)

$$r_1 = \sqrt{1 - e^{-2r_0}},$$

where r_0 is defined as follows,

$$r_0 = \int\int (p(x,y)\log p(x,y) - p(x)q(y)\log[p(x)q(y)])dxdy,$$

where X, Y are continuous variables with joint probability density distribution $p(x,y)$ and $p(x), q(y)$ are the probability density distributions of X, Y.

By Linfoot [1957], it is invariant under a change of parameterization $X' = f(X), Y' = g(Y)$. All the measures of associations here are defined when (X, Y) are continuous, see Mesfioui and Quessy [2010] for the situation when (X, Y) are not continuous.

4.2 Existing Multivariate Dependence Measures

4.2.1 Multivariate Dependence Measures Depending on the Copula Only

All the measures of association mentioned in Section 4.1 have different versions of multivariate generalizations, we list some of them mentioned in Schmid et al. [2010].

Definition 4.2.1. (Generalization of Kendall's tau by Joe [1990])

$$\tau(\mathbf{X}) = \Sigma_{k=d'}^{d} w_k P\{(D_1, ..., D_d) \in B_{k,d-k}\}$$

with $d' = \lfloor (d+1)/2 \rfloor$ and $B_{k,d-k}$ being the subset of $\mathbf{x} = (x_1, ..., x_d)$ in $\mathbb{R}^d$ with k positive components and $d-k$ negative or k negative components and $d-k$ positive. There are some additional constraints about w_k given in Joe [1990]. τ's representation given in terms of the copula is given by

$$\tau(C) = \frac{1}{2^{d-1}-1} \left\{ 2^d \int_{[0,1]^d} C(\mathbf{u}) dC(\mathbf{u}) - 1 \right\}.$$

Some other multivariate versions of Kendall's tau can be found in Nelsen [1996], Nelsen [2002] and Taylor [2007].

Definition 4.2.2. (Generalizations of Spearman's rho by Schmid and Schmidt [2007])

$$\rho_1(C) = \frac{\int_{[0,1]^d} C(\mathbf{u}) d\mathbf{u} - \int_{[0,1]^d} \prod(\mathbf{u}) d\mathbf{u}}{\int_{[0,1]^d} M(\mathbf{u}) d\mathbf{u} - \int_{[0,1]^d} \prod(\mathbf{u}) d\mathbf{u}} = h_\rho(d) \left\{ 2^d \int_{[0,1]^d} C(\mathbf{u}) d\mathbf{u} - 1 \right\},$$

where $h_\rho(d) = (d+1)/\{2^d - (d+1)\}$.
Another version is

$$\rho_2(C) = h_\rho(d) \left\{ 2^d \int_{[0,1]^d} \prod(\mathbf{u}) dC(\mathbf{u}) - 1 \right\},$$

where $\prod(\mathbf{u}) = \prod_{i=1}^{n} u_i$ denotes the independent copula.

Nelsen [1996] proposes another version as $\rho_3 = (\rho_1 + \rho_2)/2$. Some other multivariate extensions can be found in Ruymgaart and van Zuijlen [1978], Wolff [1980] and Stepanova [2003].

Definition 4.2.3. (Generalization of Blomqvist's beta)

$$\begin{aligned} \beta(C) &= \frac{C(\mathbf{1/2}) - \prod(\mathbf{1/2}) + \overline{C}(\mathbf{1/2}) - \overline{\prod}(\mathbf{1/2})}{M(\mathbf{1/2}) - \prod(\mathbf{1/2}) + \overline{M}(\mathbf{1/2}) - \overline{\prod}(\mathbf{1/2})} \\ &= h_\beta(d)\{C(\mathbf{1/2}) + \overline{C}(\mathbf{1/2}) - 2^{1-d}\}, \end{aligned} \tag{4.1}$$

where $\overline{C}$ denotes the survival function, $\overline{C}(\mathbf{u}) = P(\mathbf{U} \geq \mathbf{u})$, $h_\beta(d) := \frac{2^{d-1}}{2^{d-1}-1}$ and $\mathbf{1/2} :=$

$(1/2, ..., 1/2)$.

Definition 4.2.4. (Generalization of Gini's gamma by Behboodian et al. [2007])

$$\gamma(C) = \frac{1}{b(d) - a(d)} \left[\int_{[0,1]^d} \{A(\mathbf{u}) + \overline{A}(\mathbf{u})\} dC(\mathbf{u}) - a(d) \right],$$

where normalization constants $a(d)$ and $b(d)$ are of the form:

$$a(d) = \int_{[0,1]^d} \left\{ A(\mathbf{u}) + \overline{A}(\mathbf{u}) \right\} du_1 ... du_d = \frac{1}{d+1} + \frac{1}{2(d+1)!} + \sum_{i=0}^{d} (-1)^i \binom{d}{i} \frac{1}{2(i+1)!},$$

and

$$b(d) = \int_{[0,1]^d} \{A(\mathbf{u}) + \overline{A}(\mathbf{u})\} dM(\mathbf{u}) = 1 - \sum_{i=1}^{d-1} \frac{1}{4i},$$

where $A(\mathbf{u}) = \{M(\mathbf{u}) + W(\mathbf{u})\}/2, \mathbf{u} \in [0,1]^d$ and $\overline{A}$ denotes the survival function of A.

According to Schmid et al. [2010], Joe [1989] introduces relative entropy (also known as Kullback-Leibler divergence, see Kullback and Leibler [1951], Kullback [1997]) as an information-based measure of association, defined as follows.

Definition 4.2.5. (Relative entropy of a random vector $\mathbf{X} = (X_1, ..., X_d)$ with copula C)

$$\delta(\mathbf{X}) = \int_{\mathbb{R}^d} \log \left[\frac{f(\mathbf{x})}{\prod_{i=1}^d f_i(x_i)} \right] f(\mathbf{x}) d\mathbf{x},$$

where f is the density of the distribution of $\mathbf{X}$ and $(f_i)_{1 \leq i \leq d}$ are the marginal distributions.

$$\delta(\mathbf{X}) = \delta(C) = \int_{[0,1]^d} \log[c(\mathbf{u})] c(\mathbf{u}) d\mathbf{u},$$

where c is the density of the copula C.

4.2.2 Multivariate Dependence Measures Depending on the Multivariate Distribution

Some other multivariate dependence measures depending on the multivariate distribution include:

Definition 4.2.6. (From Dhaene et al. [2014], dependence measure ρ_c of a random vector $\mathbf{X}$ with non-degenerate margins)

$$\rho_c(\mathbf{X}) = \frac{\text{Var}(S) - \text{Var}(S^{\perp})}{\text{Var}(S^c) - \text{Var}(S^{\perp})} = \frac{\Sigma_{i=1}^d \Sigma_{j<i} \text{Cov}(X_i, X_j)}{\Sigma_{i=1}^d \Sigma_{j<i} \text{Cov}(X_i^c, X_j^c)},$$

where $\mathbf{X}^c = (X_1^c, X_2^c, ..., X_d^c)$ is a random vector with the same marginal distributions as $\mathbf{X}$ but with comonotonic components. $S^c = \sum_{i=1}^d X_i^c$, and $S^{\perp} = \sum_{i=1}^d X_i^{\perp}$, where $\mathbf{X}^{\perp} = (X_1^{\perp}, X_2^{\perp}, ..., X_d^{\perp})$ is a random vector with the same marginal distributions as $\mathbf{X}$ but with independent components.

By Dhaene et al. [2014], ρ_c measures dependence in $\mathbf{X}$ indirectly through the distribution of the sum S of its components, focusing on the aggregate risk rather than on the copula or joint distribution function itself.

The following dependence measure based on comonotonicity by means of product moment is motivated by ρ_c.

Definition 4.2.7. (From Zhang and Yin [2014]) The dependence measure $\rho(\mathbf{X})$ of a random vector $\mathbf{X}$ with non-degenerate margins is defined as

$$\rho(\mathbf{X}) = \frac{E[\prod_{i=1}^d X_i] - \prod_{i=1}^d E[X_i]}{E[\prod_{i=1}^d X_i^C] - \prod_{i=1}^d E[X_i]},$$

provided the expectations exist, where X^c is a comonotonic random vector.

4.3 A New Multivariate Dependence Measure Depending on the Multivariate Distribution

Now we introduce a new multivariate dependence measure, denoted as ϱ in the Definition 4.3.1 below. It also focuses on the aggregate risk just as ρ_c in Definition 4.2.6.

