Current open questions in complete mixability

Ruodu Wang[*]

Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada. e-mail: wang@uwaterloo.ca

Abstract: Complete and joint mixability has raised considerable interest in recent few years, in both the theory of distributions with given margins, and applications in discrete optimization and quantitative risk management. We list various open questions in the theory of complete and joint mixability, which are mathematically concrete, and yet accessible to a broad range of researchers without specific background knowledge. In addition to the discussions on open questions, some results contained in this paper are new.

MSC 2010 subject classifications: Primary 60C05, 60E05; secondary 60E15.
Keywords and phrases: Complete mixability, joint mixability, dependence, optimization, Fréchet problems.

Contents

1. Complete and joint mixability

1.1. Motivation of the paper

In this paper we summarize the current state-of-the-art of research in complete and joint mixability, and list about a dozen of open questions which the author considers challenging and with potential impact. Many of the open questions emerged via private communications with other researchers[1]. Most of the questions are formulated in terms of complete mixability due to its nicer mathematical properties, and this is already reflected in the title of this paper. The paper carries dual purposes: to stimulate research activities leading to developments in the listed challenges, and to introduce the topic of complete and joint mixability to a broader range of scholars, especially from different fields of study other than applied probability. Both purposes serve to advance the study of this topic as well as of the expanding field of dependence modeling.

One nice feature of complete and joint mixability is that the topic requires very little background knowledge, and the mathematical questions are still concrete and challenging. The definitions, existing results and open questions can be easily understood by any graduate students in analysis, combinatorics, probability or statistics. This paper hopefully enhances the accessibility of the subject to a broad range of researchers at all levels and from all fields of mathematics and statistics.

Although the main purpose of this paper is to discuss open questions, some useful results listed in Section 2 (in particular, Theorems 1–5) are original in this paper.

It is the author's sincere hope that, after one or a few decades, most of the questions listed in this paper would have been answered and connected to other fields of mathematics and its applications, inspiring new research directions that are unseen from today. Some of the questions listed here may be easily answered by some experts, especially those from other fields, due to the obvious limitation of the author's knowledge. The opinions expressed in this paper as well as any errors are solely the responsibility of the author.

1.2. An optimization problem

We start the story of mixability by a simple optimization problem. Suppose that there are n steps in the production of an equipment of a certain type. A company employs m workers specialized in each of its n production steps, that is, mn workers in total; as such the company is able to produce m equipments simultaneously. We use (i, j) for the i-th worker employed in the j-th production step. Suppose that the time for worker (i, j) to finish her job is a positive number $a_{i,j}$. Each time, the company produces m equipments and then send them out to a buyer. Naturally, the company is interested in minimizing the time T of production of the m equipments, namely, $T = \max\{t_1, \ldots, t_m\}$ where t_i stands

---

[1]Some names are listed in the Acknowledgement.

for the time taken in the production of the i-th equipment, $i = 1, \ldots, m$. What is the optimal arrangement of workers for each equipment?

Each equipment is assigned with n workers with one from each step. That is, for some m-permutations $\sigma_1, \ldots, \sigma_n$, workers $(\sigma_1(i), 1), \ldots, (\sigma_n(i), n)$ are assigned to the i-th equipment, and hence $t_i = \sum_{j=1}^n a_{\sigma_j(i),j}$. The problem is

$$\text{to minimize} \quad T = \max \left\{ \sum_{j=1}^n a_{\sigma_j(i),j} : i = 1, \ldots, m \right\} \quad \text{over all } \sigma_1, \ldots, \sigma_n \in S_m,$$

(1)

where S_m is the set of m-permutations. This optimization target is a minimax, and is very often consistent with variance reduction problems: to minimize quantities such as the sample variance of $t_1, \ldots, t_n$.

Intuitively, the optimal arrangement should be such that $t_1, \ldots, t_m$ are close enough, and ideally equal. Since

$$t_1 + \cdots + t_m = \sum_{i=1}^m \sum_{j=1}^n a_{i,j},$$

we have that

$$T^* = \min_{\sigma_1, \ldots, \sigma_n \in S_m} T \geqslant \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n a_{i,j},$$

and ideally, T^* would be almost equal to $\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n a_{i,j}$. Here we have two general questions:

(i) What is an optimal arrangement $(\sigma_1, \ldots, \sigma_n)$ for (1)? How could we calculate an optimal arrangement $(\sigma_1, \ldots, \sigma_n)$?
(ii) Under what conditions, $T^* = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n a_{i,j}$? How could we calculate T^*?

Both questions are related to the concept of joint mixability, the main focus of this paper.

1.3. Definitions and terminologies

Throughout, n and d are positive integers, and we assume a atomless space $(\Omega, \mathcal{A}, \mathbb{P})$ of random variables taking values in a semigroup $\mathbb{G}$ which can be chosen as $\mathbb{R}^d$ in most cases. In the literature, complete mixability and joint mixability are defined for distributions on $\mathbb{R}$. Theoretically, the concepts of mixability do not require any extra mathematical (topological, algebraic, analytical) structure on the underlying set $\mathbb{G}$ of study, other than an addition ($+$); in view of applications, only the case $\mathbb{G} = \mathbb{R}^d$ is particularly relevant. In the following we use the term "distributions" for probability measures.

Definition 1 (Joint mixability). An n-tuple $(F_1, \ldots, F_n)$ of distributions on $\mathbb{G}$ is *jointly mixable* (JM) if there exists a distribution H on $\mathbb{G}^n$ with margins $F_1, \ldots, F_n$ such that H is supported in $\{(x_1, \ldots, x_n) \in \mathbb{G}^n : x_1 + \cdots + x_n = K\}$ for some $K \in \mathbb{G}$.

In the literature, mixability is often defined using the language of random variables. The two definitions are equivalent.

Definition 2 (Definition given in [46]). A random vector $(X_1, \ldots, X_n)$ satisfying

$$X_1 + \cdots + X_n = K \quad \text{for some} \quad K \in \mathbb{G}, \tag{2}$$

is called a *joint mix*. An n-tuple $(F_1, \ldots, F_n)$ of distributions on $\mathbb{G}$ is *jointly mixable* (JM) if there exists a joint mix with marginal distributions $F_1, \ldots, F_n$. When $\mathbb{G} = \mathbb{R}^d$, K in (2) is called a *joint center* of $(F_1, \ldots, F_n)$.

