J. Appl. Prob. **49,** 430–440 (2012)
Printed in England

ADVANCES IN COMPLETE MIXABILITY

GIOVANNI PUCCETTI,* *University of Firenze*
BIN WANG,** *Peking University*
RUODU WANG,*** *Georgia Institute of Technology*

Abstract

The concept of complete mixability is relevant to some problems of optimal couplings with important applications in quantitative risk management. In this paper we prove new properties of the set of completely mixable distributions, including a completeness and a decomposition theorem. We also show that distributions with a concave density and radially symmetric distributions are completely mixable.

Keywords: Complete mixability; multivariate dependence; concave density; radially symmetric distribution; optimal coupling

2010 Mathematics Subject Classification: Primary 60E05
Secondary 91B30

1. Introduction

A distribution function F is called *n-completely mixable* (n-CM) if there exist n random variables $X_1, \dots, X_n$, identically distributed as F, having constant sum, that is, satisfying

$$\mathrm{P}(X_1 + \cdots + X_n = nk) = 1.$$

If F has finite first moment μ then $k = \mu$. The concept of complete mixability is related to some optimization problems in the theory of optimal couplings.

(i) Assume that F has finite first moment μ. For a (strictly) convex function $f : \mathbb{R} \to \mathbb{R}$, we have

$$\inf\{\mathrm{E}[f(X_1 + \cdots + X_n)];\ X_i \sim F,\ 1 \le i \le n\} \ge f(n\mu), \tag{1.1}$$

and the equality holds if (and only if) F is n-CM.

(ii) Assume that F has finite first moment, and let F^{-1} be the generalized inverse of F. Define the function $\Psi(a) = \mathrm{E}[X \mid X \ge F^{-1}(a)]$ for $a \in [0, 1]$ and $X \sim F$. For any $s \in \mathbb{R}$, we have

$$\sup\{\mathrm{P}(X_1 + \cdots + X_n \ge s);\ X_i \sim F,\ 1 \le i \le n\} \le 1 - \Psi^{-1}\left(\frac{s}{n}\right), \tag{1.2}$$

and the sup is attained if and only if F is n-CM on the interval $(F^{-1}(\Psi^{-1}(s/n)), F^{-1}(1))$.

Received 13 September 2011.
* Postal address: Department of Mathematics for Decision Theory, University of Firenze, via Lombroso 6/17, 50134 Firenze, Italy.
** Postal address: Department of Mathematics, Peking University, Beijing, 100871, P. R. China.
*** Postal address: School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA. Email address: ruodu.wang@math.gatech.edu

For more details on the solutions of these problems and a brief history of the concept of the complete mixability, we refer the reader to the recent papers Wang and Wang (2011) and Wang *et al.* (2011).

Problems (1.1) and (1.2) have relevant applications in quantitative risk management, where they are needed to assess the aggregate risk of a portfolio of losses for regulatory issues. For more details on the motivation of these problems within quantitative risk management, we refer the reader to Embrechts and Puccetti (2010). Other important applications are related to the theory of dependence measures; see Nelson and Úbeda-Flores (2012).

In view of these applications, it would be of interest to characterize the class of completely mixable distributions. Only partial characterizations, which we summarize in Section 2, are known in the literature. In our paper we give a contribution in the direction of a complete characterization of completely mixable distributions. In Section 3 we give a completeness and a decomposition theorem for completely mixable distributions. In Sections 4 and 5 we prove complete mixability for two new classes of distributions, namely continuous distributions with a concave density and radially symmetric distributions.

2. Some preliminaries on complete mixability

In this section we give a summary of the existing results on completely mixable distributions which we will use in what follows. Throughout the paper, we identify probability measures with the corresponding distribution functions.

Definition 2.1. A distribution function F on $\mathbb{R}$ is called *n-completely mixable* (n-CM) if there exist n random variables $X_1, \dots, X_n$ identically distributed as F such that

$$\mathrm{P}(X_1 + \cdots + X_n = nk) = 1 \tag{2.1}$$

for some $k \in \mathbb{R}$. Any such k is called a center of F and any vector $(X_1, \dots, X_n)$ satisfying (2.1) with $X_i \sim F$, $1 \le i \le n$, is called an *n-complete mix.*

If F is n-CM and has finite first moment μ, then its center is unique and equal to μ. We denote by $\mathcal{M}_n(\mu)$ the set of all n-CM distributions with center μ, and by $\mathcal{M}_n = \bigcup_{\mu \in \mathbb{R}} \mathcal{M}_n(\mu)$ the set of all n-CM distributions on $\mathbb{R}$. As proved in Wang and Wang (2011), the set $\mathcal{M}_n(\mu)$ is convex, while the set $\mathcal{M}_n$ is not. Some straightforward examples of completely mixable distributions are given in Wang and Wang (2011).