Definition 4.3.1. (New multivariate dependence measure ϱ)

$$\varrho(\mathbf{X}) = \frac{\sum\limits_{\text{all partitions S}} \phi(\sum\limits_{i \in S} \mathbf{X_i}, \sum\limits_{i \in \overline{S}} \mathbf{X_i})}{\text{number of ways to partition } S \text{ into 2 parts}},$$

where ϕ is some existing 2-dimensional dependence measure. In particular, given a matrix of data $\mathbf{X} = (X_1, X_2, ..., X_{n-1}, X_n)$, divide $\{1, ..., n\}$ into 2 parts, $\{1, ..., n\} = S \bigcup \overline{S}, S \bigcap \overline{S} = \emptyset$. We denote by $\mathcal{C}$ the set that consists of all these partitions such that neither S nor $\overline{S}$ is empty,

$$\varrho(\mathbf{X}) = \frac{1}{2^n - 2} \sum_{S \in \mathcal{C}} \phi(\sum_{i \in S} \mathbf{X_i}, \sum_{i \in \overline{S}} \mathbf{X_i}).$$

Note that each partition appears twice in the summation.

4.3.1 Properties

We now introduce its properties. We assume that the 2-dimensional dependence measure ϕ in Definition 4.3.1 is a concordance measure and recall its definition from Lee and Ahn [2014].

Definition 4.3.2. (2-dimensional measure of concordance)
Let $\mathbf{X}$ and $\mathbf{X}^*$ be bivariate random vectors with distribution functions $H = C(F_1, F_2)$ and $H^* = C^*(F_1^*, F_2^*)$, respectively. A numeric measure $\kappa(H)$ (or $\kappa(\mathbf{X})$) is a *measure of concordance* if it satisfies the following axioms. (Here $X = (X_1, X_2), X^* = (X_1^*, X_2^*)$ are 2-variate random vectors and F_i, F_i^* are the marginals for $X_i, X_i^*, i = 1, 2$, C, C^* are the copulas for $(\mathbf{X}, \mathbf{X}^*)$. See Definition 1.2.1 for definition of a copula.)
$[S1]$. $-1 \leq \kappa(H) \leq 1$, $\kappa(M(F_1, F_2)) = 1$, $\kappa(\prod(F_1, F_2)) = 0$, and $\kappa(W(F_1, F_2)) = -1$ where M, W are defined in Theorem 1.2.8, $\prod$ is the independence copula.
$[S2]$. If $C \prec C^*$(e.g, C is smaller than C^* in concordance order, $C(\mathbf{u}) \leq C^*(\mathbf{u})$ for all $\mathbf{u} \in [0, 1]^2$.), then $\kappa(H) \leq \kappa(H^*)$.
$[S3]$. If a sequence of joint distribution functions, $\{H_1, H_2, ...\}$, converges pointwise to H, then $\lim_{i\to\infty} \kappa(H_i) = \kappa(H)$;
$[S4]$. $\kappa(X_1, X_2) = \kappa(X_2, X_1)$[1];
$[S5]$. $\kappa(C_1^{\#}(F_1, F_2)) = \kappa(C_2^{\#}(F_1, F_2)) = -\kappa(H)$ where $C_i^{\#}$ is the copula associated with a random vector $(X_1, ..., X_{i-1}, -X_i, X_{i+1}, ..., X_d)$.

Taylor [2010] extends Definition 4.3.2 to a multivariate measure of concordance.

Definition 4.3.3. (Multivariate measure of concordance)
Let $\mathbf{X}$ and $\mathbf{X}^*$ be d-variate random vectors with distribution functions $H = C(F_1, ..., F_d)$

[1]Observe that it is a permutation of $\{1, 2\}$.

and $H^* = C^*(F_1^*, ..., F_d^*)$, respectively. A numeric measure $\kappa(H)$ (or $\kappa(\mathbf{X})$) is a *measure of concordance* if it satisfies the following axioms.
$[A1]$. (Normalization) $\kappa_d(H) \leq 1$, $\kappa_d(M(F_1, ..., F_d)) = 1$ and $\kappa_d(\prod(F_1, ..., F_d)) = 0$;
$[A2]$. (Monotonicity) If $(X_1, ..., X_d) \prec (Y_1, ..., Y_d)$, then $\kappa_d(X_1, ..., X_d) \leq \kappa_d(Y_1, ..., Y_d)$;
$[A3]$. (Continuity) If F_k is the joint distribution function of the random vector $(X_{k1}, ..., X_{kd})$ and F is the distribution function for $(X_1, ..., X_d)$. If $F_k \to F$, then $\kappa_d(X_{k1}, ..., X_{kd}) \to \kappa_d(X_1, ..., X_d)$;
$[A4]$. (Permutation Invariance) If $(i_1, ..., i_d)$ is a permutation of $(1, ..., d)$, then $\kappa_d(X_{i_1}, ..., X_{i_d}) = \kappa(X_1, ..., X_d)$;
$[A5]$. (Duality) $\kappa_d(-X_1, ..., -X_d) = \kappa_d(X_1, ..., X_d)$;
$[A6]$. (Reflection Symmetry Property; RSP) $\sum_{\varepsilon_1, ..., \varepsilon_d} \kappa_d(\varepsilon_1 X_1, ..., \varepsilon_d X_d) = 0$ where each $\varepsilon_i = \pm 1$ and the sum is over all possible combinations of ± 1;
$[A7]$. (Transition Property; TP) There exists a sequence of numbers $\{r_d\}$, where $d \geq 2$, such that for every n-tuple of continuous random variables $(X_1, ..., X_d)$, we have

$$r_{d-1}\kappa_{d-1}(X_2, ..., X_d) = \kappa_d(X_1, X_2, ..., X_d) + \kappa_d(-X_1, X_2, ..., X_d).$$

To prove Proposition 4.3.7, we need the following propositions.

Proposition 4.3.4. *(Recalled from p199 of Embrechts et al. [2005])*
$X_1, ..., X_d$ are comonotonic if and only if

$$(X_1, ..., X_d) =^d (v_1(Z), ..., v_d(Z)) \tag{4.2}$$

for some random variables Z and increasing functions $v_1, ..., v_d$, where notation " $=^d$ " means both sides of the equation have the same distributions.

Proposition 4.3.5. *Given random vectors $\mathbf{X} = (X_1, ..., X_d)$, if each component is comonotonic, then for a fixed partition $S \cup \overline{S} = \{1, ..., d\}$, $\sum_{i \in S} X_i$ and $\sum_{i \in \overline{S}} X_i$ are still comonotonic.*

Proof. For $U \sim U[0, 1]$, since $X_1, ..., X_d$ are comonotonic, $(X_1, ..., X_d) =^d (F_{X_1}^{-1}(U), ..., F_{X_d}^{-1}(U))$, then

$$\sum_{i \in S} X_i =^d \sum_{i \in S} F_{X_i}^{-1}(U) = g_1(U), \tag{4.3}$$

where $g_1(\cdot) := \sum_{i\in S} F_{X_i}^{-1}(\cdot)$ is increasing, and

$$\sum_{i\in\overline{S}} X_i =^d \sum_{i\in\overline{S}} F_{X_i}^{-1}(U) = g_2(U) = g_2 \circ g_1^{-1}(\sum_{i\in S} X_i), \tag{4.4}$$

where $g_2(\cdot) := \sum_{i\in\overline{S}} F_{X_i}^{-1}(\cdot)$ is increasing. Since $g_2 \circ g_1^{-1}$ is still increasing, by Proposition 4.3.4, $\sum_{i\in S} X_i$ and $\sum_{i\in\overline{S}} X_i$ are comonotonic. □

Proposition 4.3.6. *If random vectors* $\mathbf{X}$ *and* $\mathbf{Y}$ *satisfy*

$$\mathbf{X} = (X_1, ..., X_d) \prec \mathbf{Y} = (Y_1, ..., Y_d), \tag{4.5}$$

then for a fixed partition $S \cup \overline{S}$,

$$\mathbf{A} =: \left(\sum_{i\in S} X_i, \sum_{j\in\overline{S}} X_j\right) \prec \left(\sum_{i\in S} Y_i, \sum_{j\in\overline{S}} Y_j\right) := \mathbf{B}.$$

Proof. For convenience, denote the cardinality $|S| = p, |\overline{S}| = q$ where $p + q = d$. We list

$$S = \{i_1, ..., i_p\}, \overline{S} = \{j_1, ..., j_q\}.$$

By Equation (4.5),

$$\frac{\partial^d}{\partial t_1 \partial t_2 ... \partial t_d} F_{X_1,...,X_d}(t_1, ..., t_d) \leq \frac{\partial^d}{\partial t_1 \partial t_2 ... \partial t_d} F_{Y_1,...,Y_d}(t_1, ..., t_d),$$

so the joint density functions of $\mathbf{X}$ and $\mathbf{Y}$ satisfy,

$$f_{X_1,...,X_d}(t_1, ..., t_d) \leq f_{Y_1,...,Y_d}(t_1, ..., t_d),$$

for all $(t_1, ..., t_d) \in \mathbb{R}^d$.