Joint mixability is supported by many applications, including the optimization problem in Section 1.2; see also Section 1.5 below. Suppose that for $j = 1, \ldots, n$, F_j is a discrete distribution on $\mathbb{R}$ supported on distinct points $a_{1,j}, \ldots, a_{m,j}$ with point-mass $1/m$ each. Let us recall question (ii) in Section 1.2. If $T^* = \frac{1}{m} \sum_{i=1}^{m} \sum_{j=1}^{n} a_{i,j}$, then for some $\sigma_1^*, \ldots, \sigma_n^* \in S_m$,

$$T^* = \sum_{j=1}^{n} a_{\sigma_j^*(i),j}, \quad i = 1, \ldots, m.$$

Now let U be a discrete uniform random variable in $\{1, \ldots, m\}$, and define $X_j = a_{\sigma_j^*(U),j}, j = 1, \ldots, n$. It follows that $X_j \sim F_j, j = 1, \ldots, n$, and $X_1 + \cdots + X_n = T^*$. That is, $(F_1, \ldots, F_n)$ is jointly mixable, and the optimal arrangement $(\sigma_1^*, \ldots, \sigma_n^*)$ in question (i) corresponds to a joint mix $(X_1, \ldots, X_n)$. Question (ii) in Section 1.2 is a special question of joint mixability.

Below we give the definition of complete mixability, which is the homogeneous case of joint mixability when all marginal distributions are identical.

Definition 3 (Complete mixability). A distribution F on $\mathbb{G}$ is called *n-completely mixable* (*n*-CM) if the n-tuple $(F, \ldots, F)$ is jointly mixable. When $\mathbb{G} = \mathbb{R}^d$, $\mu = K/n$ is called a *center* of F, where K is the joint center of the n-tuple $(F, \ldots, F)$. A joint mix with identical margins F is called a *complete mix*.

The reason why distributions in Definitions 1–3 are called *mixable* is that we are curious about whether one is *able* to find a joint *mix* with the given constraints on margins.

Although complete mixability is a special case of joint mixability, the two concepts are studied separately in the literature as they require mathematical techniques at significantly different levels; see for example the results on monotone densities in [43] and [44]. In addition, n-complete mixability is a property of a single distribution, allowing us to study the property by letting n vary.

We denote by $\mathcal{M}_n(\mu)$ the set of all n-CM distributions on $\mathbb{R}$ with center μ, and by $\mathcal{J}_n(K)$ the set of all n-tuples of JM distributions with joint center K, that is,

$$\mathcal{J}_n(K) = \{(F_1, \ldots, F_n) : (F_1, \ldots, F_n) \text{ is JM with joint center } K\}.$$

Apparently, $F \in \mathcal{M}_n(\mu)$ if and only if $(F, \ldots, F) \in \mathcal{J}_n(n\mu)$. For $n = 1$ or $n = 2$, the sets $\mathcal{J}_n(K)$ and $\mathcal{M}_n(\mu)$ are fully characterized. For $n \geqslant 3$, a full characterization of either set is still an open question.

Remark 1. In Definitions 1–3, both the summation $x_1 + \cdots + x_n$ and the constant constraint K in the support of H are for mathematical tractability and practical relevance. Other constraints may be chosen for different purposes of applications or theoretical studies.

1.4. Related literature

In this section we provide a non-exhaustive brief list of related literature on complete and joint mixability, especially for the reader who is new to this topic. Except for a few early milestone studies, most papers listed here are within the recent few years.

Probability measures with given margins have been studied since the early work by Fréchet [17] and Hoeffding [20]; see also the milestone papers [41, 42]. The first study of questions specifically related to complete mixability was given in [18] where uniform distributions were shown to be n-CM for $n \geqslant 2$. Relevant contributions from the perspectives of mass transportation, variance reduction and stochastic orders can be found in [37, 40, 27, 23]. The terms *complete mixability* and *joint mixability* were introduced in [43] and [46] respectively, along with properties and results on the complete mixability of monotone densities. Recent advances on complete and joint mixability can be found in [32, 33, 44].

As opposed to the strongest positive dependence (see for instance *comonotonicity* in [9]), a universal notion of the strongest negative dependence does not exist for a collection of more than two random variables, and the corresponding optimization problems are generally much more complicated than those involving the strongest positive dependence. Complete and joint mixes are sometimes argued to have the strongest negative dependence structure as they naturally solve a large class of optimization problems. Recent studies searching for a notion of extremal negative dependence can be found in [8, 43, 5, 6, 24]. A recent review on extremal dependence concepts is given in [35].

Algorithms related to mixability have been designed for questions (i) and (ii) in Section 1.2. An early study on *rearrangement methods* is found in [38]; some recent research includes [28, 12, 34, 19]. In particular, [19] showed that question (i) in Section 1.2 is NP-complete even in the case when all $a_{i,j} \in \mathbb{Z}$. As such, an analytical characterization of joint mixability is of considerable importance.

1.5. Applications

The concepts of complete and joint mixability are closely related to many optimization problems with marginal constraints. We discuss a few of them in this section. For the reader who is only interested in mathematical challenges, this section may be skipped.

Let $\mathcal{X}$ be a convex cone of random variables (taking values in $\mathbb{R}$) of interest; $\mathcal{X}$ can be chosen as L^1 (the set of integrable random variables) or L^∞ (the set of bounded random variables) in most applications. For some univariate distributions $F_1, \ldots, F_n$, define the *aggregation set*

$$\mathcal{D}_n = \{X_1 + \cdots + X_n : X_i \in \mathcal{X}, \ X_i \sim F_i, \ i = 1, \ldots, n\} \subset \mathcal{X}.$$

Many optimization problems, including variance reduction problems, convex functionals minimization, and maximin and minimax problems such as (1) described in Section 1.2, boil down to the search for the smallest element in $\mathcal{D}_n$ with respect to convex order.

Definition 4 (Convex order). Let X and Y two random variables with finite mean. X is smaller than Y *in convex order*, denoted by $X \prec_{\mathrm{cx}} Y$, if for all convex functions f,

$$\mathbb{E}[f(X)] \leqslant \mathbb{E}[f(Y)], \tag{3}$$

whenever both sides of (3) are well-defined.

It is well-known that the convex ordering largest element in $\mathcal{D}_n$ is always obtained by $F_1^{-1}(U) + \cdots + F_n^{-1}(U)$ for a random variable $U \sim \mathrm{U}[0,1]$. However, it remains open in general to find the smallest element in $\mathcal{D}_n$ with respect to convex order for $n \geqslant 3$. [2] gave an example where $\mathcal{D}_n$ does not contain a smallest element in this sense.

When $(F_1, \ldots, F_n)$ is JM with joint center K, it is easy to see that $K \in \mathcal{D}_n$ and $K \prec_{\mathrm{cx}} S$ for all $S \in \mathcal{D}_n$. In [43] and [21], the smallest element with respect to convex order in $\mathcal{D}_n$ is characterized based on joint mixability when $F_1, \ldots, F_n$ have monotone densities even if $(F_1, \ldots, F_n)$ is not jointly mixable.

The concepts of complete and joint mixability have also raised a considerable interest in quantitative risk management, as it plays an important role in the context of *risk aggregation with dependence uncertainty*. A typical question in this field concerns the calculation of

$$\sup\{\rho(S) : S \in \mathcal{D}_n\} \quad \text{and} \quad \inf\{\rho(S) : S \in \mathcal{D}_n\} \tag{4}$$

for a law-determined risk measure $\rho : \mathcal{X} \to \mathbb{R}$; see the book [26, Section 6.2] and the early work on the risk measure Value-at-Risk (VaR) in [10]. Here the set $\mathcal{D}_n$ represents the set of possible aggregate risks under model uncertainty at the level of dependence, a common setup in risk management practice.

