

# Centers of probability measures without the mean

Giovanni Puccetti\*, Pietro Rigo<sup>†</sup>, Bin Wang<sup>‡</sup>, Ruodu Wang<sup>§</sup>

January 11, 2018

## Abstract

In the recent years, the notion of mixability has been developed with applications to operations research, optimal transportation, and quantitative finance. An  $n$ -tuple of distributions is said to be jointly mixable if there exist  $n$  random variables following these distributions and adding up to a constant, called center, with probability one. When the  $n$  distributions are identical, we speak of complete mixability. If each distribution has finite mean, the center is obviously the sum of the means. In this paper, we investigate the set of centers of completely and jointly mixable distributions not having a finite mean. In addition to several results, we show the (possibly counterintuitive) fact that, for each  $n \geq 2$ , there exist  $n$  standard Cauchy random variables adding up to a constant  $C$  if and only if

$$|C| \leq \frac{n \log(n-1)}{\pi}.$$

**MSC2000 subject classification:** Primary 60E05.

**Keywords:** Cauchy distribution; Complete mixability; Joint mixability; Multivariate dependence

## 1 Introduction

In the recent years, the field of complete and joint mixability has been rapidly developing starting from the paper Wang and Wang [23]. Mixability serves as a building block for the solutions of many optimization problems under marginal-distributional constraints. Applications are found in operations research (Haus [12], Bellini et al. [1], Boudt et al. [4]), as well as in optimal transportation (Rüschendorf [20]) and quantitative finance (Embrechts et al. [9], Bernard et al. [3]).

One of the open theoretical challenges in this field is to identify the set of all possible *centers* of an  $n$ -tuple of distributions  $(\mu_1, \dots, \mu_n)$ . More precisely, let  $\Gamma(\mu_1, \dots, \mu_n)$  be the collection of probability measures on  $\mathbb{R}^n$  having one-dimensional marginals  $\mu_1, \dots, \mu_n$ . A center of  $(\mu_1, \dots, \mu_n)$  is any constant  $C$  such that, for some  $\lambda \in \Gamma(\mu_1, \dots, \mu_n)$ ,

$$\lambda(\{x \in \mathbb{R}^n : x_1 + \dots + x_n = C\}) = 1. \quad (1)$$

---

\*Department of Economics, Management and Quantitative Methods, University of Milano, Italy.

<sup>†</sup>Department of Mathematics, University of Pavia, Italy.

<sup>‡</sup>Key Laboratory of Random Complex Structures and Data Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China.

<sup>§</sup>Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Canada.

If condition (1) holds for some  $C \in \mathbb{R}$  and some  $\lambda \in \Gamma(\mu_1, \dots, \mu_n)$ , the  $n$ -tuple  $(\mu_1, \dots, \mu_n)$  is said to be *jointly mixable* (JM, Wang et al. [26]). Similarly, a probability measure  $\mu$  on  $\mathbb{R}$  is  *$n$ -completely mixable* ( $n$ -CM, Wang and Wang [23]) if the  $n$ -tuple  $(\mu, \dots, \mu)$  is JM; in this case,  $C/n$  is called an  $n$ -center of  $\mu$ . Clearly,  $C/n$  coincides with the mean of  $\mu$  provided the latter exists and is finite. In this sense, the notion of  $n$ -center can be seen as a generalization of the notion of mean. The term  $n$ -center is used to stress the dependence on  $n$ . However, the “ $n$ -” notation will be dropped when clear from the context.

The historical motivation for investigating mixability was to minimize  $\text{var}(\sum_{i=1}^n X_i)$ , where  $X_1, \dots, X_n$  are real random variables with given marginal distributions. In fact, the idea of building random variables with constant sum, or at least whose sum has minimum variance, goes back to Gaffke and Rüschemdorf [10], where complete mixability of the uniform distribution was shown. Random sums with minimal variance were further investigated in Rüschemdorf and Uckelmann [21], where complete mixability of symmetric unimodal distributions was established. Complete mixability and joint mixability of distributions with monotone densities are characterized in Wang and Wang [23] and Wang and Wang [24], respectively.

From the analytical viewpoint, mixability can be seen as an extension of the concept of countermonotonicity to dimensions  $n \geq 3$ . For  $n = 2$ , [strongest possible negative dependence](#) can be naturally described by the notion of countermonotonicity. Two real random variables  $X_1$  and  $X_2$  are said to be countermonotonic if  $X_1$  (resp.  $X_2$ ) is a.s. an increasing (resp. decreasing) function of some common random factor  $Z$ . But, for  $n \geq 3$ , there is no clear manner to define a satisfactory notion of [strongest possible negative dependence, i.e. multidimensional countermonotonicity](#) (see Puccetti and Wang [17]). Joint mixability just provides one possible way to do that.

[Multidimensional countermonotonicity among random variables could for instance also be described as a dependence that yields a sum that is minimum in the sense of convex order](#) (see Bernard et al. [2, 3]) or as a dependence that yields, after suitable increasing transformations of the variables, a sum that is constant (Lee and Ahn [13]).

[In quantitative finance, the study of mixability has played a fundamental role in the optimization of quantile-based risk measures under marginal constraints. For example, the maximum quantile for the sum of  \$n \geq 2\$  random variables with given marginal distributions is attained by a joint distribution in the case it shows \(in part\) a jointly mixable behavior; see Embrechts et al. \[9\] and Rüschemdorf \[20\]. The paper Bernard et al. \[3\] strengthens this result by showing it also holds when the joint distribution is such that a sum is obtained that is \(in part\) minimum in the sense of convex order.](#) The idea of minimizing variance also underpins some meaningful procedures in operations research, such as the algorithms in Graham [11], Coffman and Yannakakis [6], and Puccetti and Rüschemdorf [15].

In this paper, we study the set of centers of completely and jointly mixable distributions. It directly follows from the definition that a distribution with *finite mean* has either no center (i.e. it is not mixable) or a unique center (its mean). We find that a distribution with *infinite mean* (see Section 1.1) cannot be mixable (see Corollary 3.7). Moreover, we show that a distribution with *no mean* can have many centers. If it is discrete, the centers can be finitely many (see Proposition 3.2); if it is diffuse, the centers could possibly be infinitely many (see Example 2.5).

One popular example of a probability measure with no mean is the Cauchy distribution, namely, the law of the ratio of two independent standard normal random variables. The family of Cauchy distributions is stable and closed under convolution. Furthermore, the Cauchy has several applications, mainly in physical problems. For instance, it arises as the distribution the  $x$ -intercept of a ray issuing from a uniformly distributed angle, or else as the solution to the differential equation describing forced resonance.

Our main result (Theorem 4.2) is that, for every  $n \geq 2$ , the set of  $n$ -centers of the standard Cauchy distribution is the interval

$$\left[ -\frac{\log(n-1)}{\pi}, \frac{\log(n-1)}{\pi} \right].$$

Even if apparently innocuous, the proof (or at least our proof) of such a result is quite involved. In general, the set of  $n$ -centers of a probability measure is compact (Proposition 3.1) and is finite in case the probability measure is discrete (Proposition 3.2). Moreover, we derive general bounds on the set of centers for jointly mixable distributions (Proposition 3.3).

In addition to the results mentioned above, various other useful facts are proved. Amongst them, we mention Example 2.6, which provides the first (to our knowledge) explicit construction of two joint (complete) mixes having the same marginal distributions and different centers.

We finally remark that, still today, the Cauchy distribution continues to exhibit some rather unexpected properties; see Pillai and Meng [14] for another example.

## 1.1 Notation

Throughout this paper,  $n$  is a positive integer. For any  $A \subset \mathbb{R}^n$ , we say "a probability measure on  $A$ " to mean "a probability measure on the Borel  $\sigma$ -field of  $A$ ". We write  $X \sim \nu$  to mean that  $\nu$  is the probability distribution of the random variable  $X$  and  $X \sim Y$  to mean that  $X$  and  $Y$  have the same law. Also, for any set  $Z$ ,  $\delta_z$  stands for the point mass at  $z \in Z$ .