Now for any fixed pairs of $(x, y) \in \mathbb{R}^2$,

$$
\begin{aligned}
&P(\sum_{i\in S} X_i \le x, \sum_{j\in \overline{S}} X_j \le y)\\
=&P(X_{i_1} \le x - \sum_{i\in S\setminus\{i_1\}} X_i, X_{j_1} \le y - \sum_{j\in \overline{S}\setminus\{j_1\}} X_j)\\
=&\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \int_{-\infty}^{y-\sum_{j\in\overline{S}\setminus\{j_1\}} x_j} \int_{-\infty}^{x-\sum_{i\in S\setminus\{i_1\}} x_i} f_{X_1,...,X_d}(x_1, ..., x_d) dx_{j_1} dx_{i_1} ... dx_{i_p} dx_{j_q}\\
\le&\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \int_{-\infty}^{y-\sum_{j\in\overline{S}\setminus\{j_1\}} x_j} \int_{-\infty}^{x-\sum_{i\in S\setminus\{i_1\}} x_i} f_{Y_1,...,Y_d}(x_1, ..., x_d) dx_{j_1} dx_{i_1} ... dx_{i_p} dx_{j_q}\\
=&P(\sum_{i\in S} Y_i \le x, \sum_{j\in \overline{S}} Y_j \le y).
\end{aligned}
$$

Similarly,

$$
P(\sum_{i\in S} X_i > x, \sum_{j\in \overline{S}} X_j > y) \le P(\sum_{i\in S} Y_i > x, \sum_{j\in \overline{S}} Y_j > y).
$$

So $\mathbf{A} \prec \mathbf{B}$. □

Proposition 4.3.7. *In Definition 4.3.1, given a d-variate random vector* $\mathbf{X} = (X_1, ..., X_d)$, *when* $d = 2$, $\varrho_d(\mathbf{X}) = \phi(\mathbf{X})$; *when* $d \ge 3$, *if* ϕ *is a measure of concordance, then* ϱ_d *is not necessarily a measure of concordance. In particular,* ϱ *satisfies axioms A1-A6 but not A7.*

Proof. When $d = 2$,

$$
\varrho_2\left((X_1, X_2)\right) = \frac{\phi(X_1, X_2) + \phi(X_2, X_1)}{2} = \phi(X_1, X_2).
$$

When $d \ge 3$, we will check A1-A7.

[*A*1]. $\varrho_d(\mathbf{X}) \le 1$ since each $\phi(\sum_{i\in S} X_i, \sum_{i\in\overline{S}} X_i) \le 1$.
If $\mathbf{X}$ has comonotonic components, for each partition $S \cup \overline{S} = \{1, ..., d\}$, by Proposition 4.3.5, $\sum_{i\in S} X_i$ and $\sum_{i\in\overline{S}} X_i$ are still comonotonic Then each $\phi(\sum_{i\in S} X_i, \sum_{i\in\overline{S}} X_i) = 1$, so $\varrho_d(M(F_1, ..., F_d)) = 1$. Similarly, $\varrho_d(\prod(F_1, ..., F_d)) = 0$ since when $(X_1, ..., X_d)$ is independent, $\sum_{i\in S} X_i$ and $\sum_{i\in\overline{S}} X_i$ are independent.

$[A2]$. If $(X_1, ..., X_d) \prec (Y_1, ..., Y_d)$, then for each partition $S \cup \overline{S}$, by Proposition 4.3.6,

$$\left(\sum_{i\in S} X_i, \sum_{j\in\overline{S}} X_j\right) \prec \left(\sum_{i\in S} Y_i, \sum_{j\in\overline{S}} Y_j\right),$$

so

$$\phi\left(\sum_{i\in S} X_i, \sum_{j\in\overline{S}} X_j\right) \leq \phi\left(\sum_{i\in S} Y_i, \sum_{j\in\overline{S}} Y_j\right)$$

by the monotonicity of ϕ. Then $\varrho_d(X_1, ..., X_d) \leq \varrho_d(Y_1, ..., Y_d)$.

$[A3]$. Consider a sequence of random vectors $\{(X_{k1}, ..., X_{kd})\}_{k\in\mathbb{N}}$ and $(X_1, ..., X_d)$ with joint distribution functions F and F_k such that $F_k \to F$, this also means $(X_{k1}, ..., X_{kd}) \to^d (X_1, ..., X_d)$. Now for each fixed partition $S \cup \overline{S}$, define function $g_S : \mathbb{R}^d \to \mathbb{R}^2$:

$$g_S(X_1, ..., X_d) = (\sum_{i\in S} X_i, \sum_{i\in\overline{S}} X_i).$$

Clearly, g_S is continuous. By continuous mapping theorem (see Theorem 2.3 of Van der Vaart [2000]),

$$g_S(X_{k1}, ..., X_{kd}) \to^d g_S(X_1, ..., X_d),$$

then by continuity of ϕ,

$$\phi(\sum_{i\in S} X_{ki}, \sum_{i\in\overline{S}} X_{ki}) \to^d \phi(\sum_{i\in S} X_i, \sum_{i\in\overline{S}} X_i),$$

for every S. So

$$\varrho(X_{k1}, ..., X_{kd}) \to^d \varrho(X_1, ..., X_d).$$

$[A4]$. If $(i_1, ..., i_d)$ is a permutation of $(1, ..., d)$, then clearly $\varrho(X_{i_1}, ..., X_{i_d}) = \varrho(X_1, ..., X_d)$ since we still get the same set of partitions after permutation.

[$A5$]. By the duality of ϕ,

$$\begin{aligned}\varrho(-X_1, ..., -X_d) =& \frac{1}{2^n-2}\sum_{S\in\mathbf{C}}\phi(-\sum_{i\in S}X_i, -\sum_{i\in\overline{S}}X_i)\\ =& \frac{1}{2^n-2}\sum_{S\in\mathcal{C}}\phi(\sum_{i\in S}X_i, \sum_{i\in\overline{S}}X_i) = \varrho(X_1, ..., X_d).\end{aligned}$$

[$A6$].

$$\sum_{\varepsilon_1,...,\varepsilon_d}\varrho(\varepsilon_1X_1, ..., \varepsilon_dX_d) = \frac{1}{2^n-2}\sum_{\varepsilon_1,...,\varepsilon_d}(\sum_{S\in\mathcal{C}}\phi(\sum_{i\in S}\varepsilon_iX_i, \sum_{i\in\overline{S}}\varepsilon_iX_i)),$$

there is an even number of terms in the sum. So for a fixed set of $\{\varepsilon_1, ..., \varepsilon_d\}$ and each S, we can pair $\phi(\sum_{i\in S}\varepsilon_iX_i, \sum_{i\in\overline{S}}\varepsilon_iX_i)$ with $\phi(\sum_{i\in S}-\varepsilon_iX_i, \sum_{i\in\overline{S}}\varepsilon_iX_i)$, by ϕ's property (see S5 of Definition 4.3.2), we get

$$\phi(\sum_{i\in S}\varepsilon_iX_i, \sum_{i\in\overline{S}}\varepsilon_iX_i) + \phi(\sum_{i\in S}-\varepsilon_iX_i, \sum_{i\in\overline{S}}\varepsilon_iX_i) = 0.$$

So $\sum_{\varepsilon_1,...,\varepsilon_d}\varrho(\varepsilon_1X_1, ..., \varepsilon_dX_d) = 0$.

A7 clearly does not hold for different ϕs (since then ϱ is a completely different dependence measure). For fixed ϕ, A7 still does not hold, see the counterexample in Example 4.3.8. □

Example 4.3.8. Given $d = 3$, assume (Y, Y, Y) has marginals $U[0, 1]$. (X_1, X_2, X_2) has marginals $exp(1)$, where X_1 and X_2 are independent. Assuming A7 holds, we can get a contradiction.