In the case when ρ is a convex risk measure, ρ typically respects convex order; see for instance Föllmer and Schied [16, Section 4]. Thus (4) boils down to questions of convex order in $\mathcal{D}_n$ as discussed above. Convex risk measures include the Expected Shortfall, a popular risk measure used in banking regulation; see [13].

Among non-convex risk measures, the Value-at-Risk (VaR) is of particular interest in portfolio management. The Value-at-Risk of a random variable X at level $p \in (0,1)$ is defined as the (left-continuous) inverse distribution function

$$\mathrm{VaR}_p(X) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \leqslant x) \geqslant p\}.$$

Quantities of interest are

$$\begin{aligned}
\overline{\mathrm{VaR}}_p &= \sup\{\mathrm{VaR}_p(S) : S \in \mathcal{D}_n\}, \quad \text{and} \\
\underline{\mathrm{VaR}}_p &= \inf\{\mathrm{VaR}_p(S) : S \in \mathcal{D}_n\}, \quad p \in (0,1).
\end{aligned} \tag{5}$$

For more discussions and applications of this topic, see [12, 13]. The following result on $\overline{\mathrm{VaR}}_p$ is given in [46]; the case of $\underline{\mathrm{VaR}}_p$ is symmetric.

(1) For each $p \in (0,1)$, let $\Phi(p) = \frac{1}{1-p} \sum_{i=1}^{n} \int_{p}^{1} \mathrm{VaR}_q(X) \mathrm{d}q$. It holds that $\overline{\mathrm{VaR}}_p \leqslant \Phi(p)$.

(2) For each $p \in (0,1)$, the equality $\overline{\mathrm{VaR}}_p = \Phi(p)$ holds if and only if the n-tuple of the distributions of $F_1^{-1}(W), \ldots, F_n^{-1}(W)$ is jointly mixable, where $W \sim \mathrm{U}[p,1]$.

The above result can also be applied to find minimal or maximal probability function of random variables in $\mathcal{D}_n$.

Although in many cases $(F_1, \ldots, F_n)$ is not jointly mixable, solutions of (4) and (5) can still be obtained based on conditional complete or joint mixability in many cases; see [46, 30, 12, 2, 21] for work in this direction. Some other recent research on (4) and (5) involving mixability can be found in [29, 7, 31, 33, 3, 1, 14]. We refer to [13] for a recent review of this subject in the context of banking regulation, and the book [39] contains a comprehensive treatment of many related problems.

2. Current open questions

In this section, we discuss some open questions in complete and joint mixability. Unless otherwise specified, we consider $\mathbb{G} = \mathbb{R}$, and F is a distribution on $\mathbb{R}$. We use L^0 for the set of all random variables in $(\Omega, \mathcal{A}, \mathbb{P})$ taking values in $\mathbb{R}$, and we use I_A to denote the indicator function of a set A. It is not necessary to read the following questions in a particular order.

2.1. Uniqueness of the center

Suppose that F is n-CM. It is obvious that if F has finite mean μ, then its center is unique and equal to μ. It is shown that if $x\mathbb{P}(|X| \geqslant x) \to 0$ as $x \to \infty$ for $X \sim F$, then the center of F is also unique; see [43, Proposition 2.1]. This uniqueness can be easily extended to the case of $\mathbb{G} = \mathbb{R}^d$. For a generic Abelian group $\mathbb{G}$, the uniqueness is not guaranteed; an example can be easily built for finite cyclic groups. For instance, consider a Bernoulli distribution $\mathrm{Bern}(1/2)$ on $\mathbb{Z}_2$ with $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = 1/2$ for $X \sim \mathrm{Bern}(1/2)$. It is obvious that $X + X = 0$ and $X + (1 - X) = 1$ on $\mathbb{Z}_2$, hence the center is not unique in this setting.

We are interested in whether the center μ is always unique for the case $\mathbb{G} = \mathbb{R}$ or $\mathbb{R}^d$. Non-uniqueness may only happen in the case that the support of F is unbounded from both sides, and F does not have finite mean. Note that the index n in complete mixability is irrelevant; indeed if a distribution F is n-CM with center μ_1 and k-CM with center μ_2, $\mu_1 \neq \mu_2$, then F is also nk-CM with centers μ_1 and μ_2. Therefore, it suffices to determine whether a distribution can be n-CM with different centers for any $n \in \mathbb{N}$.

Open Problem 1. Is the center of mixability always unique for a distribution on $\mathbb{R}$ (or $\mathbb{R}^d$)? In other words, for $\mu, \nu \in \mathbb{R}$, $\mu \neq \nu$, is it true that $\mathcal{M}_n(\mu) \cap \mathcal{M}_n(\nu) = \varnothing$?

2.2. Generic proofs of some theorems

Below we list some main results on complete and joint mixability in the recent literature.

(a) [40] Any continuous distribution function F having a symmetric and uni-modal density is n-CM for $n \geqslant 2$.

(b) [43] Suppose that F admits a monotone density on its essential support $[a, b]$ with mean μ. Then F is n-CM if and only if

$$a + \frac{b-a}{n} \leqslant \mu \leqslant b - \frac{b-a}{n}. \tag{6}$$

(c) [32] Suppose that F admits a concave density on its essential support. Then F is n-CM for $n \geqslant 3$.

(d) [33] Suppose that F admits a density f on a finite interval $[a, b]$, and $f(x) \geqslant \frac{3}{n(b-a)}$ on $[a, b]$. Then F is n-CM.

(e) [44] Suppose that $F_1, \ldots, F_n$ all admit increasing (or decreasing) densities on their essential supports $[a_i, b_i]$ and have mean μ_i, $i = 1, \ldots, n$, respectively. Then $(F_1, \ldots, F_n)$ is JM if and only if

$$\sum_{i=1}^{n} a_i + \max_{i=1,\ldots,n} (b_i - a_i) \leqslant \sum_{i=1}^{n} \mu_i \leqslant \sum_{i=1}^{n} b_i - \max_{i=1,\ldots,n} (b_i - a_i). \tag{7}$$

(f) [44] Suppose that $F_i \sim E_1(\mu_i, \sigma_i^2, \phi)$, where E_1 is the 1-elliptical distribution (for definition, see [15]) with parameters $\mu_i \in \mathbb{R}$, $\sigma_i \geqslant 0$, $i = 1, \ldots, n$, and ϕ is a characteristic generator for an n-elliptical distribution. Then $(F_1, \ldots, F_n)$ is JM if and only if

$$\sum_{i=1}^{n} \sigma_i \geqslant 2 \max_{i=1,\ldots,n} \sigma_i. \tag{8}$$

Note that all results (a)–(e) include uniform distributions as a special case. The proofs of (a) and (f) are analytical and reasonably straightforward due to the symmetric nature of the underlying distributions. The dependence structure of a corresponding joint mix in (a) and (f) is clear.