We always denote by  $\mathcal{B}$  the Borel  $\sigma$ -field of  $\mathbb{R}$  and by  $\mu$  (with or without indices) a probability measure on  $\mathcal{B}$ . Recall that  $\mu$  has a (possibly infinite) mean if  $\int x^+ \mu(dx) \wedge \int x^- \mu(dx) < \infty$ , in which case  $\int x^+ \mu(dx) - \int x^- \mu(dx)$  is the mean of  $\mu$ . In particular,  $\mu$  has a finite mean if  $\int |x| \mu(dx) < \infty$  while  $\mu$  does not have the mean if  $\int x^+ \mu(dx) = \int x^- \mu(dx) = \infty$ .

For  $x \in \mathbb{R}^n$  and  $i = 1, \dots, n$ , the  $i$ -th coordinate of  $x$  is denoted by  $x_i$ . If  $\lambda$  is a probability measure on  $\mathbb{R}^n$ , the  $i$ -th one-dimensional marginal of  $\lambda$  is the probability measure on  $\mathbb{R}$  given by  $A \mapsto \lambda(\{x \in \mathbb{R}^n : x_i \in A\})$ .

Finally, all random variables are defined on a common probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ .

## 2 (Non)-Uniqueness of the center

A *joint mix* for  $(\mu_1, \dots, \mu_n)$  with center  $C$  is an  $n$ -tuple  $(X_1, \dots, X_n)$  of real random variables such that  $X_i \sim \mu_i$ ,  $1 \leq i \leq n$ , and  $\sum_{i=1}^n X_i \stackrel{\text{a.s.}}{=} C$ . Similarly,  $(X_1, \dots, X_n)$  is a  *$n$ -complete mix* for  $\mu$  with center  $c$  if  $X_i \sim \mu_i$ ,  $1 \leq i \leq n$ , and  $\sum_{i=1}^n X_i \stackrel{\text{a.s.}}{=} nc$ .

Not all probability measures on  $\mathbb{R}$  are completely mixable. For instance,  $\mu$  is necessarily symmetric if it is 2-CM and centered at 0. Or else,  $\mu$  is not  $n$ -CM for any  $n$  if the support

of  $\mu$  is bounded above (below) but not below (above). For a broad list of jointly and completely mixable distributions we refer to Puccetti and Wang [17] and the references therein. Here, we start by noting that the existence of a joint mix always delivers the existence of a complete mix (with average marginal distribution) and that a complete mix with a given marginal can always be taken to be exchangeable.

**Proposition 2.1.** (i) *If  $(\mu_1, \dots, \mu_n)$  is JM with center  $C$ , then  $\mu = (\mu_1 + \dots + \mu_n)/n$  is  $n$ -CM with center  $C/n$ .*

(ii) *Each  $n$ -CM probability measure on  $\mathbb{R}$  admits an exchangeable  $n$ -complete mix.*

*Proof.* (i) Let  $X = (X_1, \dots, X_n)$  be a joint mix for  $(\mu_1, \dots, \mu_n)$  with center  $C$ . Define

$$Z = (X_{\pi_1}, \dots, X_{\pi_n})$$

where  $\pi = (\pi_1, \dots, \pi_n)$  is a uniform random permutation of  $\{1, \dots, n\}$  independent of  $X$ . A uniform random permutation of  $\{1, \dots, n\}$  is a random permutation  $\pi$  such that  $\mathbb{P}(\pi = \sigma) = 1/n!$  for each permutation  $\sigma$  of  $\{1, \dots, n\}$ . Then,  $\sum_{i=1}^n Z_i = \sum_{i=1}^n X_{\pi_i} = \sum_{i=1}^n X_i \stackrel{\text{a.s.}}{=} C$ . By independence of  $X$  and  $\pi$  and recalling that  $X_i \sim \mu_i$ , one obtains

$$\mathbb{P}(Z_i \in A) = \sum_{\sigma \in P_n} \mathbb{P}(X_{\sigma_i} \in A, \pi = \sigma) = \frac{1}{n!} \sum_{\sigma \in P_n} \mathbb{P}(X_{\sigma_i} \in A) = \frac{1}{n} \sum_{i=1}^n \mathbb{P}(X_i \in A) = \mu(A),$$

for each  $i$  and  $A \in \mathcal{B}$ , where  $P_n$  is the set of all permutations of  $\{1, \dots, n\}$ . Therefore,  $\mu$  is  $n$ -CM with center  $C/n$  and  $Z$  is a joint mix for  $\mu$ .

(ii) Given an  $n$ -CM probability  $\nu$  on  $\mathbb{R}$ , take  $\mu_1 = \dots = \mu_n = \nu$  in (i). Then,  $Z$  is an exchangeable joint mix for  $\nu$ .  $\square$

The next example, even if obvious, is helpful in proving Theorem 4.2 below.

**Example 2.2.** Let  $\nu$  and  $\gamma$  be probability measures on  $\mathbb{R}$ . Suppose  $\nu$  is  $k$ -CM and  $\gamma$  is  $(n-k)$ -CM, where  $1 \leq k < n$ . Define  $\mu_i = \nu$  for  $1 \leq i \leq k$  and  $\mu_i = \gamma$  for  $k < i \leq n$ . Then,  $(\mu_1, \dots, \mu_n)$  is clearly JM, so that

$$\frac{k\nu + (n-k)\gamma}{n}$$

is  $n$ -CM by Proposition 2.1. In particular,  $\frac{k\delta_x + (n-k)\delta_y}{n}$  is  $n$ -CM for any  $x, y \in \mathbb{R}$ .

An intriguing question is whether the center of mixable distributions is unique. Obviously, if  $(\mu_1, \dots, \mu_n)$  is JM and each  $\mu_i$  has finite mean, then  $(\mu_1, \dots, \mu_n)$  has a unique center  $C$ , namely  $C = \sum_{i=1}^n \int x \mu_i(dx)$ . Analogously, if  $\mu$  is  $n$ -CM and has finite mean,  $\int x \mu(dx)$  is the only center of  $\mu$ . Uniqueness of the center is also clear for  $n = 1$ , since  $\mu$  is 1-CM if and only if it is degenerate.

In view of Simons [22], if  $X$  and  $Y$  are real random variables such that  $\mathbb{E}(X+Y)$  exists (finite or infinite) then  $\mathbb{E}(X+Y)$  depends only on the marginal distributions of  $X$  and  $Y$ , in the sense that  $\mathbb{E}(U+V) = \mathbb{E}(X+Y)$  provided  $U \sim X$ ,  $V \sim Y$  and  $\mathbb{E}(U+V)$  exists. It follows

that the center is unique for  $n = 2$ : if  $(X_1, X_2)$  and  $(Y_1, Y_2)$  are joint mixes for  $(\mu_1, \mu_2)$  with  $X_1 + X_2 \stackrel{\text{a.s.}}{=} C_1$  and  $Y_1 + Y_2 \stackrel{\text{a.s.}}{=} C_2$ , then

$$X_2 - C_1 \stackrel{\text{a.s.}}{=} -X_1 \sim -Y_1 \stackrel{\text{a.s.}}{=} Y_2 - C_2 \sim X_2 - C_2,$$

which clearly implies  $C_1 = C_2$ . More generally, one obtains the following result.

**Proposition 2.3.** *Suppose that  $(\mu_1, \dots, \mu_n)$  is JM and at least  $n - 2$  of  $\mu_1, \dots, \mu_n$  have finite mean. Then, the center of  $(\mu_1, \dots, \mu_n)$  is unique.*

*Proof.* Let  $n > 2$  and let  $(X_1, \dots, X_n)$  be a joint mix for  $(\mu_1, \dots, \mu_n)$  with center  $C$ . Without loss of generality, assume that  $\mu_1, \dots, \mu_{n-2}$  have finite mean. Then,  $\sum_{i=1}^{n-2} X_i$  is integrable and  $X_{n-1} + X_n \stackrel{\text{a.s.}}{=} C - \sum_{i=1}^{n-2} X_i$ . Thus,  $\mathbb{E}(X_{n-1} + X_n)$  is finite, so that  $\mathbb{E}(X_{n-1} + X_n)$  only depends on  $\mu_{n-1}$  and  $\mu_n$  (because of Simons [22]). Hence,

$$C = \sum_{i=1}^{n-2} \mathbb{E}(X_i) + \mathbb{E}(X_{n-1} + X_n)$$

is the only center of  $(\mu_1, \dots, \mu_n)$ . □

Another uniqueness criterion can be obtained by increasing  $n - 2$  to  $n - 1$  but replacing in the above proposition the existence of the mean with the slightly weaker condition

$$\lim_{x \rightarrow \infty} x \mu(\{y \in \mathbb{R} : |y| > x\}) = 0. \quad (2)$$

**Proposition 2.4.** *Suppose that  $(\mu_1, \dots, \mu_n)$  is JM and at least  $n - 1$  of  $\mu_1, \dots, \mu_n$  satisfy condition (2). Then, the center of  $(\mu_1, \dots, \mu_n)$  is unique.*

*Proof.* Let  $(X_1, \dots, X_n)$  be a joint mix for  $(\mu_1, \dots, \mu_n)$  with center  $C$ . Without loss of generality, assume that  $\mu_1, \dots, \mu_{n-1}$  satisfy condition (2). Then,  $\mu_n$  also satisfies condition (2). In fact, for every  $x \geq n|C|$ , one obtains

$$\mathbb{P}(|X_n| > x) = \mathbb{P}\left(|C - \sum_{i=1}^{n-1} X_i| > x\right) \leq \sum_{i=1}^{n-1} \mathbb{P}(|X_i| > x/n).$$

Hence,  $x \mathbb{P}(|X_n| > x) \rightarrow 0$  as  $x \rightarrow \infty$ .