Proposition 2.1. (Wang and Wang (2011).) *The following statements hold.*

(a) *F is 1-CM if and only if F is the distribution of a constant.*

(b) *F is 2-CM if and only if F is symmetric, i.e. $X \sim F$ and $a - X \sim F$ for some constant $a \in \mathbb{R}$.*

(c) *Any linear transformation of an n-CM distribution is n-CM.*

(d) *The binomial distribution $B(n, p/q)$, $p, q \in \mathbb{N}$, is q-CM.*

(e) *The uniform distribution on the interval (a, b) is n-CM for any $n \ge 2$ and $a < b$.*

(f) *The Gaussian and the Cauchy distributions are n-CM for $n \ge 2$.*

Some other families of completely mixable distribution are described in the following theorems.

https://doi.org/10.1239/jap/1339878796 Published online by Cambridge University Press

Theorem 2.1. (Rüschendorf and Uckelmann (2002).) *Any continuous distribution function having a symmetric and unimodal density is n-CM for any* $n \geq 2$.

Theorem 2.2. (Wang and Wang (2011).) *Suppose that* F *is a distribution function on the real interval* $[a, b]$ *having mean* μ, $a = \sup\{t\colon F(t) = 0\}$, *and* $b = \inf\{t\colon F(t) = 1\}$. *A necessary condition for* F *to be n-CM is that*

$$a + \frac{b-a}{n} \leq \mu \leq b - \frac{b-a}{n}. \tag{2.2}$$

If F *is also continuous with a monotone density on* $[a, b]$, *condition (2.2) is also sufficient.*

For example, according to Theorem 2.2, the $\mathrm{Beta}(\alpha, \beta)$ distribution with parameters $\alpha, \beta > 0$ satisfying $(\alpha - 1)(\beta - 1) \leq 0$ and $1/n \leq \alpha/(\alpha + \beta) \leq (n-1)/n$ is n-CM.

3. Completeness and decomposition theorems

In this section we show that any n-CM distribution can be obtained as the limit of a convex combination of discrete n-CM distributions. First, we show that the sets $\mathcal{M}_n(\mu)$ and $\mathcal{M}_n$ are complete under weak convergence, that is, any n-CM distributions can be seen as the limit of n-CM discrete distributions.

Theorem 3.1. *The following statements hold for weak convergence.*

(a) *The limit of a sequence of n-CM distribution functions (with center* μ*) is n-CM (with center* μ*).*

(b) *Any n-CM distribution function with center* μ *is the limit of a sequence of discrete n-CM distribution functions with center* μ.

(c) *A distribution function is n-CM (with center* μ*) if and only if it is the limit of a sequence of discrete n-CM distribution functions (with center* μ*).*

Proof. (a) Denote by F^k, $k \in \mathbb{N}$, a sequence of n-CM distributions having limit F. Since $F^k \in \mathcal{M}_n$ for any $k \in \mathbb{N}$, it is possible to find $X_1^k, \dots, X_n^k$ such that $X_i^k \sim F^k$, $1 \leq i \leq n$, and

$$\mathrm{P}(X_1^k + \cdots + X_n^k = c_k) = 1 \tag{3.1}$$

for some $c_k \in \mathbb{R}$. As $F^k \xrightarrow{\mathrm{w}} F$, there also exist n random variables $X_1, \dots, X_n$, identically distributed as F, for which $X_i^k \xrightarrow{\mathrm{w}} X_i$, $1 \leq i \leq n$, and, therefore, such that

$$(X_1^k + \cdots + X_n^k) \xrightarrow{\mathrm{w}} (X_1 + \cdots + X_n). \tag{3.2}$$

Combining (3.1) and (3.2), we find that $X_1 + \cdots + X_n = c = \lim c_k$ holds almost surely. Since $X_i \sim F$, $1 \leq i \leq n$, this implies that F is n-CM. If we have $c_k = n\mu$ for all $k \in \mathbb{N}$ then $c = n\mu$.