However, the proofs of the recent results on complete mixability, namely (b)–(d), are all based on combinatorics and discretization of distributions. We outline the common logic of the proofs as follows. To show that a distribution F is n-CM, first we find a sequence of discretizations of this distribution, say F_N, with $F_N \to F$ (sufficiently in the weak sense) as $N \to \infty$. Then, for a fixed N, we try to show that F_N can be decomposed to a convex combination of n-discrete uniform distributions with the same mean, or a convex combination of known-to-be-n-CM discrete distributions with the same mean. This often involves mathematical induction on the number of points in the support of F_N. The proof of result (e) in [44] is even more complicated; it involves decomposition of $F_1, \ldots, F_n$ into combination of distributions with step density functions

(which are not jointly mixable, but in some sense close to being jointly mixable), and a mathematical induction on the number of effective steps is used. The proofs for the above-mentioned results are typically very long and technical, and more importantly the details of the dependence structure for a joint mix are always unclear. These rather unfortunate features significantly reduce the accessibility of the theory of mixability for the general reader.

Through private communications with many scholars interested in this topic, the author believes that generic (probabilistic, analytic) proofs without involving combinatorics or mathematical induction is in demand for the future development of the theory.

Open Problem 2. Is there a generic (probabilistic, analytic) proof of the main results in complete and joint mixability?

Some duality theorems on probability measures with given margins in the literature can be applied to complete and joint mixability. Recent studies on complete mixability using duality methods are found in [29, 30, 45]. The following theorem was essentially established in [41, 37]. How they could be used to generate new results on mixability is still unclear.

Theorem 1 ([41, 37]). *For distributions $F_1, \ldots, F_n$ on $\mathbb{R}$, the following statements are equivalent:*

(i) *$(F_1, \ldots, F_n)$ is jointly mixable with joint center K.*
(ii) *For all measurable functions $f_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, n$,*

$$\sum_{i=1}^{n} \int f_i \mathrm{d}F_i \geqslant \inf \left\{ \sum_{i=1}^{n} \mathbb{E}[f_i(Y_i)] : Y_1, \ldots, Y_n \in L^0, \ \sum_{i=1}^{n} Y_i = K \right\},$$

whenever both sides of the above equation are finite.
(iii) *For all measurable functions $f_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, n$ such that $\sum_{i=1}^{n} f_i(x_i) \geqslant \mathrm{I}_{\{x_1 + \cdots + x_n = K\}}$ for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$,*

$$\sum_{i=1}^{n} \int f_i \mathrm{d}F_i \geqslant 1,$$

whenever the left-hand side of the above equation is finite.

Proof.

(a) (i)$\Rightarrow$(iii): Let $(X_1, \ldots, X_n)$ be a joint mix with joint center K, and $X_i \sim F_i$, $i = 1, \ldots, n$. Then for measurable functions $f_1, \ldots, f_n$ in (iii),

$$\sum_{i=1}^{n} \int f_i \mathrm{d}F_i = \sum_{i=1}^{n} \mathbb{E}[f_i(X_i)] \geqslant \mathbb{E}[\mathrm{I}_{\{X_1 + \cdots + X_n = K\}}] = 1.$$

(b) (iii)$\Rightarrow$(ii): For measurable functions $f_i : \mathbb{R} \to \mathbb{R}_+$, $i = 1, \ldots, n$, let

$$\xi = \inf \left\{ \sum_{i=1}^{n} \mathbb{E}[f_i(Y_i)] : Y_1, \ldots, Y_n \in L^0, \ \sum_{i=1}^{n} Y_i = K \right\}.$$

It follows that

$$\xi \leqslant \inf \left\{ \sum_{i=1}^{n} f_i(y_i) : y_1, \ldots, y_n \in \mathbb{R}, \ \sum_{i=1}^{n} y_i = K \right\}.$$

That is, $\sum_{i=1}^{n} f_i(x_i)/\xi \geqslant I_{\{x_1 + \cdots + x_n = K\}}$ for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$. It follows from (iii) that

$$\sum_{i=1}^{n} \int f_i \mathrm{d} F_i \geqslant \xi = \inf \left\{ \sum_{i=1}^{n} \mathbb{E}[f_i(Y_i)] : Y_i \in L^0, \ i = 1, \ldots, n, \ \sum_{i=1}^{n} Y_i = K \right\}. \tag{9}$$

Now we have shown that (9) holds for non-negative functions $f_1, \ldots, f_n$. Note that (9) is invariant under a shift in any of $f_1, \ldots, f_n$, and hence it holds also for all functions $f_1, \ldots, f_n$ bounded from below. For functions that are unbounded from below, a standard approximation argument using monotone convergence theorem would show that (9) still holds.

(c) (ii)$\Rightarrow$(i): this directly follows from Theorem 7 of [41]. $\qquad\square$

Remark 2. Indeed, in Theorem 1, (i)$\Leftrightarrow$(ii) can be obtained from a particular case of [41, Theorem 7], and (i)$\Leftrightarrow$(iii) can be obtained from a particular case of [37, Equation (4)]; see also [36, Theorem 1]. None of the results in [41] and [37] are stated specifically for the case of mixability.

2.3. Representation and decomposition

There are two types of decomposition of complete and joint mixability into simple objects, as shown in Theorems 2 and 3 below. Although similar ideas may be found in the literature, the theorems themselves are new in this paper.

In the following, we say a distribution F is an *n-discrete uniform distribution* on $(a_1, \ldots, a_n) \in \mathbb{R}^n$ if $\mathbb{P}(X = x) = \#\{i = 1, \ldots, n : a_i = x\}/n$ for $X \sim F$. Theorem 3.2 of [32] says that a discrete distribution F is n-CM with center μ if and only if it has a decomposition:

$$F = \sum_{i=1}^{\infty} b_i F_i,$$

where $\sum_{i=1}^{\infty} b_i = 1$, $b_i \geqslant 0$, $i \in \mathbb{N}$ and F_i, $i \in \mathbb{N}$ are n-discrete uniform distributions with mean μ. A stronger result can be obtained for any CM distributions.

Theorem 2. *A distribution F on $\mathbb{R}$ is n-CM with center μ if and only if it has the following representation*

$$F = \int_{\mathbb{R}^n} F_{\mathbf{a}} \mathrm{d} h(\mathbf{a}), \tag{10}$$

where $F_{\mathbf{a}}$, $\mathbf{a} \in \mathbb{R}^n$ are n-discrete uniform distributions with mean μ, h is a probability measure on $\mathbb{R}^n$, and for a fixed $x \in \mathbb{R}$, $F_{\mathbf{a}}(x)$ is measurable in $\mathbf{a} \in \mathbb{R}^n$.