Proof. We pick $\phi = \rho_s$, the Spearman's rho, which is a measure of concordance satisfying

A1-A7. Assuming A7 holds, then for the 3-tuple of continuous random variable (Y, Y, Y),

$$\begin{aligned}
r_2\varrho_2(Y,Y) &= \varrho_3(Y,Y,Y) + \varrho_3(-Y,Y,Y) \\
r_2 &= \frac{1}{2^3-2}\left[2\rho_s(Y,2Y) + 2\rho_s(Y,2Y) + 2\rho_s(Y,2Y)\right] \\
&+ \frac{1}{2^3-2}\left[2\rho_s(-Y,2Y) + 2\rho_s(Y,0) + 2\rho_s(Y,0)\right] \\
r_2 &= \rho_s(Y,2Y) + \frac{1}{3}\left[\rho_s(-Y,2Y) + 2\rho_s(Y,0)\right].
\end{aligned} \tag{4.6}$$

Y and 0 are independent, so $\rho_s(Y,0) = 0$, $\rho_s(-Y,2Y) = -\rho_s(Y,2Y)$ (This is by A5 and A6 and it only works in dimension 2). By the fact that Y and $2Y$ are comonotonic, we get $\rho_s(Y,2Y) = 1$, so $r_2 = \frac{2}{3}$.
Then we take another pair of 3-tuple random variables (X_1, X_2, X_2), if A7 holds, this r_2 should stay the same.

$$\begin{aligned}
r_2\varrho_2(X_2,X_2) &= \varrho_3(X_1,X_2,X_2) + \varrho_3(-X_1,X_2,X_2) \\
r_2 &= \frac{1}{3}\left[\rho_s(X_1,2X_2) + \rho_s(X_2,X_1+X_2) + \rho_s(X_2,X_1+X_2) + \rho_s(-X_1,2X_2)\right] \\
&+ \frac{1}{3}\left[\rho_s(X_2,-X_1+X_2) + \rho_s(X_2,-X_1+X_2)\right] \\
r_2 &= \frac{2}{3}\left[\rho_s(X_2,X_1+X_2) - \rho_s(X_2,X_1-X_2)\right].
\end{aligned} \tag{4.7}$$

Then we take (X_2, X_1, X_2),

$$\begin{aligned}
r_2\varrho(X_1,X_2) &= \varrho(X_2,X_1,X_2) + \varrho(-X_2,X_1,X_2) \\
0 &= \frac{1}{3}[\rho_s(X_2,X_1+X_2) + \rho_s(X_1,2X_2) + \rho_s(X_2,X_1+X_2) + \rho_s(-X_2,X_1+X_2) \\
&+ \rho_s(X_1,0) + \rho_s(X_2,X_1-X_2)] \\
-\rho_s(X_2,X_1-X_2) &= \rho_s(X_2,X_1+X_2).
\end{aligned} \tag{4.8}$$

So Equation (4.7) becomes

$$\begin{aligned} r_2 &= \frac{4}{3}\rho_s(X_2, X_1+X_2) \\ &= \frac{4}{3}\rho\left(F_{X_1}(X_1), F_{X_2+X_3}(X_2+X_3)\right) \\ &= \frac{4}{3}\rho(1-e^{-X_1}, 1-e^{-(X_1+X_2)}(1+X_1+X_2)). \end{aligned} \tag{4.9}$$

Since both $F_{X_1}(X_1)$ and $F_{X_2+X_3}(X_2+X_3)$ follow $U[0,1]$,

$$E\left(F_{X_1}(X_1)\right) = E\left(F_{X_2+X_3}(X_2+X_3)\right) = \frac{1}{2},$$

$$\mathrm{Var}\left(F_{X_1}(X_1)\right) = \mathrm{Var}\left(F_{X_2+X_3}(X_2+X_3)\right) = \frac{1}{12},$$

and X_1, X_2 has moment generating function

$$E(e^{tX_1}) = E(e^{tX_2}) = \frac{1}{1-t},$$

for $t < 1$.
So

$$E(e^{-X_1}) = E(e^{-X_2}) = \frac{1}{2}, E(e^{-2X_1}) = \frac{1}{3},$$

and

$$E(X_1e^{-2X_1}) = \int_0^\infty x_1e^{-3x_1}dx_1 = \frac{1}{9},$$

$$E(X_1e^{-X_1}) = \int_0^\infty x_1e^{-x_1}dx_1 = \frac{1}{4},$$

$$\begin{aligned} E(F_{X_1}(X_1)F_{X_1+X_2}(X_1+X_2) &= E[1-e^{-X_1}+e^{-2X_1-X_2}-e^{-X_1-X_2}+ \\ &\quad e^{-2X_1-X_2}X_1 - e^{-X_1-X_2}X_1 + e^{-2X_1-X_2}X_2 - e^{-X_1-X_2}X_2] \\ &= \frac{11}{36}. \end{aligned} \tag{4.10}$$

So by Equation (4.9),

$$\begin{aligned} r_2 &= \frac{4}{3} \cdot \frac{E(F_{X_1}(X_1)F_{X_1+X_2}(X_1+X_2)) - E(F_{X_1}(X_1)E(F_{X_1+X_2}(X_1+X_2))}{\sqrt{\mathrm{Var}(F_{X_1}(X_1))\mathrm{Var}(F_{X_1+X_2}(X_1+X_2))}} \\ &= \frac{4}{3} \cdot \frac{\frac{11}{36} - \frac{1}{2}\cdot\frac{1}{2}}{\sqrt{\frac{1}{12}\cdot\frac{1}{12}}} = \frac{8}{9}, \end{aligned} \tag{4.11}$$

which is different from the r_2 we get from Equation (4.6) ($r_2 = \frac{2}{3}$). There is a contradiction, so A7 can not hold. □

Remark 4.3.9. Note by Proposition 4.3.5 and property A1 of ϱ, given $\mathbf{X}$, ϱ can be written as

$$\varrho(\mathbf{X}) = \frac{\sum_{S\in\mathcal{C}} \phi\left(\sum_{i\in S} X_i, \sum_{i\in\overline{S}} X_i\right)}{\sum_{S\in\mathcal{C}} \phi\left(\sum_{i\in S} X_i^c, \sum_{i\in\overline{S}} X_i^c\right)},$$

where $\mathbf{X}^c = (X_1^c, ..., X_d^c)$ is a comonotonic random vector and $S, \mathcal{C}$ are defined in Definition 4.3.1.

Definition 4.3.10. (d-countermonotonic from Lee and Ahn [2014]) A d-variate random vector $\mathbf{X}$ is called d-countermonotonic with $(f_1, ..., f_d)$ if there exist non-decreasing continuous functions $f_1, ..., f_d$ on $\mathbb{R}$ and a support, B, of $\mathbf{X}$, which satisfy the following conditions:
i. $f_1(s_1), ..., f_d(s_d)$ are strictly increasing functions at $(s_1, ..., s_d) = (x_1, ..., x_d)$,
ii. $\sum_{j=1}^d f_i(x_i) = 1$,
for any $(x_1, ..., x_d) \in B$.

We give a counterexample in Example 4.3.11 that d-countermonotonic is not the necessary condition to ensure $\varrho(\mathbf{X}) = -1$.

Example 4.3.11. If $X_1 \sim Pareto(2), X_2, X_3 \sim Pareto(1)$, denote $U_i = F_i(X_i), i = 1, 2, 3,$ and

$$U_1 + U_2 + U_3 = c$$

for some constant c, then X_1, X_2, X_3 are 3-countermonotonic. But $X_1 + X_2$ and X_3 are not 2-countermonotonic (anticomonotonic).