(b) Let $\boldsymbol{X} = (X_1, \dots, X_n)$ be an n-complete mix on $\mathbb{R}^n$ with $X_i \sim F$, $1 \leq i \leq n$, and

$$X_1 + \cdots + X_n = n\mu \quad \text{almost surely.}$$

As $\boldsymbol{X}$ is supported on the set $S_n(\mu) = \{\boldsymbol{x} \in \mathbb{R}^n\colon \sum_{i=1}^n x_i = n\mu\} \subset \mathbb{R}^n$, we can find a sequence F^k, $k \in \mathbb{N}$, of discrete distributions on $S_n(\mu)$ converging weakly to the distribution of $\boldsymbol{X}$. The theorem follows by noting that F_1^k, the first marginal of F^k, is n-CM since F^k is supported on $S_n(\mu)$ and the sequence F_1^k, $k \in \mathbb{N}$, converges weakly to F.

(c) This is a corollary of points (a) and (b).

Now, we prove a decomposition theorem for n-CM distributions. In the following, we call an *n-discrete uniform* distribution a uniform distribution on n points, that is, giving mass $1/n$ at each of the n points in its support

Lemma 3.1. *An n-discrete uniform distribution is n-CM.*

Proof. Let F be an n-discrete uniform distribution on the points $y_1, \dots, y_n$. Let $\boldsymbol{X} = (X_1, \dots, X_n)$ be a random vector uniformly distributed on the $n!$ vectors

$$(y_{\pi(1)}, \dots, y_{\pi(n)}), \qquad \pi \in \mathcal{P}_n,$$

where $\mathcal{P}_n$ is the set of all permutations of $\{1, \dots, n\}$. In the support of $\boldsymbol{X}$, there are exactly $(n-1)!$ vectors having the value y_j as the ith component. Therefore, we have

$$\mathrm{P}(X_i = y_j) = \frac{(n-1)!}{n!} = \frac{1}{n}, \qquad 1 \leq i, j \leq n.$$

As a consequence, $\boldsymbol{X}$ has marginal distributions identically distributed as F. Since $\sum_{i=1}^{n} y_{\pi(i)}$ is constant on π, $\boldsymbol{X}$ is an n-complete mix and F is n-CM.

Let $\mathcal{M}_n^S(\mu)$ be the set of all n-discrete uniform distributions with mean μ, and let $L(\mathcal{M}_n^S(\mu))$ be the set of all countable convex combinations of elements in $\mathcal{M}_n^S(\mu)$, that is,

$$L(\mathcal{M}_n^S(\mu)) = \left\{ \sum_{k=1}^{\infty} a_k F^k;\ F^k \in \mathcal{M}_n^S(\mu),\ a_k \geq 0,\ \sum_{k=1}^{\infty} a_k = 1 \right\}.$$

We show that any discrete n-CM distribution can be obtained as the countable convex combination of n-discrete uniform distributions.

Theorem 3.2. *The following statements hold.*

(a) *The countable convex combination of n-CM distribution functions with center μ is n-CM with center μ.*

(b) *If F is discrete then $F \in \mathcal{M}_n(\mu)$ if and only if $F \in L(\mathcal{M}_n^S(\mu))$.*

(c) *If $F \in L(\mathcal{M}_n^S(\mu))$ with $F = \sum_{k \in \mathbb{N}} a_k F^k$, the joint distribution G of an n-complete mix with marginals F is given by*

$$G(x_1, \dots, x_n) = \sum_{k \in \mathbb{N}} \frac{a_k}{n!} \prod_{i=1}^{n} [nF^k(x_{[i]}) - i + 1]^+,$$

where $x_{[i]}$ is the ith order statistic of $\{x_1, \dots, x_n\}$.

Proof. (a) The statement for finite convex combinations follows by induction from Proposition 2.1(3) of Wang and Wang (2011). Now let a_k, $k \in \mathbb{N}$, be a sequence of nonnegative values with $\sum_{k=1}^{+\infty} a_k = 1$, and let $F^k \in \mathcal{M}_n(\mu)$, $k \in \mathbb{N}$, be a sequence of n-CM distributions having center μ. Without loss of generality, we can assume that $a_1 > 0$ and define the new sequence

$$G^k = \frac{\sum_{i=1}^{k} a_i F^i}{\sum_{i=1}^{k} a_i}, \qquad k \in \mathbb{N}.$$

Any G^k is the finite convex sum of n-CM distributions; thus, it is n-CM. Since $G^k \xrightarrow{\text{w}} G = \sum_{k=1}^{+\infty} a_k F^k$, G is n-CM by Theorem 3.1(a).