Proof. Let $\mathbf{X_a}$ be a μ-centered complete mix with identical marginal distributions $F_{\mathbf{a}}$, $\mathbf{a} \in \mathbb{R}^n$. Take a random vector $\mathbf{A} \sim h$ be independent of $\mathbf{X_a}$, $\mathbf{a} \in \mathbb{R}^n$ and define

$$\mathbf{X_A}(\omega) = \mathbf{X}_{\mathbf{A}(\omega)}(\omega), \ \omega \in \Omega.$$

It is easy to see that $\mathbf{X_A}$ is also a μ-centered complete mix. The marginal distribution of $\mathbf{X_A}$ can be easily calculated as

$$\mathbb{P}\left(\mathbf{X_A} \leqslant (x, \infty, \ldots, \infty)\right) = \int_{\mathbb{R}^n} \mathbb{P}\left(\mathbf{X_a} \leqslant (x, \infty, \ldots, \infty)\right) \mathrm{d}h(\mathbf{a})$$
$$= \int_{\mathbb{R}^n} F_{\mathbf{a}} \mathrm{d}h(\mathbf{a}) = F(x), \quad x \in \mathbb{R}.$$

Hence, F is n-CM with center μ.

Now suppose that F is n-CM with center μ. Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a μ-centered complete mix with identical marginal distributions F. Let $F_{\mathbf{a}}$ for $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{a} \cdot \mathbf{1}_n = n\mu$ be the n-discrete uniform distribution on $\mathbf{a}$ with mean μ. It is obvious that for fixed $x \in \mathbb{R}$, $F_{\mathbf{a}}(x)$ is measurable in $\mathbf{a} \in \mathbb{R}^n$. Let U be a discrete uniform distribution on $(1, \ldots, n)$, independent of $\mathbf{X}$, and $Z = \sum_{i=1}^n X_i \mathrm{I}_{\{U=i\}}$. It is straightforward to verify that that $Z \sim F$, and for $x \in \mathbb{R}$,

$$\mathbb{P}(Z \leqslant x|\mathbf{X}) = \frac{1}{n} \sum_{i=1}^n \mathbb{P}(X_i \leqslant x|\mathbf{X}) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\mathrm{I}_{\{X_i \leqslant x\}}|\mathbf{X}] = \frac{1}{n} \sum_{i=1}^n \mathrm{I}_{\{X_i \leqslant x\}}$$
$$= F_{\mathbf{a}}(x)\big|_{\mathbf{a}=\mathbf{X}}.$$

It follows that

$$F(x) = \mathbb{E}[\mathbb{P}(Z \leqslant x|X)] = \int_{\mathbb{R}^n} F_{\mathbf{a}}(x) \mathrm{d}\mathbb{P}(\mathbf{X} \leqslant \mathbf{a}), \quad x \in \mathbb{R},$$

and $h(\mathbf{a})$ in (10) can be chosen as $\mathbb{P}(\mathbf{X} \leqslant \mathbf{a})$. $\qquad\square$

Since complete (and joint) mixability is preserved by taking weak limit (see [43]), it is often sufficient to investigate complete mixability for bounded discrete distributions on $\mathbb{Z}$ and then take a limit for general distributions; this technique was used repeatedly in [43, 32, 33]. We say a vector $\mathbf{X}$ is a *binary multinomial random vector* if $\mathbf{X}$ has a multinomial distribution with the "number of trials" parameter $n = 1$, that is, $\mathbf{X}$ takes values in $\{0, 1\}^n$ and exactly one of the components of $\mathbf{X}$ is 1. The following decomposition, which could be seen as "perpendicular" to Theorem 2, may be of help to characterize complete and joint mixability.

Theorem 3. *Suppose that $\mathbf{X}$ takes values in $\mathbb{Z}_+^n$. $\mathbf{X}$ is a joint mix with joint center $N \in \mathbb{Z}_+$ if and only if it has the following representation*

$$\mathbf{X} = \sum_{k=1}^N \mathbf{X}_k, \tag{11}$$

where $\mathbf{X}_k$, $k = 1, \ldots, N$ are binary multinomial random vectors.

Proof. Suppose that (11) holds. Since $\mathbf{X}_k \cdot \mathbf{1}_n = 1$, it is easy to see that $\mathbf{X}$ is a joint mix with center N. Now suppose that $\mathbf{X} = (X_1, \ldots, X_n)$ is a joint mix. For $k = 1, \ldots, N$ and $i = 1, \ldots, n$ let

$$Y_{k,i} = I_{\{\sum_{j=1}^{i} X_j \geqslant k\}} - I_{\{\sum_{j=1}^{i-1} X_j \geqslant k\}}$$

with the convention that $\sum_{j=1}^{0} X_j = 0$, and let $\mathbf{X}_k = (Y_{k,1}, \ldots, Y_{k,n})$. Since $X_1 + \cdots + X_n = N$, each $\mathbf{X}_k$ is binary multinomial. Then

$$\sum_{k=1}^{N} \mathbf{X}_k = \left(\sum_{k=1}^{N} Y_{k,1}, \ldots, \sum_{k=1}^{N} Y_{k,n} \right)$$

$$= \left(\sum_{k=1}^{N} I_{\{X_1 \geqslant k\}}, \sum_{k=1}^{N} I_{\{X_1 + X_2 \geqslant k\}} - \sum_{k=1}^{N} I_{\{X_1 \geqslant k\}}, \ldots, \sum_{k=1}^{N} I_{\{\sum_{j=1}^{n} X_j \geqslant k\}} \right.$$

$$\left. - \sum_{k=1}^{N} I_{\{\sum_{j=1}^{n-1} X_j \geqslant k\}} \right)$$

$$= \left(X_1, X_1 + X_2 - X_1, \ldots, \sum_{j=1}^{n} X_j - \sum_{j=1}^{n-1} X_j \right) = (X_1, X_2, \ldots, X_n).$$

Thus, $\mathbf{X}$ admits a decomposition of (11). □

As a trivial consequence of Theorem 3, any binomial distribution with parameters $(n, 1/n)$ for $n \in \mathbb{N}$ is n-CM, since it is the marginal distribution of multinomial distribution with parameters $(n; 1/n, \ldots, 1/n)$, and any multinomial random vector has a natural representation (11).

Arguments of the type of Theorem 2 has been applied extensively in the recent literature to show the complete/joint mixability of some classes of distributions. It remains a question whether Theorem 3 can be useful in a non-trivial way.

Open Problem 3. Is Theorem 3 helpful (and how) in characterizing more classes of CM and JM distributions?

2.4. Norm condition

Below we discuss the relationship between mixability and law-determined norms. First we give the definition of a law-determined norm.

Definition 5 (Law-determined norm). A law-determined norm $||\cdot||$ on L^0 maps L^0 to $[0, \infty]$, such that

  (i) $||aX|| = |a| \cdot ||X||$ for $a \in \mathbb{R}$ and $X \in L^0$;
 (ii) $||X + Y|| \leqslant ||X|| + ||Y||$ for $X, Y \in L^0$;
(iii) $||X|| = 0$ implies $X = 0$ a.s.;
 (iv) $||X|| = ||Y||$ if $X \stackrel{d}{=} Y$, $X, Y \in L^0$;
  (v) $||X|| \leqslant ||Y||$ if $0 \leqslant X \leqslant Y$ a.s.