Now take iid copies of  $(X_1, \dots, X_n)$ , denoted by  $\{(X_{1,k}, \dots, X_{n,k})\}_{k=1}^{\infty}$ . For  $m \in \mathbb{N}$ , let  $c_{i,m} = \mathbb{E}(X_i \mathbf{1}_{\{|X_i| \leq m\}})$ ,  $i = 1, \dots, n$ , and  $c_m = \sum_{i=1}^n c_{i,m}$ . By condition (2) and the weak law of large numbers, we have  $\frac{1}{m} \sum_{k=1}^m X_{i,k} - c_{i,m} \xrightarrow{\mathbb{P}} 0$ , as  $m \rightarrow \infty$ , for fixed  $i = 1, \dots, n$ . It follows that

$$C - c_m = \sum_{i=1}^n \left( \frac{1}{m} \sum_{k=1}^m X_{i,k} - c_{i,m} \right) \xrightarrow{\mathbb{P}} 0 \quad \text{as } m \rightarrow \infty.$$

Therefore,  $C = \lim_m c_m$  is unique. □

Contrary to the cases  $n = 1$  and  $n = 2$ , a JM  $n$ -tuple of distributions may have more than one center if  $n \geq 3$ .

Recall that the standard Cauchy distribution is the probability measure on  $\mathbb{R}$  with density  $f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$  with respect to the Lebesgue measure. Let  $\text{Cauchy}(\sigma)$  denote the distribution of  $\sigma X$ , where  $\sigma > 0$  and  $X$  has the standard Cauchy distribution. In Chen and Shepp [5], Chen and Shepp show the existence of two  $\text{Cauchy}(4)$  random variables  $U, V$  and a constant  $C \neq 0$  such that  $U + V + C$  is  $\text{Cauchy}(4\sqrt{2})$ . Thus,  $(-U, -V, U + V + C)$  and  $(U, V, -U - V - C)$  are both joint mixes for the triplet  $(\text{Cauchy}(4), \text{Cauchy}(4), \text{Cauchy}(4\sqrt{2}))$  with centers  $C$  and  $-C$ , respectively. From Proposition 2.1, one also obtains a 3-CM probability measure with an interval of centers.

**Example 2.5** (A probability measure with an interval of centers). Let  $\nu = \text{Cauchy}(4)$  and  $\gamma = \text{Cauchy}(4\sqrt{2})$ . By Chen and Shepp [5], the triplet  $(\nu, \nu, \gamma)$  is JM with centers  $C$  and  $-C$  for some  $C > 0$ . Take two independent joint mixes  $(X_1, X_2, X_3)$  and  $(Y_1, Y_2, Y_3)$  for  $(\nu, \nu, \gamma)$  with centers  $C$  and  $-C$ , respectively. Fix  $\alpha \in [0, 1]$  and define  $Z_i = \alpha X_i + (1 - \alpha)Y_i, 1 \leq i \leq 3$ . Using characteristic functions, it is straightforward to see that  $Z_1 \sim Z_2 \sim \nu$  and  $Z_3 \sim \gamma$ . Hence,  $(Z_1, Z_2, Z_3)$  is a joint mix for  $(\nu, \nu, \gamma)$  with center  $(2\alpha - 1)C$ . Since  $\alpha \in [0, 1]$  is arbitrary, Proposition 2.1 implies that each point in the interval  $[-C/3, C/3]$  is a 3-center of  $\mu = (2\nu + \gamma)/3$ . Note however that such  $\mu$  does not belong to the Cauchy family of distributions.

The general question of whether the center of a JM  $n$ -tuple of distributions is *always* unique was stated as an open problem in Puccetti and Wang [17] and Wang [25]. During the writing of the present paper, we became aware of the Chen-Shepp example in Chen and Shepp [5] providing an early negative answer to the question.

However, the Chen-Shepp example, while implying non-uniqueness of the center, does not provide an explicit construction for it depends on (the existence of) an orthogonal projection. Furthermore, the value of  $C$  is not explicitly given and it is not clear if and how it can be computed. We next give an example of couplings having the same marginal distributions but different sums. To the best of our knowledge, this is the first explicit construction of two joint (complete) mixes having the same marginal distributions and different centers.

**Example 2.6** (Center of a JM triplet is not unique). Let  $Z$  be a random variable with a geometric distribution with parameter  $1/2$ , that is,  $\mathbb{P}(Z = k) = 2^{-(k+1)}, k \geq 0$ , and let  $B$  be a Bernoulli random variable with parameter  $1/2$  independent of  $Z$ . Let

$$X_1 = X_2 = 2^Z, X_3 = -2^{Z+1},$$

and

$$Y_1 = B2^{Z+1} + (1 - B), Y_2 = (1 - B)2^{Z+1} + B, Y_3 = -2^{Z+1}.$$

Then,  $X_1 + X_2 + X_3 = 0, Y_1 + Y_2 + Y_3 = 1$  and  $X_3 = Y_3$ . Furthermore,  $\mathbb{P}(X_1 = 1) = 1/2 = \mathbb{P}(Y_1 = 1)$  and

$$\mathbb{P}(Y_1 = 2^k) = \mathbb{P}(B = 1)\mathbb{P}(Z = k - 1) = 2^{-(k+1)} = \mathbb{P}(X_1 = 2^k)$$

for each  $k \geq 1$ . Similarly,  $X_2 \sim Y_2$ . Thus, if  $\nu$  denotes the distribution of  $X_1 = X_2$  and  $\gamma$  that of  $X_3$ , the triplet  $(\nu, \nu, \gamma)$  is JM with centers 0 and 1. From this example and Proposition 2.1,

it also follows that the probability measure  $\mu = \frac{2}{3}\nu + \frac{1}{3}\gamma$  is 3-CM with centers 0 and 1/3. In Example 3.6 below we shall see that 0 and 1/3 are actually the only 3-centers of  $\mu$ .

### 3 The set of centers of mixable distributions

Let  $\Lambda_n(\mu)$  be the set of those  $\lambda \in \Gamma(\mu, \dots, \mu)$  such that

$$\lambda(\{x \in \mathbb{R}^n : x_1 + \dots + x_n = C\}) = 1$$

for some  $C \in \mathbb{R}$ . With a slight abuse of terminology, we call  $C/n$  the  $n$ -center of  $\lambda$ . Clearly,  $\Lambda_n(\mu) \neq \emptyset$  if and only if  $\mu$  is  $n$ -CM and each  $\lambda \in \Lambda_n(\mu)$  is the distribution of a  $n$ -complete mix for  $\mu$ . We also denote by  $\phi$  the function

$$\phi(\lambda) = \text{center}(\lambda), \quad \text{for } \lambda \in \Lambda_n(\mu).$$

**Proposition 3.1.** *Let  $\mu$  be  $n$ -CM. Then,  $\Lambda_n(\mu)$  is a compact metric space and the function  $\phi : \Lambda_n(\mu) \rightarrow \mathbb{R}$  is continuous. In particular, the set  $\phi(\Lambda_n(\mu))$  of  $n$ -centers of  $\mu$  is compact. Thus, there exist  $a \leq b$  such that  $a$  and  $b$  are  $n$ -centers of  $\mu$  but no point in  $(-\infty, a) \cup (b, +\infty)$  is an  $n$ -center of  $\mu$ .*