(b) The inclusion $L(\mathcal{M}_n^S(\mu)) \subset \mathcal{M}_n(\mu)$ follows from (a). Then it is sufficient to show that $\mathcal{M}_n(\mu) \subset L(\mathcal{M}_n^S(\mu))$. Let $\boldsymbol{X} = (X_1, \dots, X_n)$ be a complete mix with center μ and discrete marginals identically distributed as F. Denoting by $\{\boldsymbol{x}^j,\ j \in A \subset \mathbb{N}\}$ the countable support of $\boldsymbol{X}$, we have

$$
\begin{aligned}
F(s) &= \frac{1}{n}\sum_{i=1}^{n} \mathrm{P}(X_i \leq s) \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j\in A} \mathrm{P}(X_i \leq s \mid \boldsymbol{X} = \boldsymbol{x}^j)\,\mathrm{P}(\boldsymbol{X} = \boldsymbol{x}^j) \\
&= \sum_{j\in A} \mathrm{P}(\boldsymbol{X} = \boldsymbol{x}^j)\biggl(\frac{1}{n}\sum_{i=1}^{n} \mathrm{P}(X_i \leq s \mid \boldsymbol{X} = \boldsymbol{x}^j)\biggr) \\
&= \sum_{j\in A} a_j \biggl(\frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{x_i^j \leq s\}}\biggr),
\end{aligned}
$$

where x_i^j denotes the ith component of the vector $\boldsymbol{x}^j$ and $a_j = \mathrm{P}(\boldsymbol{X} = \boldsymbol{x}_j)$, $j \in A$. Note that the a_js are nonnegative, $\sum_{j\in A} a_j = 1$, and, for any $j \in A$, the function $\sum_{i=1}^{n} \mathbf{1}_{\{x_i^j \leq s\}}$ is the distribution function of a random variable uniformly distributed on $\{x_1^j, \dots, x_n^j\}$. With $\boldsymbol{X}$ being an n-complete mix, we have $\sum_{i=1}^{n} x_i^j = n\mu$ when $a_j > 0$. As a result, F can be written as a countable convex sum of distributions in $\mathcal{M}_n^S(\mu)$, that is, $F \in L(\mathcal{M}_n^S(\mu))$.

(c) First, note that G has marginals identically distributed as F since

$$
\lim_{x_i \to +\infty,\, i \neq j} R(x_1, \dots, x_n) = \sum_{k\in\mathbb{N}} a_k F^k(x_j) = F(x_j), \qquad 1 \leq j \leq n.
$$

In order to show that G is the distribution of an n-complete mix, we prove that

$$
G^k(x_1, \dots, x_n) = \frac{1}{n!}\prod_{i=1}^{n}[nF^k(x_{[i]}) - i + 1]^+
$$

is the distribution of an n-complete mix with center μ for any $k \in \mathbb{N}$.

Since $F^k \in \mathcal{M}_n^S(\mu)$, there exist $y_1^k \leq \cdots \leq y_n^k$ such that $\sum_{i=1}^{n} y_i^k = n\mu$ and $F^k(y_i^k) = (1/n)\sum_{j=1}^{n} \mathbf{1}_{\{y_j^k \leq y_i^k\}}$. Noting that

$$
\frac{1}{n!}\prod_{i=1}^{n}[nF^k(x_{[i]}) - i + 1]^+ = \frac{1}{n!}\sum_{\pi\in\mathcal{P}_n} \mathbf{1}_{\{y_{\pi(i)}^k \leq x_i,\ 1\leq i\leq n\}},
$$

we deduce that, for any $k \in \mathbb{N}$, G^k is uniformly distributed on the $n!$ vectors

$$
(y_{\pi(1)}^k, \dots, y_{\pi(n)}^k), \qquad \pi \in \mathcal{P}_n,\ k \in \mathbb{N}.
$$

Thus, G^k is the distribution of an n-complete mix with center $(1/n)\sum_{i=1}^{n} y_i^k = \mu$, from which it also follows that $G = \sum_{k\in\mathbb{N}} a_k G^k$ is the distribution of an n-complete mix with center μ.

Remark 3.1. There are some points to remark about Theorem 3.2.

(a) Similarly to as in the proof of Theorem 3.2(b), we can show that an arbitrary n-CM distribution with center μ can be written as an integral of n-discrete uniform distributions with center μ.