Proof. First, note that U_1 and U_2 can be independent, for instance, take $U_1 = U, U_2 = \tilde{U}, U_3 = 1 - U - \tilde{U}$ where U and $\tilde{U}$ are independent. We assume U_1 and U_2 are independent in this example and $c = 1$.
Then,

$$X_1 + X_2 = F_1^{-1}(U_1) + F_2^{-1}(U_2) = (1 - U_1)^{-1/2} + (1 - U_2)^{-1} - 2,$$
$$X_3 = F_3^{-1}(1 - U_1 - U_2) = (U_1 + U_2)^{-1} - 1.$$

If $X_1 + X_2$ and X_3 are independent, then there are f_1, f_2 strictly increasing on the support set such that

$$f_1\left((1 - U_1)^{-1/2} + (1 - U_2)^{-1} - 2\right) + f_2\left((U_1 + U_2)^{-1} - 1\right) = c_2 \tag{4.12}$$

for some constant c_2. Define $\tilde{f}_1(x) = f_1(x - 2), \tilde{f}_2(x) = f_2(x - 1) - c_2$, where x belongs to the support set. Then $\tilde{f}_1, \tilde{f}_2$ are still strictly increasing. Thus Equation (4.12) becomes

$$\tilde{f}_1\left((1 - U_1)^{-1/2} + (1 - U_2)^{-1}\right) = -\tilde{f}_2\left((U_1 + U_2)^{-1}\right). \tag{4.13}$$

Since $\tilde{f}_2$ has inverse function $\tilde{f}_2^{\,-1}$ on its support set, Equation (4.13) becomes,

$$\underbrace{(-\tilde{f}_2)^{-1} \circ \tilde{f}_1}_{:=g}\left((1 - U_1)^{-1/2} + (1 - U_2)^{-1}\right) = (U_1 + U_2)^{-1}.$$

Now we take 2 pairs of $(U_1, U_2) = \left(\frac{3}{4}, \frac{2}{3}\right), (\tilde{U_1}, \tilde{U_2}) = \left(\frac{8}{9}, \frac{1}{2}\right)$, then

$$(1 - U_1)^{-1/2} + (1 - U_2)^{-1} = 5 = (1 - \tilde{U_1})^{-1/2} + (1 - \tilde{U_2})^{-1}.$$

However,

$$(U_1 + U_2)^{-1} = \frac{12}{17} \neq (\tilde{U_1} + \tilde{U_2})^{-1} = \frac{18}{25}.$$

So function g is not well-defined. There are no such functions f_1, f_2. $\square$

Thus when we take ϕ as a measure of concordance, $\phi(X_1 + X_2, X_3) > -1$, so $\varrho(\mathbf{X}) > -1$ here. Now we give a sufficient condition to ensure $\varrho(\mathbf{X}) = -1$.

Definition 4.3.12. ($\Sigma -$ *countermonotonic* from Puccetti and Wang [2014])
A random vector $\mathbf{X}$ is said to be $\Sigma -$ *countermonotonic* if for any vector $\mathbf{a} \in \{0, 1\}^d$, $\mathbf{a} \cdot \mathbf{X}$

and $(\mathbf{1}-\mathbf{a})\cdot\mathbf{X}$ are countermonotonic.

Proposition 4.3.13. $\mathbf{X}$ *is* $\Sigma -$ *countermonotonic, if and only if* $\varrho(\mathbf{X}) = -1$.

Proof. "$\Rightarrow$" For any partition $S \cup \overline{S} = \{1, ..., d\}$, take $\mathbf{a} = (a_1, ..., a_d)$ such that for each $1 \le i \le d$,

$$a_i = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{if } i \in \overline{S} \end{cases}$$

then $\mathbf{a}\cdot\mathbf{X}$ and $(\mathbf{1}-\mathbf{a})\cdot\mathbf{X}$ are countermonotonic. This means $\sum\limits_{i\in S} X_i$ and $\sum\limits_{i\in\overline{S}} X_i$ are countermonotonic. So by Definition 4.3.2, $\phi(\sum\limits_{i\in S} X_i, \sum\limits_{i\in\overline{S}} X_i) = -1$, $\varrho(\mathbf{X}) = -1$.

"$\Leftarrow$" Fix some $\mathbf{a} \in \{0,1\}^d$, we look at $\mathbf{a}\cdot\mathbf{X}$ and $(\mathbf{1}-\mathbf{a})\cdot\mathbf{X}$. Denote $\mathbf{a} = (a_1, ..., a_d)$. We define the set S as following: for each $1 \le i \le d$, if $a_i = 1$, put index i into set S. Thus set S contains all the index i such that $a_i = 1$; set $\overline{S}$ contains all the index i such that $a_i = 0$. So $\mathbf{a}\cdot\mathbf{X} = \sum\limits_{i\in S} X_i$ and $(\mathbf{1}-\mathbf{a})\cdot\mathbf{X} = \sum\limits_{i\in\overline{S}} X_i$. Since $\varrho(\mathbf{X}) = -1$ and we use the same notations as in Definition 4.3.1, each $\phi\left(\sum\limits_{i\in S} X_i, \sum\limits_{i\in\overline{S}} X_i\right) = -1$, thus $\sum\limits_{i\in S} X_i$ and $\sum\limits_{i\in\overline{S}} X_i$ are *countermonotonic*.

Since this $\mathbf{a}$ can be any vector in $\{0,1\}^d$, we get $\mathbf{a}\cdot\mathbf{X}$ and $(\mathbf{1}-\mathbf{a})\cdot\mathbf{X}$ are $\Sigma-$*countermonotonic*. $\square$

In Proposition 4.3.7, we prove that $\varrho_d(\mathbf{X}) = 1$ if $\mathbf{X} = (X_1, ..., X_d)$ is a comonotonic random vector. Its reverse direction also holds.

Proposition 4.3.14. *For any random vector* (X, Y)*, if* ϕ *satisfies:* $\phi(X, Y) = 1 \Rightarrow (X, Y)$ *are comonotonic. Then given a d-variate random vector* $\mathbf{X} = (X_1, ..., X_d)$*, if* $\varrho_d(\mathbf{X}) = 1$*, we have* $\mathbf{X}$ *is a comonotonic random vector.*

Proof. When $d = 2$, $\varrho(\mathbf{X}) = \phi(\mathbf{X})$, this automatically holds.
When $d \ge 3$, if $\varrho(\mathbf{X}) = 1$, since for each S, $\phi(\sum_{i\in S} X_i, \sum_{i\in\overline{S}} X_i) \le 1$, then $\phi(\sum_{i\in S} X_i, \sum_{i\in\overline{S}} X_i) = 1$. So $\sum_{i\in S} X_i$ and $\sum_{i\in\overline{S}} X_i$ comonotonic. In particular, there exist increasing functions $g_i, 1 \le i \le d$ such that $g_i(X_i) = \sum_{j\ne i} X_j$. For $k \ne i$,

$$X_k + g_k(X_k) = X_i + g_i(X_i).$$

As g_i, g_k are increasing, when X_k is increasing, $X_k + g_k(X_k)$ is increasing, which is when X_i is increasing. So X_k is an increasing function of X_i, thus X_k and X_i are comonotonic. This holds for all the i and k. By Theorem 3 of Dhaene et al. [2002], the random vector $\mathbf{X}$ is comonotonic when its component is pairwise comonotonic.

□

Definition 4.3.15. A random vector $\mathbf{X} = (X_1, ..., X_d)$ is said to be positively orthant dependent (POD) if both

$$P(X_1 \leq x_1, ..., X_d \leq x_d) \geq \prod_{i=1}^{d} P(X_i \leq x_i), \qquad \text{for all } (x_1, x_2, ..., x_d) \in \mathbb{R}^d$$

and

$$P(X_1 > x_1, ..., X_d > x_d) \geq \prod_{i=1}^{d} P(X_i > x_i), \qquad \text{for all } (x_1, x_2, ..., x_d) \in \mathbb{R}^d$$

hold.

Proposition 4.3.16. *If a random vector* $\mathbf{X}$ *is POD, then* $\varrho(\mathbf{X}) \geq 0$.