*Proof.* Let  $\mathcal{P}_n$  be the set of all probability measures on  $\mathbb{R}^n$ , equipped with the topology of weak convergence, and let  $S(x) = \sum_{i=1}^n x_i$ , for  $x \in \mathbb{R}^n$ . Since  $\Lambda_n(\mu) \subset \mathcal{P}_n$  and  $\mathcal{P}_n$  is a Polish space,  $\Lambda_n(\mu)$  is compact if and only if it is closed and tight. Fix  $\lambda_k \in \Lambda_n(\mu)$  and  $\lambda \in \mathcal{P}_n$  such that  $\lambda_k \rightarrow \lambda$  weakly as  $k \rightarrow \infty$ . Then,  $\lambda \in \Gamma(\mu, \dots, \mu)$  since  $\lambda_k \in \Gamma(\mu, \dots, \mu)$  for all  $k$  and the coordinate maps  $x \mapsto x_i$  are continuous for all  $i = 1, \dots, n$ . Similarly, since  $S$  is continuous,  $\lambda_k \circ S^{-1} \rightarrow \lambda \circ S^{-1}$  weakly. In addition,  $\lambda_k \in \Lambda_n(\mu)$  implies  $\lambda_k \circ S^{-1} = \delta_{n\phi(\lambda_k)}$  for all  $k$ . Since  $\{\delta_x : x \in \mathbb{R}^n\}$  is a closed subset of  $\mathcal{P}_n$ , it follows that  $\lambda \circ S^{-1} = \delta_x$  for some  $x \in \mathbb{R}^n$  and  $n\phi(\lambda_k) \rightarrow x$  as  $k \rightarrow \infty$ . Hence,  $\lambda \in \Lambda_n(\mu)$  and  $\phi(\lambda) = x/n = \lim_k \phi(\lambda_k)$ . This proves that  $\Lambda_n(\mu)$  is closed and  $\phi$  is continuous. Finally,  $\Lambda_n(\mu)$  is tight since all its elements have the same one-dimensional marginals (all equal to  $\mu$ ). Thus,  $\Lambda_n(\mu)$  is a compact metric space.  $\square$

If  $\mu$  is  $n$ -CM with an unique center, then  $a = b$  in Proposition 3.1. In case  $a < b$ , a natural question is: which points in the interval  $[a, b]$  are centers of  $\mu$ ?

A probability measure supported by the integers, such as  $\mu$  in Example 2.6, has at most finitely many centers because of Proposition 3.1. This conclusion can be actually generalized to any discrete distribution.

**Proposition 3.2.** *If  $(\mu_1, \dots, \mu_n)$  is JM and each  $\mu_i$  is discrete, then  $(\mu_1, \dots, \mu_n)$  has finitely many centers.*

*Proof.* Since  $\mu_i$  is discrete, there is a finite set  $A_i$  such that  $0 \in A_i$  and  $\mu_i(A_i) > 1 - 1/n$ . Let  $(X_1, \dots, X_n)$  be a joint mix for  $(\mu_1, \dots, \mu_n)$  with center  $C$  and  $B = \bigcap_{i=1}^n \{X_i \in A_i\}$ . Then,

$$C\mathbf{1}_B \stackrel{\text{a.s.}}{=} (X_1 + \dots + X_n)\mathbf{1}_B \in A_1 + \dots + A_n,$$

where  $A_1 + \dots + A_n = \{\sum_{i=1}^n x_i : x_i \in A_i, 1 \leq i \leq n\}$ . Since  $\mathbb{P}(B) > 0$ , it follows that  $C$  belongs to the finite set  $A_1 + \dots + A_n$ .  $\square$

The situation is quite different for diffuse distributions, which can have infinitely many centers; see for instance Example 2.5. A more interesting case is exhibited by Theorem 4.2 below, where  $\mu$  is the standard Cauchy and each point in  $[a, b]$  is an  $n$ -center of  $\mu$ .

We next obtain two useful bounds for  $a$  and  $b$  in Proposition 3.1. Define the quantile functional

$$q_\mu(t) = \inf\{x \in \mathbb{R} : \mu((-\infty, x]) \geq t\}, \quad t \in (0, 1).$$

For  $0 < \alpha < \beta < 1$ , define also the average quantile functional

$$R_{[\alpha, \beta]}(\mu) = \frac{1}{\beta - \alpha} \int_\alpha^\beta q_\mu(t) dt.$$

Note that, for fixed  $\alpha$  and  $\beta$ , the map  $\mu \mapsto R_{[\alpha, \beta]}(\mu)$  is continuous with respect to weak convergence.

**Proposition 3.3.** *Suppose that  $(\mu_1, \dots, \mu_n)$  is JM with center  $C$ . Then, for any  $\beta_1, \dots, \beta_n \in (0, 1)$  such that  $\beta := \beta_1 + \dots + \beta_n < 1$ , one obtains*

$$\sum_{i=1}^n R_{[\beta_i, 1-\beta+\beta_i]}(\mu_i) \leq C \leq \sum_{i=1}^n R_{[\beta-\beta_i, 1-\beta_i]}(\mu_i). \quad (3)$$

*Proof.* The first inequality follows from the second by noting that  $-C$  is a center of  $(\mu_1^*, \dots, \mu_n^*)$ , where  $\mu_i^*(A) = \mu_i(-A)$  for each  $A \in \mathcal{B}$ . Hence, we only prove the second inequality.

By applying  $\alpha_i = \beta_i$ ,  $\beta_i = 1 - \beta - \epsilon$  in Theorem 1 of Embrechts et al. [7] (with their notation  $\text{RVaR}_{a,b}(X) = R_{[1-a-b, 1-a]}(\mu)$  for  $X \sim \mu$ ,  $a, b > 0$ ,  $a + b < 1$ ), we obtain, for any integrable random variables  $Y_1 \sim \nu_1, \dots, Y_n \sim \nu_n$ , with  $\sum_{i=1}^n Y_i = Y \sim \nu$ , that

$$R_{[\epsilon, 1-\beta]}(\nu) = \text{RVaR}_{\beta, 1-\beta-\epsilon}(Y) \leq \sum_{i=1}^n \text{RVaR}_{\beta_i, 1-\beta-\epsilon}(Y_i) = \sum_{i=1}^n R_{[\beta-\beta_i+\epsilon, 1-\beta_i]}(\nu_i),$$

for all  $\epsilon \in (0, 1 - \beta)$ .

Take a joint mix  $(X_1, \dots, X_n)$  for  $(\mu_1, \dots, \mu_n)$  with center  $C$  and a sequence  $\{(X_{1,k}, \dots, X_{n,k})\}_{k=1}^\infty$  satisfying

$$\begin{aligned} &X_{i,k} \text{ is integrable for all } i, k \text{ and} \\ &(X_{1,k}, \dots, X_{n,k}) \xrightarrow{d} (X_1, \dots, X_n) \quad \text{as } k \rightarrow \infty \end{aligned}$$

where  $\xrightarrow{d}$  stands for convergence in distribution. Then,  $\sum_{i=1}^n X_{i,k} \xrightarrow{d} \sum_{i=1}^n X_i \stackrel{\text{a.s.}}{=} C$  and  $X_{i,k} \xrightarrow{d} X_i$  for all  $i$ . By continuity of the average quantile functional with respect to weak convergence, it follows that

$$C = R_{[\epsilon, 1-\beta]}(\delta_C) \leq \sum_{i=1}^n R_{[\beta-\beta_i+\epsilon, 1-\beta_i]}(\mu_i) \quad \text{for any } \epsilon \in (0, 1 - \beta).$$



Finally, by taking  $\epsilon \downarrow 0$ , one obtains

$$C \leq \sum_{i=1}^n R_{[\beta-\beta_i, 1-\beta_i]}(\mu_i),$$

which concludes the proof.  $\square$

Letting  $\beta_1 = \dots = \beta_n$  and  $\mu_1 = \dots = \mu_n$ , Proposition 3.3 has the following useful consequence.