(b) Using the notation introduced in the proof of Theorem 3.2(c), the distribution G can be seen as the distribution of the random variable $\sum_{k\in\mathbb{N}} \mathbf{1}_{\{Z=k\}}\, G^k$, where Z is a discrete random variable giving mass a_k to $k \in \mathbb{N}$ and independent from the G^ks. Note, however, that the distribution of an n-complete mix for a discrete F may not be unique.

(c) A number of the n points of the support of an n-discrete distribution can be chosen to be equal. The set of n-discrete uniform distributions therefore includes all distributions giving masses $(k/n),\ k \in \mathbb{N}$, to at most n different points.

(d) The convex combination of n-discrete distributions with different centers may fail to be n-CM. For example, the Bernoulli distribution $F(s) = (\mathbf{1}_{\{0\le s\}} + \mathbf{1}_{\{1\le s\}})/2$ is the convex sum of two 1-CM distributions but it is not 1-CM. Therefore, the assumption of a common center cannot be dropped in Theorem 3.2(a)–(c).

As a corollary of Theorem 3.1(c) and Theorem 3.2(b), we present the main result of this section.

Corollary 3.1. *A distribution is n-CM with center μ if and only if it is the limit of a sequence of countable convex combinations of n-discrete uniform distributions with center μ.*

4. Distributions with a concave density

In this section we show that any continuous distribution with a concave density is completely mixable. Similarly to the method used in the proof of Theorem 2.4 of Wang and Wang (2011), we will first prove complete mixability of a particular class of discrete distributions with concave mass function.

Theorem 4.1. *Suppose that F is a discrete distribution on the set*

$$S_{N,M} = \{-N, -N+1, \dots, -1, 0, 1, \dots, M-1, M\}, \qquad N, M \in \mathbb{N}_0,$$

having mean $\mu = 0$ and mass function $f\colon S_{N,M} \to [0, 1]$ satisfying $f(-N), f(M) > 0$ and

$$f(i-1) + f(i+1) \le 2f(i), \qquad -N+1 \le i \le M-1. \tag{4.1}$$

Then, F is n-CM for any $n \ge 3$.

In order to prove Theorem 4.1, we need the following lemma.

Lemma 4.1. *Under the assumptions of Theorem 4.1, we have*

$$M \le 2N \quad \textit{and} \quad N \le 2M.$$

Proof. We need to only prove that $M \le 2N$, as $N \le 2M$ follows by symmetry. The condition $\mu = 0$ implies that $M = 0$ if and only if $N = 0$; thus, we can assume that both M and N are positive. It is easy to see that (4.1) is equivalent to

$$A(v) \ge \frac{(w-v)A(u) + (v-u)A(w)}{w-u} \tag{4.2}$$

for all $u, v, w \in S_{N,M}$ such that $u \le v \le w$ and $u < w$. For instance, the two inequalities

$$f(v) \ge \frac{f(v-1)+f(v+1)}{2} \quad \text{and} \quad f(v-1) \ge \frac{f(v-2)+f(v)}{2}$$

imply that

$$f(v) \ge \frac{f(v-2)+2f(v+1)}{3}.$$

As particular cases of (4.2), we obtain

$$f(i) \ge \frac{(M-i)f(0)+if(M)}{M} > \frac{M-i}{M}f(0), \qquad 0 \le i \le M, \tag{4.3a}$$

$$f(0) \ge \frac{Mf(-j)+jf(M)}{M+j} > \frac{M}{M+j}f(-j), \qquad 0 \le j \le N. \tag{4.3b}$$

Since $\mu = \sum_{i \in S_{N,M}} if(i) = 0$, (4.3) implies that

$$\begin{aligned}
\frac{f(0)M(M-1)(M+1)}{6M} &= \frac{f(0)}{M}\sum_{i=1}^{M} i(M-i) \\
&< \sum_{i=1}^{M} if(i) \\
&= \sum_{j=0}^{N} jf(-j) \\
&< \frac{f(0)}{M}\sum_{j=1}^{N} j(M+j) \\
&= \frac{f(0)N(N+1)(3M+2N+1)}{6M},
\end{aligned}$$

from which we have

$$M(M+1)(M-1) < N(N+1)(3M+2N+1).$$

In the above equation, the right-hand side is increasing in N and equality holds when $N = (M+1)/2$. Therefore, we have $N > (M-1)/2$, namely, $M \le 2N$.