The L^p-norms $p \in [1, \infty)$, $|| \cdot ||_p : L^0 \to [0, \infty]$, $X \mapsto (\mathbb{E}[|X|^p])^{1/p}$ and the L^∞-norm $|| \cdot ||_\infty : L^0 \to [0, \infty]$, $X \mapsto \text{ess-sup}(|X|)$ are law-determined norms. Here, we allow $|| \cdot ||$ to take a value of ∞, which means that the non-negative functional $|| \cdot ||$ is not necessarily a norm in the common sense; we slightly abuse the terminology here since all natural examples are norms in respective proper spaces. We obtain a necessary condition for complete and joint mixability based on law-determined norms. In what follows, $(\cdot)_+ = \max\{\cdot, 0\}$ and $(\cdot)_- = -\min\{\cdot, 0\}$.

Theorem 4. *Suppose that $(F_1, \ldots, F_n)$ is JM with joint center K, $X_i \sim F_i$, $i = 1, \ldots, n$, $|| \cdot ||$ is any law-determined norm. Then we have that*

$$||(X_i - \mu_i)_+|| \leqslant \sum_{j=1, j \neq i}^{n} ||(X_j - \mu_j)_-|| \quad and \quad ||(X_i - \mu_i)_-|| \leqslant \sum_{j=1, j \neq i}^{n} ||(X_j - \mu_j)_+||,$$
(12)

for all $i = 1, \ldots, n$, and all $\mu_1, \ldots, \mu_n \in \mathbb{R}$ such that $\mu_1 + \cdots + \mu_n = K$.

Proof. Since $(F_1, \ldots, F_n)$ is JM, there exist random variables $X_1 \sim F_1, \ldots, X_n \sim F_n$ such that $X_1 + \cdots + X_n = K$. It follows that $X_1 - \mu_1 = -((X_2 + \cdots + X_n) - (\mu_2 + \cdots + \mu_n))$ and hence

$$(X_1 - \mu_1)_+ = \left(\sum_{i=2}^{n} (X_i - \mu_i) \right)_- \leqslant \sum_{i=2}^{n} (X_i - \mu_i)_-.$$
(13)

Applying $|| \cdot ||$ on both sides of (13), we obtain

$$||(X_1 - \mu_1)_+|| = \left\| \left(\sum_{i=2}^{n} (X_i - \mu_i) \right)_- \right\| \leqslant \left\| \sum_{i=2}^{n} (X_i - \mu_i)_- \right\| \leqslant \sum_{i=2}^{n} ||(X_i - \mu_i)_-||.$$

The rest parts are obtained by symmetry. $\square$

A similar version of Theorem 4 for complete mixability is listed below.

Theorem 5. *Suppose that F is n-CM with center μ, $X \sim F$ and $|| \cdot ||$ is any law-determined norm. Then we have that*

$$||(X - t)_+|| \leqslant (n-1)||(X - s)_-|| \quad and \quad ||(X - t)_-|| \leqslant (n-1)||(X - s)_+||, \quad (14)$$

for all $t \in \mathbb{R}$ and $s = (n\mu - t)/(n-1)$.

It is worth noting that if we take $|| \cdot || = || \cdot ||_\infty$ and $s = t = \mu$ in Theorem 5, then we obtain that

$$||(X - \mu)_+||_\infty \leqslant (n-1)||(X - \mu)_-||_\infty,$$

which is $b - \mu \leqslant (n-1)(\mu - a)$, where $a = \sup\{t \in \mathbb{R} : F(t) = 0\}$ and $b = \inf\{t \in \mathbb{R} : F(t) = 1\}$. Combining with the other inequality in (14), we obtain

$$a + \frac{b - a}{n} \leqslant \mu \leqslant b - \frac{b - a}{n}.$$
(15)

(15) is the *mean condition* obtained in [43], one of the key necessary conditions for complete mixability, and is also a sufficient condition if F has monotone density, shown in [43]; see also (b) in Section 2.2. If we take $|| \cdot || = || \cdot ||_2$ and $\mu_i = \mathbb{E}[X_i]$ in Theorem 4, and assume that each F_i is $N(\mu_i, \sigma_i^2)$, $i = 1, \ldots, n$, then we obtain that

$$2 \max_{i=1,\ldots,n} \sigma_i \leqslant \sum_{i=1}^{n} \sigma_i.$$

which is a sufficient condition for $F_1, \ldots, F_n$ to be JM, shown in [46]; see also (f) in Section 2.2. Both examples indicate that special cases of (12) and (14) may be sufficient for particular classes of distributions.

Open Problem 4. Suppose that F has mean μ, and (14) holds for all law-determined norms $|| \cdot ||$ and all $t \in \mathbb{R}$. (With what extra conditions, possibly some smoothness conditions) is it sufficient for F to be n-CM?

This problem induces another interesting question which is not directly related to mixability: how can we characterize all possible law-determined norms in Definition 5?

2.5. Mixability in vector spaces

Most of the literature has a focus on complete and joint mixability on $\mathbb{R}$ for its relevance in applications. Clearly, concepts of mixability can be naturally generalized to distributions on $\mathbb{R}^d$. However, existing non-trivial results in the multi-dimensional setting are very limited; an early study in this direction can be found in [40].

A simple observation is listed below. Its proof is straightforward and omitted.

Proposition 6. *Suppose that F on $\mathbb{R}^d$ is n-CM. Then the projection of F to any subspace of $\mathbb{R}^d$ is n-CM.*

We have the following conjecture, with its rationale explained below.

Open Problem 5. Is a uniform distribution on a convex set $C \subset \mathbb{R}^d$ necessarily n-CM for all $n \geqslant d + 1$?

Obviously, the above conjecture is equivalent to say that a uniform distribution on a convex set $C \subset \mathbb{R}^d$ is $(d + 1)$-CM. The trivial cases $d = 0$ and $d = 1$ are explained below. When $d = 0$, C degenerates to a singleton, on which a distribution is always n-CM for $n \geqslant 1$. When $d = 1$, C is an interval, and a uniform distribution on an interval is n-CM for $n \geqslant 2$; this was already shown in [18]. When $d = 2$, any projection of a uniform distribution on C to a line has a concave density. [32] showed that a distribution with a concave density is n-CM for $n \geqslant 3$. Of course, this is not sufficient for such a distribution to be n-CM on $\mathbb{R}^d$. However, we wonder how this type of dimension reduction would help to characterize complete mixability.

Open Problem 6. Suppose that F is a distribution on $\mathbb{R}^d$, $d > 1$, and the projection of F to any essential subspace of $\mathbb{R}^d$ is n-CM. (With what extra conditions) is it sufficient for F to be n-CM on $\mathbb{R}^d$?