Proof. We take a random vector $\mathbf{Y} = (Y_1, ..., Y_d)$, which has the same marginal as $\mathbf{X}$. Assume $\mathbf{Y}$ has independent components, then

$$P(X_1 \leq x_1, ..., X_d \leq x_d) \geq \prod_{i=1}^{d} P(X_i \leq x_i) = \prod_{i=1}^{d} P(Y_i \leq x_i) = P(Y_1 \leq x_1, ..., Y_d \leq x_d)$$

and

$$P(X_1 > x_1, ..., X_d > x_d) \geq \prod_{i=1}^{d} P(X_i > x_i) = \prod_{i=1}^{d} P(Y_i > x_i) = P(Y_1 > x_1, ..., Y_d > x_d)$$

for all $(x_1, x_2, ..., x_d) \in \mathbb{R}^d$.
So $\mathbf{Y} \prec \mathbf{X}$. By Proposition 4.3.7, A1 and A2 properties of $\varrho(Y)$, we get $\varrho(\mathbf{X}) \geq \varrho(\mathbf{Y}) = 0$. □

4.3.2 Estimation

In this section, we introduce some ways to estimate ϱ given a sample of $\mathbf{X}$. One simple way is replacing bivariate measure ϕ by its sample version $\hat{\phi}$ to get the estimate of ϱ.
Another way is described as follows. Given a matrix of data of d-dimensional $\mathbf{X} = (\mathbf{X_1}, ..., \mathbf{X_d})$, each column $\mathbf{X_j} = (X_{1,j}, X_{2,j}, ..., X_{n,j})^T, 1 \leq j \leq d$. Let $S_i = \sum_{j=1}^{d} X_{i,j}$ for each $1 \leq i \leq n$, and $(\tilde{I_{i,j}})_{1\leq i\leq n, 1\leq j\leq d} \sim Bernoulli(0.5)$ independently. We select those $0 < \sum_{j=1}^{d} \tilde{I_{i,j}} < d$, and denote $I_{i,j} = \left(\tilde{I_{i,j}} | 0 < \sum_{j=1}^{d} \tilde{I_{i,j}} < d\right)$. Now we define the estimator $\hat{\varrho}$ of ϱ as

$$\hat{\varrho}(X) = \frac{1}{n} \sum_{i=1}^{n} \phi\left(\sum_{j=1}^{d} I_{i,j} X_{i,j}, S_i - \sum_{j=1}^{d} I_{i,j} X_{i,j}\right). \tag{4.14}$$

Proposition 4.3.17. *Recall Definition 4.3.1 of ϱ, $\hat{\varrho}$ is an unbiased estimator of ϱ. In particular,*

$$E(\hat{\varrho}) = \frac{1}{2^d - 2} \sum_{S \in \mathcal{C}} \phi\left(\sum_{j \in S} X_j, \sum_{j \in \overline{S}} X_j\right)$$

Proof. Fix the ith sample, then we omit i in the subscript now, $(I_j)_{1\leq j\leq d} = (I_{i,j})_{1\leq j\leq d}, (X_j)_{1\leq j\leq d} = (X_{i,j})_{1\leq j\leq d}$.
If $i_1 = i_2 = ... = i_d = 0$ or 1, then

$$P(I_1 = i_1, I_2 = i_2, ..., I_d = i_d) = P(\tilde{I}_1 = i_1, \tilde{I}_2 = i_2, ..., \tilde{I}_d = i_d | 0 < \sum_{j=1}^{d} \tilde{I}_j < d) = 0.$$

If $i_j \in \{0,1\}, j = 1, ..., d$ and $0 < \sum_{j=1}^{d} i_j < d$,

$$
\begin{aligned}
P(I_1 = i_1, I_2 = i_2, ..., I_d = i_d) &= P(\tilde{I}_1 = i_1, \tilde{I}_2 = i_2, ..., \tilde{I}_d = i_d | 0 < \sum_{j=1}^{d} \tilde{I}_j < d) \\
&= \frac{P(\tilde{I}_1 = i_1, \tilde{I}_2 = i_2, ..., \tilde{I}_d = i_d)}{P(0 < \sum_{j=1}^{d} \tilde{I}_j < d)} \\
&= \frac{\prod_{j=1}^{d} P(\tilde{I}_j = i_j)}{1 - P(\sum_{j=1}^{d} \tilde{I}_j = 0) - P(\sum_{j=1}^{d} \tilde{I}_j = d)} \\
&= \frac{\frac{1}{2^d}}{1 - 2 \cdot \frac{1}{2^d}} \\
&= \frac{1}{2^d - 2}.
\end{aligned}
$$

So

$$
\begin{aligned}
&E\left[\phi\left(\sum_{j=1}^{d} I_j X_j, S_i - \sum_{j=1}^{d} I_j X_j\right)\right] \\
=&\phi(X_1, X_2 + ...X_d) \cdot P(I_1 = 1, I_2 = ... = I_d = 0) \\
&+ ... + \phi(X_1 + X_2 + ...X_{d-1}, X_d) \cdot P(I_1 = I_2 = ...I_{d-1} = 1, I_d = 0) \\
=&\frac{1}{2^d - 2} \sum_{S \in \mathcal{C}} \phi\left(\sum_{j \in S} X_j, \sum_{j \in \overline{S}} X_j\right).
\end{aligned}
$$

□

4.4 Conclusion of Chapter 4 and Future Work

In this chapter, we propose a new multivariate dependence measure ϱ focusing on the sum of random variables. We have derived its properties and studied how to estimate it. Future work on ϱ includes deriving more properties, other estimation methods and its applications. The asymptotic normality, robustness and statistical inference associated with the estimator can be studied. See Bernard and Mcleish [2014] for one application to designing algorithms to minimize the variance of the sum.

Chapter 5

Conclusion and Future Work

This thesis studies bounds on variance under partial information available, the compatible matrix problem and gives a new multivariate dependence measure. Conclusion and future directions of each part have been given in Sections 2.6, 3.4 and 4.4. The following topics are beyond the scope of this thesis and thus left to future work:

1. In both Chapters 2 and 3, we study bounds on variance with partial information on dependence: in Chapter 2, we know the copula on some restricted area S. In Chapter 3, we assume that we know the interaction with some background factor Z. It is natural to ask whether there is some relationship between the two problems.

2. The covariance matrix given in Chapter 3 can be studied in terms of other probability distributions. When $n \geq 4$, the matrix is not unique anymore. When $n \to \infty$, do the entries of the matrix converge to any value or follow any distribution? This problem is related to random matrix theory (see Tao [2012]).

3. This thesis deals almost exclusively with the variance but it can be extended to other risk measures, or to study the distribution of the sum.

4. We have only worked on sums of random variables. Our work can be extended to bounds on $f(X_1, ..., X_n)$ for other functions f such as stop-loss premiums, exotic options.

5. This thesis only focuses on dependence of continuous random variables, see Nešlehová [2004] for discussion on non-continuous random variables.

References

K. Aas and G. Puccetti. Bounds on total economic capital: the DNB case study. *Extremes*, pages 1–23, 2014.

H. Albrecher, P. Mayer, and W. Schoutens. General lower bounds for arithmetic Asian option prices. *Applied Mathematical Finance*, 15(2):123–149, 2008.

C. Alexander and J. M. Sarabia. Quantile uncertainty and Value-at-Risk model risk. *Risk Analysis*, 32(8):1293–1308, 2012.

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.

N. Bäuerle and A. Müller. Stochastic orders and risk measures: consistency and bounds. *Insurance: Mathematics and Economics*, 38(1):132–148, 2006.

J. Behboodian, A. Dolati, and M. Úbeda-Flores. A multivariate version of Gini's rank association coefficient. *Statistical Papers*, 48(2):295–304, 2007.

G. Beliakov, T. Calvo, and J. Lazaro. Pointwise construction of Lipschitz aggregation operators with specific properties. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 15(02):193–223, 2007.

C. Bernard and D. Mcleish. Algorithms or finding copulas minimizing the variance of sums. *Working paper*, 2014.

C. Bernard and S. Vanduffel. Optimal investment under probability constraints. In *Proc. 2011 Actuarial and Financial Mathematics Conf*, pages 3–14, 2011.

C. Bernard and S. Vanduffel. A new approach to assessing model risk in high dimensions. *Working paper*, 2014.

C. Bernard, X. Jiang, and S. Vanduffel. Note on “Improved Fréchet bounds and model-free pricing of multi-asset options” by Tankov (2011). *Journal of Applied Probability*, 49(3): 866–875, 2012.

C. Bernard, Y. Liu, N. MacGillivray, and J. Zhang. Bounds on capital requirements for bivariate risk with given marginals and partial information on the dependence. *Dependence Modelling*, 1:37–53, 2013a.

C. Bernard, L. Rüschendorf, and S. Vanduffel. VaR bounds with a variance constraint. *Working Paper*, 2013b.

C. Bernard, P. Boyle, and S. Vanduffel. Explicit representation of cost-efficient strategies. *Finance*, 2014a.

C. Bernard, J. Chen, and S. Vanduffel. Optimal portfolios under worst case scenarios. *Quantitative Finance*, 2014b.

C. Bernard, X. Jiang, and R. Wang. Risk aggregation with dependence uncertainty. *Insurance: Mathematics and Economics*, 54:93–108, Jan. 2014.