Proof of Theorem 4.1. We will prove the theorem by induction over $M+N$, the cardinality of the set $S_{N,M}$. Note that, if $M = N = 0$, F is the unit mass at 0 and is thus completely mixable for any n. Moreover, the case $M + N = 1$ is not allowed by the zero-mean condition. Therefore, the first step of the induction will be $M+N = 2$. In this case the zero-mean condition combined with (4.1) forces F to be supported on $\{-1, 0, 1\}$ with masses $f(-1) = f(1) = a$ and $f(0) = 1 - 2a$ with $a < 0 \le \frac{1}{3}$. We can write F as

$$F = (3a)G + (1-3a)H, \tag{4.4}$$

where G is the uniform distribution on $\{-1, 0, 1\}$ and H is the unit mass at 0. Being a unit mass, H is n-CM for any $n \in \mathbb{N}$, while G satisfies the assumptions of Lemma 2.8 of

Wang and Wang (2011) with $d = n - 1$ and is thus n-CM for any $n \geq 2$. Equation (4.4) states that F is the convex sum of two n-CM distributions with center $\mu = 0$. By Theorem 3.2(a), F is n-CM for any $n \geq 2$.

We now assume that the theorem holds for any distribution H satisfying the assumption of the theorem with $N + M \leq K - 1$ points in $S_{N,M}$ and prove that it holds for any distribution F with K points in $S_{N,M}$, $K \geq 3$. As illustrated for $N + M = 2$, the idea of the proof is to decompose F as the convex sum of such an H and another n-CM distribution G.

Let F be a distribution satisfying the assumption of the theorem with $N + M = K$, $K \geq 3$. Without loss of generality, in what follows we assume that $M \geq N$ (the theorem holds symmetrically for $M \leq N$). We denote by G the discrete distribution having mass function $g\colon S_{N,M} \to [0, 1]$ given by

$$g(-N) = \frac{M - N + 1}{M + N + 1}, \qquad g(-N + 1) = \cdots = g(M) = \frac{2N}{(M + N + 1)(M + N)}.$$

Elementary calculations show that the distribution G has first moment $\mu = 0$ and, since $M \geq N$, that g is decreasing. From Lemma 4.1, we have $M \leq 2N \leq (n - 1)N$ for any $n \geq 3$ and, so, the distribution G satisfies the assumption of Lemma 2.8 of Wang and Wang (2011) with $d = n - 1$. As a consequence, G is n-CM. Now, we define the function $\hat{f}\colon S_{N,M} \to \mathbb{R}$ as

$$\hat{f} = f - k_1 g, \tag{4.5}$$

where

$$k_1 = \min\left\{\frac{f(-N)}{g(-N)}, \frac{f(M)}{g(M)}\right\} > 0.$$

Note that we have

$$\hat{f}(-N) = f(-N) - k_1 g(-N) \geq f(-N) - \frac{f(-N)}{g(-N)} g(-N) = 0, \tag{4.6a}$$

$$\hat{f}(M) = f(M) - k_1 g(M) \geq f(M) - \frac{f(M)}{g(M)} g(M) = 0. \tag{4.6b}$$

Since g is convex on $S_{N,M}$, the function $\hat{f}$ is the sum of two concave densities and, therefore, is concave. The concavity of $\hat{f}$, combined with (4.6), implies that $\hat{f}$ is also nonnegative on $S_{N,M}$. At this point, it is possible to define the discrete distribution H as the distribution having concave mass function

$$h = \frac{\hat{f}}{k_2}, \tag{4.7}$$

where

$$k_2 = \sum_{i \in S_{N,M}} \hat{f}(i).$$

Note that the distribution H has mean $\mu = 0$ as

$$\sum_{i=-N}^{M} i h(i) = \frac{1}{k_2}\left(\sum_{i=-N}^{M} i f(i) - k_1 \sum_{i=-N}^{M} i g(i)\right) = 0.$$

Moreover, at least one of the values $\hat{f}(-N)$ and $\hat{f}(M)$ is equal to 0. In conclusion, H is a distribution function on a subset of $S_{N,M}$ containing at most $K - 1$ points, having mean $\mu = 0$

and concave mass function h. By the induction assumption, H is n-CM. Combining (4.5) and (4.7), we obtain

$$F = k_1 G + k_2 H \quad \text{with } k_1 + k_2 = 1.$$

Thus, F is the convex combination of two n-CM distributions and, so, F is n-CM.

Theorem 4.2. *Any continuous distribution on a bounded interval* (a, b) *having a concave density is* n*-CM for any* $n \geq 3$.