2.6. Asymptotic mixability

Let F be an arbitrary distribution with bounded support. It has been observed that [e.g. 3, 45] when n is large, it is more likely that F becomes n-CM. [33] showed that a distribution on a bounded interval $[a, b]$ with a positive density function $f(x) \geqslant 3/(n(b-a))$ is n-CM. As a consequence, any continuous distribution with a density bounded away from zero is n-CM for n sufficiently large. It is left open to answer whether this condition of a density bounded away from zero can be removed.

Open Problem 7. Are all absolutely continuous distributions on a bounded interval n-CM for large enough n?

2.7. Copula of a complete mix

The major results in [43, 32, 33, 44] are based on combinatorics and mathematical induction. The dependence structure hidden in the proofs are unclear. It was noted in [43] that an explicit form of a copula (which is generally not unique) of a complete or joint mix is very difficult to write down. Since complete and joint mixability naturally give bounds to many optimization problems, it would be nice to have a copula of a complete mix, or a sampling method for simulation. The two questions are of course very much related.

Open Problem 8. Suppose that F satisfies one of the sufficient conditions (for instance, (b) in Section 2.2) and hence is n-CM. What is a possible copula of a n-complete mix with margins F (or a joint mix with given margins)?

Open Problem 9. Suppose that F satisfies one of the sufficient conditions (for instance, (b) in Section 2.2) and hence is n-CM. Could we simulate sample from a n-complete mix with margins F?

2.8. Characterizing more classes of CM/JM distributions

It is a general task to characterize more classes of CM/JM distributions with their corresponding necessary and sufficient conditions. One particular question often discussed concerns the unimodal densities, as it is relevant to many optimization problems outlined in [2]. [34] gave counter-examples where the mean condition (15) is not sufficient for the complete mixability of a distribution with a unimodal density on a bounded interval.

Open Problem 10. Under what extra conditions a unimodal distribution on a bounded interval is n-CM?

This question is particularly relevant to optimization problems when one of the inequalities in the mean condition (15) is attained by an equality, as noted in [2]. That is, the mean of the distribution exactly divides the support $[a, b]$ of the distribution into two parts with lengths $(b-a)/n$ and $(b-a)(n-1)/n$, respectively.

2.9. Convex order problems

When $(F_1, \ldots, F_n)$ is not JM, it is generally not clear whether there exists an element $S_0 \in \mathcal{D}_n$ such that $S_0 \prec_{\mathrm{cx}} S$ for all $S \in \mathcal{D}_n$, where $\prec_{\mathrm{cx}}$ stands for convex order in Definition 4. A counter-example is given in [2] showing an aggregation set $\mathcal{D}_n$ does not necessarily contain a smallest element with respect to convex order. However, for all commonly used distributions $F_1, \ldots, F_n$, $\mathcal{D}_n$ seems to contain such a smallest element, as shown either theoretically or numerically. For instance, if each F_i has a decreasing density, $i = 1, \ldots, n$, then a smallest element with respect to convex order in $\mathcal{D}_n$ can be obtained; this was shown in [21]. It remains unclear under what conditions such a smallest element exists.

Open Problem 11. What are necessary and sufficient conditions for $\mathcal{D}_n$ to contain a smallest element with respect to convex order?

2.10. Characterizing the aggregation set

The last question is a general question concerning Fréchet classes. We use the aggregation set $\mathcal{D}_n$ as in the previous problem, and define $\mathcal{D}_n^* = \{S/n : S \in \mathcal{D}_n\}$. It is obvious that the joint mixability of $F_1, \ldots, F_n$ is equivalent to the inclusion of a degenerate random variable in $\mathcal{D}_n$. In the case when $F = F_1 = F_2 = \cdots$ and F has finite mean, [25] showed that $\mathcal{D}_n^*$ has an upper limit of $\mathcal{C}_F = \{S : S \prec_{\mathrm{cx}} X, \ X \sim F\}$ as $n \to \infty$. However, it is also noted that for a finite n, $\mathcal{D}_n^* \subset \mathcal{C}_F$ but is generally not equal to $\mathcal{C}_F$. The only fully-characterized classes of $\mathcal{D}_n$ are when $n = 2$ and the marginal distributions are Bernoulli; see [25].

Open Problem 12. How can one characterize $\mathcal{D}_n$ (maybe for some simple marginal distributions)? That is, for a given distribution G, determine whether $S \in \mathcal{D}_n$ for some $S \sim G$.

This question summarizes all challenges in complete and joint mixability. It is generally open for all $n \geqslant 2$.

2.11. Some other open questions

We conclude this paper by some other questions that are beyond the expertise of the author. To avoid misleading the reader with the author's naivety and ignorance, we simply list some possible directions.

1. Algorithms to determine whether some distributions are jointly mixable, or solving question (1) in Section 1: see for instance [28, 12, 34, 19]. The conditions under which the swapping algorithms in [28] converges are still

unclear. Interestingly, [19] showed that the determination of the joint mixability of different discrete uniform distributions on $\mathbb{Z}$ is NP-complete.
2. Mixability under higher-dimensional constraints: for fixed bivariate or higher-dimensional marginal distributions, determine whether a joint mix exists and develop algorithms for numerical determination. Note that even to justify the existence of a joint distribution with given multivariate margins is not easy; see for instance [41, 22, 11].
3. Other multivariate functions replacing the summation of random variables in the definition of mixability; see [4].
4. The influence of the algebraic structure of a semigroup $\mathbb{G}$ on complete and joint mixability defined on $\mathbb{G}$.

Acknowledgement

References

[1] AAS, K. and PUCCETTI, G. (2014). Bounds for total economic capital: the DNB case study. *Extremes*, **17**(4), 693–715. MR3282692

[2] BERNARD, C., JIANG, X. and WANG, R. (2014). Risk aggregation with dependence uncertainty. *Insurance: Mathematics and Economics*, **54**, 93–108. MR3145855

[3] BERNARD, C., RÜSCHENDORF, L. and VANDUFFEL, S. (2013). Value-at-Risk bounds with variance constraints. *Journal of Risk and Insurance*, forthcoming. Available at http://ssrn.com/abstract=2342068.