**Corollary 3.4.** *If  $\mu$  is  $n$ -CM with center  $c$ , then  $a^* \leq c \leq b^*$ , where*

$$a^* = \sup_{\alpha \in (0, \frac{1}{n})} R_{[\alpha, 1-(n-1)\alpha]}(\mu) \quad \text{and} \quad b^* = \inf_{\alpha \in (0, \frac{1}{n})} R_{[(n-1)\alpha, 1-\alpha]}(\mu).$$

**Example 3.5** (The mean inequality). A remarkable consequence of Corollary 3.4 is the mean inequality (Proposition 2.1(7) of Wang and Wang [23]), arguably the most important necessary condition for complete mixability, which is also sufficient for probability measures with monotone densities. Let  $x = \lim_{\epsilon \downarrow 0} q_\mu(\epsilon)$  and  $y = \lim_{\epsilon \downarrow 0} q_\mu(1 - \epsilon)$  be the left and right end-points of  $\mu$ . Assume  $x$  and  $y$  are finite and denote by  $c$  the mean of  $\mu$ . If  $\mu$  is  $n$ -CM, Corollary 3.4 yields  $c \leq b^*$ . Hence, for  $\alpha \in (0, 1/n)$ , we have that  $c \leq R_{[(n-1)\alpha, 1-\alpha]}(\mu)$ , that is

$$\frac{c - (1 - n\alpha)R_{[(n-1)\alpha, 1-\alpha]}(\mu)}{\alpha} \leq nc. \quad (4)$$

On the other hand,

$$\begin{aligned} \lim_{\alpha \downarrow 0} \frac{1}{\alpha} (c - (1 - n\alpha)R_{[(n-1)\alpha, 1-\alpha]}(\mu)) &= \lim_{\alpha \downarrow 0} \frac{1}{\alpha} \left( \int_0^1 q_\mu(t) dt - \int_{(n-1)\alpha}^{1-\alpha} q_\mu(t) dt \right) \\ &= \lim_{\alpha \downarrow 0} \frac{1}{\alpha} \left( \int_{1-\alpha}^1 q_\mu(t) dt + \int_0^{(n-1)\alpha} q_\mu(t) dt \right) = y + (n-1)x. \end{aligned}$$

Therefore, inequality (4) yields  $y + (n-1)x \leq nc$ , one side of the mean inequality in Wang and Wang [23] (the other follows similarly).

**Example 3.6** (Example 2.6 revisited). The probability measure  $\mu$  defined in Example 2.6 is 3-CM with 3-centers 0 and  $1/3$ . We now prove that 0 and  $1/3$  are actually the only 3-centers of  $\mu$ . For  $n = 3$ , one can compute that

$$a^* \geq \lim_{\alpha \downarrow 0} R_{[\alpha, 1-2\alpha]}(\mu) = 0 \quad \text{and} \quad b^* \leq \lim_{\alpha \downarrow 0} R_{[2\alpha, 1-\alpha]}(\mu) = 2/3.$$

By Corollary 3.4, it follows that  $a = 0$  and  $b \leq 2/3$ . Let  $(X_1, X_2, X_3)$  be a complete mix for  $\mu$ . Since  $\mu(\mathbb{Z}) = 1$ ,  $a = 0$  and  $b \leq 2/3$ , then  $X_1 + X_2 + X_3 \in [0, 2] \cap \mathbb{Z}$  a.s. Thus, to see that 0 and  $1/3$  are the only 3-centers of  $\mu$ , it suffices to show that  $\mathbb{P}(X_1 + X_2 + X_3 = 2) < 1$ . Since  $X_2 + X_3 \neq 1$  a.s., then  $\mathbb{P}(X_1 + X_2 + X_3 = 2) \leq \mathbb{P}(X_1 \neq 1) < 1$ .

Another consequence of Proposition 3.3 is that a distribution with an infinite mean

cannot be  $n$ -CM for any  $n \in \mathbb{N}$ . The following corollary can be shown by letting  $\beta_i \downarrow 0$ ,  $i = 1, \dots, n$  in (3), so that the left-hand side of (3) goes to infinity.

**Corollary 3.7.** *If  $\mu_1, \dots, \mu_n$  have means  $m_1, \dots, m_n \in (-\infty, \infty]$ , respectively, and  $m_i = \infty$  for at least one  $i = 1, \dots, n$ , then  $(\mu_1, \dots, \mu_n)$  is not JM. In particular, a probability measure on  $\mathbb{R}$  with an infinite mean is not  $n$ -CM for any  $n \in \mathbb{N}$ .*

We conclude this section by characterizing the set of  $n$ -centers of a probability measure based on a duality argument.

Let  $\mu$  be a probability measure on  $\mathbb{R}$ . Recall that  $S(x) = \sum_{i=1}^n x_i$  for  $x \in \mathbb{R}^n$  and write  $\{S = nc\}$  to denote the set  $\{x \in \mathbb{R}^n : S(x) = nc\}$ . By definition, a real number  $c$  is an  $n$ -center of  $\mu$  if and only if  $M(c) = 1$ , where

$$M(c) = \sup\{\lambda(S = nc) : \lambda \in \Gamma(\mu, \dots, \mu)\}.$$

Based on Theorem 5 of Rüschemdorf [19] and Remark 2 in Gaffke and Rüschemdorf [10],  $M(c)$  has the dual representation

$$M(c) = n \inf \left\{ \int g \, d\mu : g \in \mathcal{D}(c) \right\}, \quad (5)$$

where  $\mathcal{D}(c)$  denotes the class of bounded, Borel-measurable functions  $g : \mathbb{R} \rightarrow \mathbb{R}$  such that  $\sum_{i=1}^n g(x_i) \geq \mathbf{1}_{\{S=nc\}}(x)$  for all  $x \in \mathbb{R}^n$ . The value of (5) is not easy to compute in general. However, restricting to a subset of  $\mathcal{D}(c)$  (as done for instance in Embrechts and Puccetti [8]) leads to an upper bound for  $M(c)$ .

We consider the following class of piecewise-linear functions defined, for  $t < c$ , as

$$g_t(x) = \begin{cases} 0 & \text{if } x < t, \\ \frac{x-t}{n(c-t)} & \text{if } t \leq x \leq nc - (n-1)t, \\ 1 & \text{otherwise.} \end{cases}$$

Since  $g_t \in \mathcal{D}(c)$  for all  $t < c$ , we obtain

$$M(c) \leq D(c) := \inf_{t < c} \left\{ n \int g_t \, d\mu \right\} = \inf_{t < c} \left\{ \frac{\int_t^{nc-(n-1)t} \bar{F}(x) \, dx}{c-t} \right\}$$

where  $\bar{F}(x) = \mu((x, \infty))$ . If  $D(c) < 1$ , then  $c$  is not an  $n$ -center of  $\mu$ . Therefore, for  $c$  to be an  $n$ -center of  $\mu$ , it is necessary that

$$\frac{\int_t^{nc-(n-1)t} \bar{F}(x) \, dx}{c-t} \geq 1 \quad \text{for all } t < c.$$

The above inequality is another necessary condition for the center of  $n$ -CM probability measures, in addition to that of Corollary 3.4. These two necessary conditions are not equivalent in general.

## 4 The Cauchy distribution

From now on, we let  $\mu = \text{Cauchy}(1)$ , the standard Cauchy distribution. It is shown in Rüschemdorf and Uckelmann [21] that  $\mu$  is  $n$ -CM with center 0, for each  $n \geq 2$ , as it is symmetric and unimodal. In this section, we characterize the set of  $n$ -centers of  $\mu$ . We start by observing that such set is a closed interval contained in  $[-\frac{\log(n-1)}{\pi}, \frac{\log(n-1)}{\pi}]$ .

**Example 4.1.** As in Proposition 3.1, let  $a$  and  $b$  be the minimum and the maximum of the set of  $n$ -centers of  $\mu$ . Let  $X = (X_1, \dots, X_n)$  and  $Y = (Y_1, \dots, Y_n)$  be two independent complete mixes for  $\mu$  such that

$$\sum_{i=1}^n X_i \stackrel{\text{a.s.}}{=} na \quad \text{and} \quad \sum_{i=1}^n Y_i \stackrel{\text{a.s.}}{=} nb.$$

Fix  $\alpha \in [0, 1]$  and define  $Z_i = \alpha X_i + (1 - \alpha)Y_i$  for  $1 \leq i \leq n$ . Then,  $Z_i \sim \mu$  for each  $i$  and  $\sum_{i=1}^n Z_i = n(\alpha a + (1 - \alpha)b)$  so that  $\alpha a + (1 - \alpha)b$  is a center of  $\mu$ . Hence,  $\phi(\Lambda_n(\mu)) = [a, b]$ , namely, each point in  $[a, b]$  is a center of  $\mu$ .