Proof. The proof is analogous to the part of the proof of Theorem 2.4 of Wang and Wang (2011) following Lemma 2.8. For any F with a concave density, we find a sequence of discrete concave distributions that goes to F. Note that a distribution with concave density on (0, 1) is n-CM for all $n \geq 3$; hence, the mean condition

$$\frac{1}{n} \leq \mu \leq 1 - \frac{1}{n}$$

is automatically satisfied for $n \geq 3$.

According to Theorem 4.2, The Beta(α, β) distribution with parameters $1 \leq \alpha, \beta \leq 2$ is n-CM for $n \geq 3$. Any triangular distribution has a concave density and, is hence n-CM for $n \geq 3$.

5. Radially symmetric distributions

In this section we show that any n-radially symmetric distribution is completely mixable. The following definition of an n-radially symmetric distribution is an extension of the definition introduced in Knott and Smith (2006).

Definition 5.1. Suppose that U is a random variable uniformly distributed on (0, 1), and let $\boldsymbol{A} = (A_1, \dots, A_n)$ and $\boldsymbol{B} = (B_1, \dots, B_n)$ be two random vectors on $\mathbb{R}^n$ independently distributed from U. A random variable X and its distribution are called n*-radially symmetric* if

$$X = a + \sum_{k=1}^{n} (A_k \cos(2\pi k U) + B_k \sin(2\pi k U)) \tag{5.1}$$

for some constant $a \in \mathbb{R}$.

In Definition 5.1, the random vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ can be chosen to have an arbitrary distribution on $\mathbb{R}^n$.

Theorem 5.1. *Any* n*-radially symmetric distribution is* m*-CM for any* $m \geq n + 1$.

Proof. Let F be the n-radially symmetric distribution of a random variable X of the form (5.1) for some U uniformly distributed on (0, 1) and $\boldsymbol{A}$ and $\boldsymbol{B}$ distributed independently from U. Fix an integer $m \geq n + 1$, and let the m random variables $X_1, \dots, X_m$ be defined as

$$X_i = a + \sum_{k=1}^{n} \bigg(A_k \cos\bigg(2\pi k \bigg(V + \frac{i}{m}\bigg)\bigg) + B_k \sin\bigg(2\pi k \bigg(V + \frac{i}{m}\bigg)\bigg)\bigg), \qquad 1 \leq i \leq m,$$

where V is a random variable uniformly distributed on (0, 1) and independent from $\boldsymbol{A}$ and $\boldsymbol{B}$. Note that

$$\cos\bigg(2\pi k \bigg(V + \frac{i}{m}\bigg)\bigg) \sim \cos(2\pi k U)$$

https://doi.org/10.1239/jap/1339878796 Published online by Cambridge University Press

and

$$\sin\left(2\pi k\left(V+\frac{i}{m}\right)\right)\sim\sin(2\pi kU)$$

for $1\le i\le m$ and $1\le k\le n$. Therefore, the X_is are identically distributed as F. To complete the proof, we show that their sum is, almost surely, the constant ma.

For $1\le i\le m$, let $\xi_i=\mathrm{e}^{\mathrm{i}2\pi ki/m}$, where i is the imaginary unit. We denote by $d_k=\gcd(k,m)$ the greatest common divisor of k and m. Since $m\ge n+1$, we have $k\le n\le m-1$ and, thus, $d_k<m$ for $1\le k\le n$. When $d_k=1$, the m values $\xi_1,\dots,\xi_m$ are the roots of the equation $\xi^m=1$ and, therefore, $\sum_{i=1}^m\xi_i=0$. If, instead, $1<d_k<m$, then the m/d_k values $\xi_1,\dots,\xi_{m/d_k}$ are the roots of the equation $\xi^{m/d_k}=1$ and, again, we have

$$\sum_{i=1}^{m}\xi_i=d_k\sum_{i=1}^{m/d_k}\xi_i=0.$$

From this, it easily follows that

$$\begin{aligned}\sum_{i=1}^{m}\left(\cos\left(2\pi k\left(V+\frac{i}{m}\right)\right)+\mathrm{i}\sin\left(2\pi k\left(V+\frac{i}{m}\right)\right)\right)&=\sum_{i=1}^{m}\mathrm{e}^{\mathrm{i}2\pi k(V+i/m)}\\&=\mathrm{e}^{\mathrm{i}2\pi kV}\sum_{i=1}^{m}\xi_i\\&=0.\end{aligned}$$