[4] BIGNOZZI, V. and PUCCETTI, G. (2014). Characterizing mixability of supermodular functions. *Statistics and Probability Letters*, **100**, 48–55. MR3324074

[5] CHEUNG, K. C., and LO, A. (2013). General lower bounds on convex functionals of aggregate sums. *Insurance: Mathematics and Economics*, **53**(3), 884–896. MR3130483

[6] CHEUNG, K. C., and LO, A. (2014). Characterizing mutual exclusivity as the strongest negative multivariate dependence structure. *Insurance: Mathematics and Economics*, **55**, 180–190. MR3179810

[7] CHEUNG, K. C. and VANDUFFEL, S. (2013). 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. MR3041120

[8] DHAENE, J. and DENUIT, M. (1999). The safest dependence structure among risks. *Insurance: Mathematics and Economics*, **25**(1), 11–21. MR1718527

[9] DHAENE, J., DENUIT, M., GOOVAERTS, M. J., KAAS, R. and VYNCHE, D. (2002). The concept of comonotonicity in actuarial science and finance: theory. *Insurance: Mathematics and Economics*, **31**(1), 3–33. MR1956509

[10] EMBRECHTS, P. and PUCCETTI, G. (2006). Bounds for functions of dependent risks. *Finance and Stochastics*, **10**, 341–352. MR2244349

[11] EMBRECHTS, P. and PUCCETTI, G. (2009). Bounds for the sum of dependent risks having overlapping marginals. *Journal of Multivariate Analysis*, **101**(1), 177–190. MR2557627

[12] EMBRECHTS, P., PUCCETTI, G. and RÜSCHENDORF, L. (2013). Model uncertainty and VaR aggregation. *Journal of Banking and Finance*, **37**(8), 2750–2764.

[13] EMBRECHTS, P., PUCCETTI, G., RÜSCHENDORF, L., WANG, R. and BELERAJ, A. (2014a). An academic response to Basel 3.5. *Risks*, **2**(1), 25–48.

[14] EMBRECHTS, P., WANG, B. and WANG, R. (2015). Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance and Stochastics*, forthcoming. Available at http://ssrn.com/abstract=2558525. MR3321476

[15] FANG, K. T., KOTZ, S. and NG, K. W. (1990). *Symmetric Multivariate and Related Distributions*. London: Chapman & Hall. MR1071174

[16] FÖLLMER, H. and SCHIED, A. (2011). *Stochastic Finance: An Introduction in Discrete Time*. Walter de Gruyter, Third Edition. MR2779313

[17] FRÉCHET, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon. Sect. A. (3)*, **14**, 53–77. MR0049518

[18] GAFFKE, N. and RÜSCHENDORF, L. (1981). On a class of extremal problems in statistics. *Math. Operationsforsch. Statist. Ser. Optim.* **12**(1), 123–135. MR0609133

[19] HAUS, U. (2015). Bounding stochastic dependence, complete mixability of matrices, and multidimensional bottleneck assignment problems. *Operations Research Letters*, **43**(1), 74—79. MR3302551

[20] HOEFFDING, W. (1940). Massstabinvariante korrelationtheorie. *Schriften Math. Inst. Univ. Berlin*, **5**(5), 181–233. MR0004426

[21] JAKOBSONS, E., HAN, X. and WANG, R. (2015). General convex order on risk aggregation. *Scandinavian Actuarial Journal*, forthcoming. DOI: 10.1080/03461238.2015.1012223.

[22] JOE, H. (1997). *Multivariate Models and Dependence Concepts*. London: Chapman & Hall. MR1462613

[23] KNOTT, M. and SMITH, C. S. (2006). Choosing joint distributions so that the variance of the sum is small. *Journal of Multivariate Analysis*, **97**(8), 1757–1765. MR2298887

[24] LEE, W. and AHN, J. Y. (2014). On the multidimensional extension of countermonotonicity and its applications. *Insurance: Mathematics and Economics*, **56**, 68–79. MR3208102

[25] MAO, T. and WANG, R. (2014). On aggregation sets and lower-convex sets. *Journal of Multivariate Analysis*, **138**, 170–181. MR3348840

[26] McNeil, A. J., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, Tools*. Princeton University Press, Princeton, NJ. MR2175089

[27] Müller, A. and Stoyan, D. (2002). *Comparison Methods for Statistical Models and Risks*. Wiley, England. MR1889865

[28] Puccetti, G. and Rüschendorf, L. (2012a). Computation of sharp bounds on the distribution of a function of dependent risks. *Journal of Computational and Applied Mathematics*, **236**(7), 1833–1840. MR2863518

[29] Puccetti, G. and Rüschendorf, L. (2012b). Bounds for joint portfolios of dependent risks. *Statistics and Risk Modeling*, **29**(2), 107–131. MR2945764

[30] Puccetti, G. and Rüschendorf, L. (2013). Sharp bounds for sums of dependent risks. *Journal of Applied Probability*, **50**(1), 42–53. MR3076771

[31] Puccetti, G. and Rüschendorf, L. (2014). Asymptotic equivalence of conservative value-at-risk-and expected shortfall-based capital charges. *Journal of Risk*, **16**(3), 3–22.

[32] Puccetti, G., Wang, B. and Wang, R. (2012). Advances in complete mixability. *Journal of Applied Probability*, **49**(2), 430–440. MR2977805

[33] Puccetti, G., Wang, B. and Wang, R. (2013). Complete mixability and asymptotic equivalence of worst-possible VaR and ES estimates. *Insurance: Mathematics and Economics*, **53**(3), 821–828. MR3130477

[34] Puccetti, G. and Wang, R. (2015a). Detecting complete and joint mixability. *Journal of Computational and Applied Mathmatics*, **280**, 174–187. MR3296054

[35] Puccetti, G. and Wang, R. (2015b). Extremal dependence concepts. *Statistical Science*, forthcoming. Available at http://ssrn.com/abstract=2436392.

[36] Ramachandran, D. and Rüschendorf, L. (1995). A general duality theorem for marginal problems. *Probability Theory and Related Fields*, **101**(3), 311–319. MR1324088

[37] Rüschendorf, L. (1982). Random variables with maximum sums. *Advances in Applied Probability*, **14**(3), 623–632. MR0665297

[38] Rüschendorf, L. (1983). Solution of a statistical optimization problem by rearrangement methods. *Metrika*, **30**, 55–61. MR0701979

[39] Rüschendorf, L. (2013). *Mathematical Risk Analysis. Dependence, Risk Bounds, Optimal Allocations and Portfolios*. Springer, Heidelberg. MR3051756

[40] Rüschendorf, L. and Uckelmann, L. (2002). Variance minimization and random variables with constant sum. In *Distributions with Given Marginals and Statistical Modelling*, pp. 211–222. Dordrecht: Kluwer Academic Publishers. MR2058994

[41] Strassen, V. (1965). The existence of probability measures with given marginals. *Annals of Mathematical Statistics*, **36**(2), 423–439. MR0177430

[42] Tchen, A. H. (1980). Inequalities for distributions with given marginals. *Annals of Probability*, **8**(4), 814–827. MR0577318

[43] Wang, B. and Wang, R. (2011). The complete mixability and convex minimization problems with monotone marginal densities. *Journal of Multivariate Analysis,* **102**(10), 1344–1360. MR2819953

[44] Wang, B. and Wang, R. (2015). Joint mixability. *Mathematics of Operations Research*, forthcoming. Available at http://ssrn.com/abstract=2557067.

[45] Wang, R. (2014). Asymptotic bounds for the distribution of the sum of dependent random variables. *Journal of Applied Probability*, **51**(3), 780–798. MR3256227

[46] Wang, R., Peng, L. and Yang, J. (2013). Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. *Finance and Stochastics*, **17**(2), 395–417. MR3038596