Next, on noting that  $q_\mu(t) = \tan(\pi(t - 1/2))$ , one obtains

$$\begin{aligned} R_{[(n-1)\alpha, 1-\alpha]}(\mu) &= \frac{1}{1 - n\alpha} \int_{(n-1)\alpha}^{1-\alpha} \tan(\pi(t - 1/2)) dt \\ &= \frac{1}{1 - n\alpha} \frac{1}{\pi} \log \left( \frac{\sin(\pi(n-1)\alpha)}{\sin(\pi\alpha)} \right) \quad \text{for } \alpha \in (0, 1/n). \end{aligned}$$

By Corollary 3.4,

$$b \leq b^* \leq \lim_{\alpha \downarrow 0} R_{[(n-1)\alpha, 1-\alpha]}(\mu) = \frac{\log(n-1)}{\pi}.$$

Since  $a = -b$  (for  $\mu$  is symmetric) one also obtains  $a \geq -\frac{\log(n-1)}{\pi}$ .

Example 4.1 says that  $\phi(\Lambda_n(\mu)) \subset [-\frac{\log(n-1)}{\pi}, \frac{\log(n-1)}{\pi}]$ . Our main result is that this inclusion is an equality.

**Theorem 4.2.** *For every  $n \geq 2$ , the set of  $n$ -centers of the standard Cauchy distribution is the interval*

$$\left[ -\frac{\log(n-1)}{\pi}, \frac{\log(n-1)}{\pi} \right].$$

The rest of this section is devoted to the proof of Theorem 4.2.

For each  $c \in \mathbb{R}$ , let  $\mathcal{M}_n(c)$  denote the collection of  $n$ -CM probability measures with center  $c$ . We first need two lemmas of possible independent interest. The first states that  $\mathcal{M}_n(c)$  is closed under arbitrary mixtures, generalizing Theorem 3.2 of Puccetti et al. [16].

**Lemma 4.3.** *Let  $(T, \mathcal{E}, Q)$  be any probability space and, for each  $t \in T$ , let  $\nu_t \in \mathcal{M}_n(c)$ . Suppose that  $t \mapsto \nu_t(B)$  is a  $\mathcal{E}$ -measurable map, for each  $B \in \mathcal{B}$ , and define*

$$\nu(B) = \int \nu_t(B) Q(dt).$$

Then,  $\nu \in \mathcal{M}_n(c)$ .

*Proof.* Let  $\mathcal{R}$  be the field on  $\mathbb{R}^n$  generated by the measurable rectangles  $B_1 \times \dots \times B_n$ , where  $B_i \in \mathcal{B}$  for all  $i$ , and let  $\gamma : \mathcal{R} \rightarrow \mathbb{R}$  be any map. By Theorem 6 of Ramachandran [18],  $\gamma$  is a  $\sigma$ -additive probability on  $\mathcal{R}$  provided it is a *finitely additive* probability and  $A \mapsto \gamma\{x \in \mathbb{R}^n : x_i \in A\}$  is a  $\sigma$ -additive probability on  $\mathcal{B}$  for all  $i$ .

Let  $H = \{x \in \mathbb{R}^n : x_1 + \dots + x_n = nc\}$ . For each  $t \in T$ , since  $\nu_t \in \mathcal{M}_n(c)$ , there is  $\lambda_t \in \Gamma(\nu_t, \dots, \nu_t)$  such that  $\lambda_t(H) = 1$ . Define

$$\lambda^*(B) = \int \lambda_t(B) Q^*(dt) \quad \text{for each } B \in \mathcal{R},$$

where  $Q^*$  is a finitely additive extension of  $Q$  to the power set of  $T$ . Then,  $\lambda^*$  is a finitely additive probability on  $\mathcal{R}$  and

$$\begin{aligned} \lambda^*\{x \in \mathbb{R}^n : x_i \in A\} &= \int \lambda_t\{x \in \mathbb{R}^n : x_i \in A\} Q^*(dt) \\ &= \int \nu_t(A) Q^*(dt) = \int \nu_t(A) Q(dt) = \nu(A) \end{aligned}$$

for all  $A \in \mathcal{B}$  and all  $i = 1, \dots, n$ , where the third equality holds because  $t \mapsto \nu_t(A)$  is  $\mathcal{E}$ -measurable. Hence,  $\lambda^*$  is  $\sigma$ -additive on  $\mathcal{R}$ . Let  $\lambda$  be the only  $\sigma$ -additive extension of  $\lambda^*$  to the Borel  $\sigma$ -field of  $\mathbb{R}^n$ . Since  $\lambda \in \Gamma(\nu, \dots, \nu)$ , to conclude the proof it suffices to see that  $\lambda(H) = 1$ . In fact, since  $H^c$  is open, it is a countable union of open rectangles, that is,  $H^c = \cup_k G_k$  with  $G_k \in \mathcal{R}$  for all  $k$ . Since  $\lambda_t(G_k) \leq \lambda_t(H^c) = 0$  for all  $t \in T$ , one obtains

$$\lambda(G_k) = \lambda^*(G_k) = \int \lambda_t(G_k) Q^*(dt) = 0 \quad \text{for all } k.$$

Therefore  $\lambda(H) = 1$ , for  $H^c$  is a countable union of  $\lambda$ -null sets. □

The second lemma, which slightly generalizes Theorem 2.4 of Wang and Wang [23], provides conditions for a certain probability measure to be  $n$ -CM.

**Lemma 4.4.** *Let  $\alpha \in [0, 1]$ ,  $x < y$ , and  $\nu$  a probability measure on  $[x, y]$  which admits a non-increasing density (with respect to the Lebesgue measure). Then,  $\alpha\delta_x + (1 - \alpha)\nu$  is  $n$ -CM if and only if*

$$\alpha \leq 1 - \frac{y - x}{n(q - x)} \quad \text{where } q = \int z \nu(dz) \text{ is the mean of } \nu. \quad (6)$$

*Proof.* Let  $\gamma = \alpha\delta_x + (1 - \alpha)\nu$ . As noted in Example 3.5, a necessary condition for complete mixability is the mean inequality in Wang and Wang [23], and condition (6) is precisely the mean inequality for  $\gamma$ . Thus, (6) holds if  $\gamma$  is  $n$ -CM. Conversely, suppose (6) holds. Let  $\gamma_\epsilon = \alpha U_{[x, x+\epsilon]} + (1 - \alpha)\nu$ , where  $\epsilon \in (0, y - x)$  and  $U_I$  stands for the uniform distribution on the interval  $I$ . By (6),  $\gamma_\epsilon$  satisfies the mean inequality. By Corollary 2.9 of Wang and Wang [23], since  $\gamma_\epsilon$  has non-increasing density and meets the mean inequality,  $\gamma_\epsilon$  is  $n$ -CM. Further, as  $\epsilon \rightarrow 0$ , the mean of  $\gamma_\epsilon$  converges to  $\alpha x + (1 - \alpha)q$  and  $\gamma_\epsilon \rightarrow \gamma$  weakly. Thus,

$$\gamma \in \mathcal{M}_n(\alpha x + (1 - \alpha)q)$$

because of Theorem 3.1 of Puccetti et al. [16]. □

We are now ready to prove Theorem 4.2.

**Proof of Theorem 4.2.** Recall that  $\mu \in \mathcal{M}_n(0)$  (see Rüschemdorf and Uckelmann [21]) for all  $n \geq 2$ . Since the case  $n = 2$  is trivial, we assume  $n \geq 3$ . Fix  $c \in (0, \frac{\log(n-1)}{\pi}]$ . By Example 4.1, it suffices to show that  $\mu \in \mathcal{M}_n(c)$ . In turn, by Lemma 4.3, it suffices to prove that  $\mu$  can be written as

$$\mu = \int \mu_t Q(dt) \tag{7}$$

where  $\mu_t \in \mathcal{M}_n(c)$  for all  $t > 0$  and  $Q$  is a probability measure on  $(0, \infty)$ .