C. Bernard, L. Rüschendorf, S. Vanduffel, and R. Wang. Risk bounds for factor models. *Working paper*, 2014.

V. Bignozzi and A. Tsanakas. Model uncertainty in risk capital measurement. *Available at SSRN 2334797*, 2013.

P. Billingsley. *Probability and measure*. John Wiley & Sons, 2008.

J. F. Carriere and L. K. Chan. The bounds of bivariate distributions that limit the value of last-survivor annuities. *Transactions of the Society of Actuaries*, 38(1):51–74, 1986.

N. R. Chaganty and H. Joe. Range of correlation matrices for dependent Bernoulli random variables. *Biometrika*, 93(1):197–206, 2006.

K. C. Cheung. Improved convex upper bound via conditional comonotonicity. *Insurance: Mathematics and Economics*, 42(2):651–655, 2008.

K. C. Cheung and A. Lo. General lower bounds on convex functionals of aggregate sums. *Insurance: Mathematics and Economics*, 53(3):884–896, 2013.

K. C. Cheung and A. Lo. Characterizing mutual exclusivity as the strongest negative multivariate dependence structure. *Insurance: Mathematics and Economics*, 55:180–190, 2014.

K. C. Cheung and S. Vanduffel. Bounds for sums of random variables when the marginal distributions and the variance of the sum are given. *Scandinavian Actuarial Journal*, 2013(2):103–118, 2013.

R. Cont. Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical finance*, 16(3):519–547, 2006.

P. Conti. On some descriptive aspects of measures of monotone dependence. *Metron*, 51 (3-4):43–60, 1993.

M. Curran. Valuing Asian and portfolio options by conditioning on the geometric mean price. *Management science*, 40(12):1705–1711, 1994.

G. Dall'Aglio. Fréchet classes: the beginnings. In *Advances in Probability Distributions with Given Marginals*, pages 1–12. Springer, 1991.

W. F. Darsow, B. Nguyen, E. T. Olsen, et al. Copulas and markov processes. *Illinois Journal of Mathematics*, 36(4):600–642, 1992.

A. d'Aspremont and L. El Ghaoui. Static arbitrage bounds on basket option prices. *Mathematical programming*, 106(3):467–489, 2006.

F. Delbaen. Coherent risk measures on general probability spaces. In *Advances in finance and stochastics*, pages 1–37. Springer, 2002.

M. Denuit, C. Genest, and É. Marceau. Stochastic bounds on sums of dependent risks. *Insurance: Mathematics and Economics*, 25(1):85–104, 1999.

J. Dhaene, D. M, M. J. Goovaerts, K. R, and D. Vyncke. The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics*, pages 1–31, Sept. 2002.

J. Dhaene, S. Vanduffel, M. J. Goovaerts, R. Kaas, and D. Vyncke. Comonotonic approximations for optimal portfolio selection problems. *Journal of Risk and Insurance*, 72(2): 253–300, 2005.

J. Dhaene, D. Linders, W. Schoutens, and D. Vyncke. A multivariate dependence measure for aggregating risks. *Journal of Computational and Applied Mathematics*, 263:78–87, 2014.

F. Durante and J. Fernández-Sánchez. Multivariate shuffles and approximation of copulas. *Statistics & probability letters*, 80(23):1827–1834, 2010.

F. Durante and J. F. Sánchez. On the approximation of copulas via shuffles of min. *Statistics & Probability Letters*, 82(10):1761–1767, 2012.

F. Durante, P. Sarkoci, and C. Sempi. Shuffles of copulas. *Journal of Mathematical Analysis and Applications*, 352(2):914–921, 2009.

P. Embrechts. Copulas: A personal view. *Journal of Risk and Insurance*, 76(3):639–650, 2009.

P. Embrechts and G. Puccetti. Bounds for functions of dependent risks. *Finance and Stochastics*, 10(3):341–352, 2006.

P. Embrechts and G. Puccetti. Bounds for the sum of dependent risks having overlapping marginals. *Journal of Multivariate Analysis*, 101(1):177–190, 2010a.

P. Embrechts and G. Puccetti. Risk aggregation. In *Copula Theory and Its Applications*, pages 111–126. Springer, 2010b.

P. Embrechts, A. McNeil, and D. Straumann. Correlation and dependence in risk management: properties and pitfalls. *Risk management: Value at Risk and beyond*, pages 176–223, 2002.

P. Embrechts, A. Höing, and A. Juri. Using copulae to bound the Value-at-Risk for functions of dependent risks. *Finance and Stochastics*, 7(2):145–167, 2003a.

P. Embrechts, F. Lindskog, and A. McNeil. Modelling dependence with copulas and applications to risk management. *Handbook of heavy tailed distributions in finance*, 8(1): 329–384, 2003b.

P. Embrechts, R. Frey, and A. McNeil. *Quantitative risk management*, volume 10. Princeton Series in Finance, Princeton, 2005.

P. Embrechts, G. Puccetti, and L. Rüschendorf. Model uncertainty and VaR aggregation. *Journal of Banking & Finance*, 37(8):2750–2764, 2013.

P. Embrechts, G. Puccetti, L. Rüschendorf, R. Wang, and A. Beleraj. An academic response to Basel 3.5. *Risks*, 2(1):25–48, 2014a.

P. Embrechts, B. Wang, and R. Wang. Aggregation-robustness and model uncertainty of regulatory risk measures. *ETH Zurich: Zurich, Switzerland*, 2014b.

M. J. Frank, R. B. Nelsen, and B. Schweizer. Best-possible bounds for the distribution of a sum - a problem of Kolmogorov. *Probability Theory and Related Fields*, 74(2):199–211, 1987.

M. Fréchet. Sur les tableaux de corrélation dont les marges sont données. *Comptes rendus herbdomadaires des seances de l cademie des sciences*, 242(20):2426–2428, 1956.

C. Genest, M. Gendron, and M. Bourdeau-Brien. The advent of copulas in finance. *The European Journal of Finance*, 15(7-8):609–618, 2009.

C. Genest, J. Nešlehová, and N. Ben Ghorbal. Spearman's footrule and Gini's gamma: a review with complements. *Journal of Nonparametric Statistics*, 22(8):937–954, 2010.

R. A. Gideon and R. A. Hollister. A rank correlation coefficient resistant to outliers. *Journal of the American Statistical Association*, 82(398):656–666, 1987.

G. H. Golub and C. F. Van Loan. *Matrix computations*, volume 3. JHU Press, 2012.

G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge university press, 1952.

D. Hobson, P. Laurence, and T.-H. Wang. Static-arbitrage upper bounds for the prices of basket options. *Quantitative finance*, 5(4):329–342, 2005.

W. Hoeffding. *Massstabinvariante korrelationstheorie*. Teubner, 1940.

M. Hofer and M. R. Iacò. Optimal bounds for integrals with respect to copulas and applications. *Journal of Optimization Theory and Applications*, 161(3):999–1011, 2014.

R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge university press, 2012.

E. Jakobsons, X. Han, and R. Wang. General convex order on risk aggregation. *Preprint, ETH Zurich*, 2014.

H. Joe. Relative entropy measures of multivariate dependence. *Journal of the American Statistical Association*, 84(405):157–164, 1989.

H. Joe. Multivariate concordance. *Journal of multivariate analysis*, 35(1):12–30, 1990.

H. Joe. *Multivariate models and multivariate dependence concepts*, volume 73. CRC Press, 1997.

P. Jorion. *Value at Risk: the new benchmark for managing financial risk*, volume 2. McGraw-Hill New York, 2007.

R. Kaas, J. Dhaene, and M. J. Goovaerts. Upper and lower bounds for sums of random variables. *Insurance: Mathematics and Economics*, 27(2):151–168, 2000.

R. Kaas, R. J. Laeven, and R. B. Nelsen. Worst VaR scenarios with given marginals and measures of association. *Insurance: Mathematics and Economics*, 44(2):146–158, 2009.

M. Keller-Ressel and C. Griessler. Convex order of discrete, continuous and predictable quadratic variation & applications to options on variance. *arXiv preprint arXiv:1103.2310*, 2011.

J. Kerkhof, B. Melenberg, and H. Schumacher. Model risk and capital reserves. *Journal of Banking & Finance*, 34(1):267–279, 2010.

G. Kimeldori and A. Sampson. Uniform representations of bivariate distributions. *Communications in Statistics–Theory and Methods*, 4(7):617–627, 1975.