The above equality implies that

$$\sum_{i=1}^{k}\cos\left(2\pi k\left(V+\frac{i}{m}\right)\right)=\sum_{i=1}^{k}\sin\left(2\pi k\left(V+\frac{i}{m}\right)\right)=0,$$

and, therefore, that

$$\begin{aligned}\sum_{i=1}^{m}X_i&=ma+\sum_{i=1}^{m}\sum_{k=1}^{n}\left(A_k\cos\left(2\pi k\left(V+\frac{i}{m}\right)\right)+B_k\sin\left(2\pi k\left(V+\frac{i}{m}\right)\right)\right)\\&=ma+\sum_{k=1}^{n}\left(A_k\sum_{i=1}^{m}\cos\left(2\pi k\left(V+\frac{i}{m}\right)\right)+B_k\sum_{i=1}^{m}\sin\left(2\pi k\left(V+\frac{i}{m}\right)\right)\right)\\&=ma.\end{aligned}$$

An interesting example of a radially symmetric distribution is given by the continuous random variable $X=\cos(2\pi U)$, where U is uniformly distributed on $(0,1)$. By Theorem 5.1, the distribution of X is n-CM for $n\ge 2$. As illustrated in Figure 1, the density of X is a convex function on the interval $[-1,1]$. Therefore, Theorem 5.1 indicates that there exist continuous n-CM distributions with a large density at both endpoints of their support. This result is new if compared with Theorem 2.1 and Theorem 2.2, where complete mixability is stated for general classes of monotone or unimodal symmetric densities. As the set of n-CM distributions with a given center is convex, Theorem 5.1 is no doubt useful for constructing new classes of completely mixable distributions.

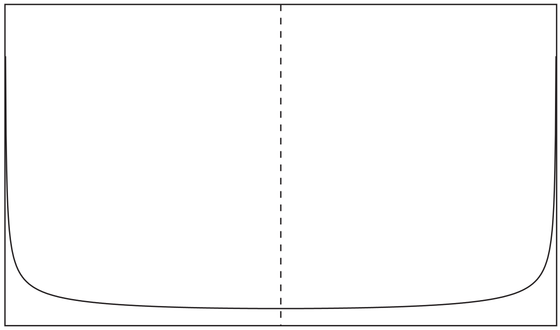

FIGURE 1: The density of the random variable $X = \cos(2\pi U)$.

6. Final remarks and open problems

In this paper we have stated three main results concerning complete mixability. First, a distribution function is n-completely mixable (n-CM) if and only if it is the limit of a sequence of countable convex combinations of n-discrete uniform distributions with the same center; see Corollary 3.1. Then, in Theorem 4.2, we stated that a continuous distribution function with a concave density is n-CM. Finally, in Theorem 5.1, we showed that radially symmetric distributions are n-CM.

In view of the relevant applications to quantitative risk management illustrated in Section 1, we believe that the above results would be useful in proving, for instance, the complete mixability of unimodal asymmetric distributions. As all the conditions implying the n-complete mixability of distributions becomes less strict when the dimension n increases, we also conjecture that any distribution F on a finite interval is n-CM for large enough n. Finally, we remark that the uniqueness of the center of n-CM distributions with infinite mean is still unproven.

Acknowledgements

The authors are grateful to Jingping Yang and the reviewers for several valuable comments on the paper.

References

EMBRECHTS, P. AND PUCCETTI, G. (2010). Risk aggregation. In *Copula Theory and Its Applications* (Lecture Notes Statist. **198**), eds P. Bickel *et al.* Springer, Berlin, pp. 111–126.

KNOTT, M. AND SMITH, C. (2006). Choosing joint distributions so that the variance of the sum is small. *J. Multivariate Anal.* **97,** 1757–1765.

NELSEN, R. B. AND ÚBEDA-FLORES, M. (2012). Directional dependence in multivariate distributions. *Ann. Inst. Statist. Math.* **64,** 677–685.

RÜSCHENDORF, L. AND UCKELMANN, L. (2002). Variance minimization and random variables with constant sum. In *Distributions with Given Marginals and Statistical Modelling*, Dordrecht, Kluwer, pp. 211–222.

WANG, B. AND WANG, R. (2011). The complete mixability and convex minimization problems with monotone marginal densities. *J. Multivariate Anal.* **102,** 1344–1360.

WANG, R., PENG, L. AND YANG, J. (2011). Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. Preprint, Georgia Institute of Technology.