Let  $f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$ ,  $x \in \mathbb{R}$ , be the standard Cauchy density and  $f^{-1}$  the function on  $[0, 1/\pi]$  given by

$$f^{-1}(0) = \infty, \quad f^{-1}(x) = \sqrt{\frac{1}{\pi x} - 1}, \quad \text{for } x \in (0, 1/\pi].$$

Also, let  $h : (0, \infty) \rightarrow \mathbb{R}$  be a  $C^1$  function such that, for each  $t > 0$ :

$$(I) \int_{c-t}^{c+(n-1)t} (x-c) \{f(x) - h(t)\}_+ dx = 0,$$

$$(II) 0 \leq h(t) \leq f(c+t),$$

$$(III) h'(t) \leq 0.$$

The existence of such  $h$  will be verified at the end of the proof. For the moment, we assume that  $h$  exists.

For  $t > 0$ , let  $\nu_t$  be the finite measure on  $\mathcal{B}$  with density

$$f_t(x) = \{f(x) - h(t)\}_+ \mathbf{1}_{[c-t, c+(n-1)t]}(x).$$

Since  $f_t(x) \leq f(x)$  and  $f(x) = \lim_{t \rightarrow \infty} f_t(x)$  for all  $t > 0$  and  $x \in \mathbb{R}$ , one obtains

$$\lim_{t \rightarrow \infty} \nu_t(B) = \lim_{t \rightarrow \infty} \int_B f_t(x) dx = \int_B f(x) dx = \mu(B)$$

for each  $B \in \mathcal{B}$ . Note also that  $\lim_{t \rightarrow 0} \nu_t(B) = 0$ , for

$$\limsup_{t \rightarrow 0} \nu_t(B) \leq \limsup_{t \rightarrow 0} \nu_t(\mathbb{R}) \leq \limsup_{t \rightarrow 0} \int_{c-t}^{c+(n-1)t} f(x) dx = 0.$$

Denote

$$K_1(t) = f(c-t) - h(t), \quad K_2(t) = (n-1) \{f(c+(n-1)t) - h(t)\}_+,$$

$$K_3(t) = \min\{c+(n-1)t, f^{-1}(h(t))\}, \quad K_4(t) = -h'(t) (K_3(t) - (c-t)).$$

Since  $f(c+t) < f(c-t)$ ,  $t > 0$ , conditions (II)-(III) imply  $K_i(t) \geq 0$  for all  $i$  with  $K_1(t) > 0$

and  $K_3(t) \geq c + t$ . Hence, for each  $t > 0$ , one can define

$$\mu_t = \frac{K_1(t)\delta_{c-t} + K_2(t)\delta_{c+(n-1)t} + K_4(t)U_{[c-t, K_3(t)]}}{K_1(t) + K_2(t) + K_4(t)}$$

where  $U_{[x,y]}$  denotes the uniform distribution on the interval  $[x, y]$ . Such  $\mu_t$  are the probability measures that we use in (7).

Next, fix  $y \in \mathbb{R}$  and define

$$I_y = (-\infty, y), \quad g(t) = \nu_t(I_y), \quad Q_t = K_1(t) + K_2(t) + K_4(t) \quad \text{for each } t > 0.$$

Then,  $g$  is continuous and satisfies, by direct calculation,

$$g'(t) = \frac{d}{dt}\nu_t(I_y) = Q_t \mu_t(I_y) \quad \text{for all } t \in T,$$

where  $T \subset (0, \infty)$  is a co-finite set (possibly depending on  $y$ ). Since  $g'$  is locally integrable (with respect to the Lebesgue measure) it follows that

$$\nu_t(I_y) = \lim_{\epsilon \rightarrow 0} \{g(t) - g(\epsilon)\} = \lim_{\epsilon \rightarrow 0} \int_{\epsilon}^t g'(s) ds = \int_0^t Q_s \mu_s(I_y) ds \quad (8)$$

for all  $t > 0$ . In particular,

$$\begin{aligned} \int_0^\infty Q_s ds &= \lim_{t \rightarrow \infty} \int_0^t Q_s \left( \lim_{y \rightarrow \infty} \mu_s(I_y) \right) ds = \lim_{t \rightarrow \infty} \lim_{y \rightarrow \infty} \int_0^t Q_s \mu_s(I_y) ds \\ &= \lim_{t \rightarrow \infty} \lim_{y \rightarrow \infty} \nu_t(I_y) = \lim_{t \rightarrow \infty} \nu_t(\mathbb{R}) = \mu(\mathbb{R}) = 1. \end{aligned}$$

Let  $Q$  be the probability measure on  $(0, \infty)$  such that  $Q((0, t]) = \int_0^t Q_s ds$  for all  $t > 0$ . Then, condition (8) yields

$$\int_0^\infty \mu_t(I_y) Q(dt) = \lim_{s \rightarrow \infty} \int_0^s Q_t \mu_t(I_y) dt = \lim_{s \rightarrow \infty} \nu_s(I_y) = \mu(I_y) \quad \text{for all } y \in \mathbb{R}.$$

Therefore,

$$\int_0^\infty \mu_t(B) Q(dt) = \mu(B) \quad \text{for each } B \in \mathcal{B}.$$

To prove  $\mu_t \in \mathcal{M}_n(c)$ , it is fundamental to note that  $\mu_t$  has mean  $c$ . Define in fact

$$\begin{aligned} \phi(t) &= \int_{c-t}^{c+(n-1)t} (x-c) \{f(x) - h(t)\}_+ dx = \int_{c-t}^{K_3(t)} (x-c)(f(x) - h(t)) dx \\ &= \int_{c-t}^{K_3(t)} (x-c)f(x) dx - h(t) \int_{c-t}^{K_3(t)} (x-c) dx. \end{aligned}$$

By condition (I),  $\phi(t) = 0$  for all  $t > 0$ . Computing  $\phi'(t)$ , one obtains

$$\begin{aligned} 0 = \phi'(t) &= -tK_1(t) + (n-1)tK_2(t) - h'(t) \int_{c-t}^{K_3(t)} (x-c)dx \\ &= Q_t \int (x-c)\mu_t(dx). \end{aligned}$$

Therefore,  $\mu_t$  has mean  $c$  for all  $t > 0$ .

Having noted this fact, fix  $t > 0$  and define

$$\mu_t^{(1)} = \frac{n-1}{n}\delta_{c-t} + \frac{1}{n}\delta_{c+(n-1)t}.$$

Such  $\mu_t^{(1)}$  has mean  $c$  and is  $n$ -CM by Example 2.2. Hence,  $\mu_t^{(1)} \in \mathcal{M}_n(c)$ . If  $K_4(t) = 0$ , then  $\mu_t$  is a convex combination of  $\delta_{c-t}$  and  $\delta_{c+(n-1)t}$ . Since  $\mu_t$  has mean  $c$ , it follows that  $\mu_t = \mu_t^{(1)} \in \mathcal{M}_n(c)$ .

Suppose now that  $K_4(t) > 0$ . Since  $K_3(t) \geq c+t$ , the mean of  $U_{[c-t, K_3(t)]}$  is not less than  $c$ . Since  $\mu_t$  has mean  $c$ , it follows that

$$\frac{K_1(t)\delta_{c-t} + K_2(t)\delta_{c+(n-1)t}}{K_1(t) + K_2(t)}$$

has a mean smaller than or equal to  $c$ , namely,  $K_1(t) \geq (n-1)K_2(t)$ . By this fact and  $K_4(t) > 0$ , one can define

$$\mu_t^{(2)} = \frac{K_1(t) - (n-1)K_2(t)}{K_1(t) - (n-1)K_2(t) + K_4(t)}\delta_{c-t} + \frac{K_4(t)}{K_1(t) - (n-1)K_2(t) + K_4(t)}U_{[c-t, K_3(t)]}.$$

Since  $\mu_t$  and  $\mu_t^{(1)}$  have mean  $c$  and

$$\mu_t = \frac{nK_2(t)}{Q_t}\mu_t^{(1)} + \frac{K_1(t) - (n-1)K_2(t) + K_4(t)}{Q_t}\mu_t^{(2)},$$

then  $\mu_t^{(2)}$  has mean  $c$  as well. By Lemma 4.4,  $\mu_t^{(2)}$  is  $n$ -CM (condition (6) follows from  $\mu_t^{(2)}$  having mean  $c$ ). Therefore,  $\mu_t^{(2)} \in \mathcal{M}_n(c)$ . Finally, since  $\mu_t^{(i)} \in \mathcal{M}_n(c)$  for  $i = 1, 2$  and  $\mathcal{M}_n(c)$  is convex (by Lemma 4.3), one obtains  $\mu_t \in \mathcal{M}_n(c)$ .