S. Kullback. *Information Theory and Statistics*. Courier Dover Publications, 1997.

S. Kullback and R. A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, pages 79–86, 1951.

S. Kusuoka. On law invariant coherent risk measures. In *Advances in mathematical economics*, pages 83–95. Springer, 2001.

W. Lee and J. Y. Ahn. On the multidimensional extension of countermonotonicity and its applications. *Insurance: Mathematics and Economics*, 56:68–79, 2014.

E. Linfoot. An informational measure of correlation. *Information and control*, 1(1):85–89, 1957.

J.-F. Mai and M. Scherer. *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications*, volume 4. World Scientific, 2012.

G. Makarov. Estimates for the distribution function of a sum of two random variables when the marginal distributions are fixed. *Theory of Probability & its Applications*, 26 (4):803–806, 1982.

M. Mesfioui and J.-F. Quessy. Bounds on the Value-at-Risk for the sum of possibly dependent risks. *Insurance: Mathematics and Economics*, 37(1):135–151, 2005.

M. Mesfioui and J.-F. Quessy. Concordance measures for multivariate non-continuous random vectors. *Journal of Multivariate Analysis*, 101(10):2398–2410, 2010.

P. Mikusinski, H. Sherwood, and M. D. Taylor. Shuffles of min. *Stochastica*, 13(1):61–74, 1992.

A. Müller. Stop-loss order for portfolios of dependent risks. *Insurance: Mathematics and Economics*, 21(3):219–223, 1997.

A. Müller and D. Stoyan. *Comparison methods for stochastic models and risks*, volume 389. Wiley, 2002.

R. B. Nelsen. Nonparametric measures of multivariate association. *Lecture Notes-Monograph Series*, pages 223–232, 1996.

R. B. Nelsen. Concordance and copulas: a survey. In *Distributions with given marginals and statistical modelling*, pages 169–177. Springer, 2002.

R. B. Nelsen. *An introduction to copulas.* Springer, 2007.

R. B. Nelsen and M. Ubeda-Flores. A comparison of bounds on sets of joint distribution functions derived from various measures of association. *Communications in Statistics-Theory and Methods*, 33(10):2299–2305, 2005.

R. B. Nelsen and M. Úbeda-Flores. How close are pairwise and mutual independence? *Statistics & Probability Letters*, 82(10):1823–1828, 2012.

R. B. Nelsen, J. J. Quesada-Molina, J. A. Rodriíguez-Lallena, and M. Ubeda-Flores. Bounds on bivariate distribution functions with given margins and measures of association. *Communications in Statistics-Theory and Methods*, 30(6):1055–1062, 2001.

R. B. Nelsen, J. J. Q. Molina, J. A. R. Lallena, and M. Ú. Flores. Best-possible bounds on sets of bivariate distribution functions. *Journal of Multivariate Analysis*, 90(2):348–358, 2004.

J. Nešlehová. Dependence of non-continuous random variables. 2004.

G. Puccetti. Sharp bounds on the expected shortfall for a sum of dependent random variables. *Statistics & Probability Letters*, 83(4):1227–1232, 2013.

G. Puccetti and L. Rüschendorf. Computation of sharp bounds on the distribution of a function of dependent risks. *Journal of Computational and Applied Mathematics*, 236 (7):1833–1840, 2012.

G. Puccetti and R. Wang. General extremal dependence concepts. *Available at SSRN 2436392*, 2014.

G. Puccetti, B. Wang, R. Wang, et al. Advances in complete mixability. *Journal of Applied Probability*, 49(2):430–440, 2012.

G. Puccetti, B. Wang, and R. Wang. Complete mixability and asymptotic equivalence of worst-possible VaR and ES estimates. *Insurance: Mathematics and Economics*, 53(3): 821–828, 2013.

S. Rachev and L. Rüschendorf. Solution of some transportation problems with relaxed or additional constraints. *SIAM Journal on Control and Optimization*, 32(3):673–689, 1994.

S. T. Rachev and L. Rüschendorf. *Mass Transportation Problems: Volume I: Theory*, volume 1. Springer, 1998.

G. Rapuch and T. Roncalli. Some remarks on two-asset options pricing and stochastic dependence of asset prices. *Groupe de Recherche Operationnelle, Credit Lyonnais, France*, 2001.

L. C. G. Rogers and Z. Shi. The value of an Asian option. *Journal of Applied Probability*, pages 1077–1088, 1995.

P. Ruankong, T. Santiwipanont, and S. Sumetkijakan. Shuffles of copulas and a new measure of dependence. *Journal of Mathematical Analysis and Applications*, 398(1): 392–402, 2013.

L. Rüschendorf. Random variables with maximum sums. *Advances in Applied Probability*, pages 623–632, 1982.

L. Rüschendorf. Bounds for distributions with multivariate marginals. *Lecture Notes-Monograph Series*, pages 285–310, 1991.

F. H. Ruymgaart and M. van Zuijlen. Asymptotic normality of multivariate linear rank statistics in the non-iid case. *The Annals of Statistics*, pages 588–602, 1978.

S. Sadooghi-Alvandi, Z. Shishebor, and H. Mardani-Fard. Sharp bounds on a class of copulas with known values at several points. *Communications in Statistics-Theory and Methods*, 42(12):2215–2228, 2013.

F. Schmid and R. Schmidt. Multivariate extensions of Spearman's rho and related statistics. *Statistics & Probability Letters*, 77(4):407–416, 2007.

F. Schmid, R. Schmidt, T. Blumentritt, S. Gaißer, and M. Ruppert. Copula-based measures of multivariate association. In *Copula theory and its applications*, pages 209–236. Springer, 2010.

B. Schweizer. Thirty years of copulas. In *Advances in probability distributions with given marginals*, pages 13–50. Springer, 1991.

B. Schweizer and E. F. Wolff. On nonparametric measures of dependence for random variables. *The Annals of Statistics*, pages 879–885, 1981.

N. Stepanova. Multivariate rank tests for independence and theiry asymptotic efficiency. *Mathematical Methods of Statistics*, 12(2):197–217, 2003.

P. Tankov. Improved Fréchet bounds and model-free pricing of multi-asset options. *Journal of Applied Probability*, 48(2):389–403, 2011.

T. Tao. *Topics in random matrix theory*, volume 132. American Mathematical Soc., 2012.

M. Taylor. Multivariate measures of concordance. *Annals of the Institute of Statistical Mathematics*, 59(4):789–806, 2007.

M. D. Taylor. Multivariate Measures of Concordance for Copulas and their Marginals. *ArXiv e-prints*, Apr. 2010.

A. H. Tchen. Inequalities for distributions with given marginals. *The Annals of Probability*, pages 814–827, 1980.

W. Trutschnig and J. Fernández Sánchez. Some results on shuffles of two-dimensional copulas. *Journal of Statistical Planning and Inference*, 143(2):251–260, 2013.

E. A. Valdez, J. Dhaene, M. Maj, and S. Vanduffel. Bounds and approximations for sums of dependent log-elliptical random variables. *Insurance: Mathematics and Economics*, 44(3):385–397, 2009.

A. W. Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

S. Vanduffel, Z. Shang, L. Henrard, J. Dhaene, and E. A. Valdez. Analytic bounds and approximations for annuities and Asian options. *Insurance: Mathematics and Economics*, 42(3):1109–1117, 2008.

M. Vanmaele, G. Deelstra, J. Liinev, J. Dhaene, and M. J. Goovaerts. Bounds for the price of discrete arithmetic Asian options. *Journal of Computational and Applied Mathematics*, 185(1):51–90, 2006.

B. Wang and R. Wang. The complete mixability and convex minimization problems with monotone marginal densities. *Journal of Multivariate Analysis*, 102(10):1344–1360, 2011.

B. Wang and R. Wang. Extreme negative dependence and risk aggregation. *Preprint, University of Waterloo*, 2013.

B. Wang and R. Wang. Joint mixability. *Preprint, University of Waterloo*, 2014.

R. Wang. Asymptotic bounds for the distribution of the sum of dependent random variables. *Journal of Applied Probability, to appear*, 2014.

R. Wang, L. Peng, and J. Yang. Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. *Finance and Stochastics*, pages 1–23, 2013.

H. Werner. Analytical bounds for two Value-at-Risk functionals. *Astin bulletin*, 32(2): 235–265, 2002.

E. F. Wolff. N-dimensional measures of dependence. *Representations*, 1:1, 1980.

Y. Zhang and C. Yin. A new multivariate dependence measure based on comonotonicity. *arXiv preprint arXiv:1410.7845*, 2014.