To conclude the proof, it remains only to prove that a  $C^1$ -function  $h$  satisfying conditions (I)-(II)-(III) actually exists. Define

$$A(t, y) = \int_{c-t}^{c+(n-1)t} (x-c)\{f(x) - y\}_+ dx \quad \text{for all } t > 0 \text{ and } y \in \mathbb{R}.$$

Then,  $A$  is a  $C^1$ -function on  $(0, \infty) \times \mathbb{R}$  and

$$\frac{\partial A}{\partial y}(t, y) = \int_{c-t}^{c+(n-1)t} -(x-c)\mathbf{1}_{\{f > y\}}(x) dx,$$

where  $\{f > y\}$  denotes the set  $\{x \in \mathbb{R} : f(x) > y\}$ . Fix  $t > 0$ . If  $y < f(c+t)$ , there is

$u \in (c+t, c+(n-1)t]$  such that  $f(x) > y$  for every  $x \in [c-t, u]$ . Thus,

$$\frac{\partial A}{\partial y}(t, y) \leq - \int_{c-t}^u (x-c) dx = - \int_{c+t}^u (x-c) dx < 0.$$

Hence, the map  $y \mapsto A(t, y)$  is continuous, strictly decreasing on  $(-\infty, f(c+t)]$ , and

$$\begin{aligned} \lim_{y \rightarrow -\infty} A(t, y) &= \infty, \quad A(t, f(c+t)) = \int_{c-t}^{c+t} (x-c)(f(x) - f(c+t)) dx \\ &= \int_{-t}^t xf(c+x) dx < 0. \end{aligned}$$

It follows that, for each  $t > 0$ , there exists a unique number  $h(t)$  satisfying  $h(t) < f(c+t)$  and  $A(t, h(t)) = 0$ . It remains to see that  $h \geq 0$ ,  $h$  is  $C^1$  and  $h' \leq 0$ .

We begin with  $h \geq 0$ . Since  $A(t, y) > A(t, 0)$  whenever  $y < 0$ , it suffices to see that  $A(t, 0) \geq 0$ . Define

$$m(t) = \frac{1}{t} \frac{d}{dt} A(t, 0) = (n-1)^2 f(c+(n-1)t) - f(c-t) \quad \text{for } t > 0.$$

Then,

$$m'(t) = (n-1)^3 f'(c+(n-1)t) + f'(c-t) = (n-1)^3 f'(c+(n-1)t) - f'(t-c).$$

Observe now that, since  $\sqrt{1/f}$  is convex (see Remark 4.5),

$$\frac{d\sqrt{1/f(x)}}{dx} = -\frac{1}{2} \frac{f'(x)}{f^{3/2}(x)}$$

is an increasing function of  $x \in \mathbb{R}$ . Therefore,

$$\frac{f'(c+(n-1)t)}{(f(c+(n-1)t))^{3/2}} \leq \frac{f'(t-c)}{(f(t-c))^{3/2}}$$

and, by rearranging terms,

$$\frac{f'(t-c)}{(n-1)^3 f'(c+(n-1)t)} \leq \left( \frac{f(t-c)}{(n-1)^2 f(c+(n-1)t)} \right)^{3/2}. \quad (9)$$

If  $m(t) \geq 0$ , then  $(n-1)^2 f(c+(n-1)t) \geq f(c-t) = f(t-c)$ , so that the right-hand member of (9) is bounded above by 1. Hence,  $m(t) \geq 0$  implies  $m'(t) \leq 0$ . Thanks to this fact and  $\lim_{t \downarrow 0} m(t) > 0$ , one concludes that there is some  $t_0 > 0$  (possibly  $t_0 = \infty$ ) such that  $m(t) \geq 0$  for  $t \leq t_0$  and  $m(t) \leq 0$  for  $t \geq t_0$ . Since  $m(t)$  and  $\frac{d}{dt} A(t, 0)$  have the same sign and

$$\lim_{t \downarrow 0} A(t, 0) = 0, \quad \lim_{t \rightarrow \infty} A(t, 0) = \lim_{t \rightarrow \infty} \int_{c-t}^{c+(n-1)t} xf(x) dx - c = \frac{\log(n-1)}{\pi} - c \geq 0,$$

one finally obtains  $A(t, 0) \geq 0$ . Therefore,  $h(t) \geq 0$  for all  $t > 0$ .

To prove  $h$  is  $C^1$ , recall that  $h(t)$  is the only real number such that  $h(t) < f(c+t)$  and  $A(t, h(t)) = 0$ . Also,  $h(t) < f(c+t)$  implies  $\frac{\partial A}{\partial y}(t, h(t)) < 0$ . Thus,  $h$  is  $C^1$  because of the



implicit function theorem.

We finally prove  $h' \leq 0$ . Since  $\frac{\partial A}{\partial y}(t, h(t)) < 0$  and

$$h'(t) = -\frac{\frac{\partial A}{\partial t}(t, h(t))}{\frac{\partial A}{\partial y}(t, h(t))},$$

it suffices to show that  $\frac{\partial A}{\partial t}(t, h(t)) \leq 0$  for  $t > 0$ . Since

$$\frac{1}{t} \frac{\partial A}{\partial t}(t, h(t)) = (n-1)^2 \{f(c + (n-1)t) - h(t)\}_+ - \{f(c-t) - h(t)\}_+,$$

it can be assumed  $f(c + (n-1)t) > h(t)$ . In this case,

$$\frac{1}{t} \frac{\partial A}{\partial t}(t, h(t)) = (n-1)^2 f(c + (n-1)t) - f(c-t) - n(n-2)h(t) = m(t) - n(n-2)h(t).$$

Recall now that  $m(t) \geq 0$  for  $t \leq t_0$  and  $m(t) \leq 0$  for  $t \geq t_0$  where  $t_0 \in (0, \infty]$ . Hence,  $\frac{\partial A}{\partial t}(t, h(t)) \leq 0$  for  $t \geq t_0$ . If  $t \in (0, t_0)$ , since  $t m(t) = \frac{d}{dt} A(t, 0)$  and  $m$  is decreasing on  $(0, t_0)$ , then

$$A(t, 0) = \int_0^t s m(s) ds \geq m(t) \int_0^t s ds = t^2 m(t)/2.$$

On the other hand, it is easily seen that

$$\frac{\partial A}{\partial y}(t, y) \geq -\int_{c-t}^{c+(n-1)t} (x-c) dx = -n(n-2)t^2/2.$$

Therefore,

$$-A(t, 0) = A(t, h(t)) - A(t, 0) = \int_0^{h(t)} \frac{\partial A}{\partial y}(t, y) dy \geq -h(t) n(n-2)t^2/2.$$

It follows that  $m(t) \leq n(n-2)h(t)$ , and again  $\frac{\partial A}{\partial t}(t, h(t)) \leq 0$ . To summarize,  $h'(t) \leq 0$  for all  $t > 0$ , namely,  $h$  satisfies conditions (I)-(II)-(III). This concludes the proof.  $\square$

**Remark 4.5.** The proof of Theorem 4.2 is valid for other probability measures in addition to the Cauchy distribution. Fix in fact a probability measure  $\nu$  on  $\mathbb{R}$  which admits a density  $g$  with respect to the Lebesgue measure. If  $g$  is strictly positive, differentiable, symmetric and strictly unimodal (that is,  $g' > 0$  on  $(-\infty, 0)$  and  $g' < 0$  on  $(0, \infty)$ ), and if  $\sqrt{1/g}$  is a convex function, then any real number  $q$  satisfying

$$|q| \leq \liminf_{t \rightarrow \infty} \int_t^{(n-1)t} x g(x) dx, \quad (10)$$

is an  $n$ -center of  $\nu$ . This follows from replacing  $\mu$  and  $c$  by  $\nu$  and  $|q|$ , respectively, in the proof of Theorem 4.2. Note that, if  $\nu$  has finite mean, the right-hand side in (10) is zero and (10) reduces to the well known fact that unimodal symmetric distributions have a center at zero (even if  $\sqrt{1/g}$  fails to be convex); see Rüschemdorf and Uckelmann [21]. Therefore, taking an arbitrary differentiable, symmetric, strictly positive and strictly convex function

$\phi$ , by defining a density  $g = \frac{\beta}{\phi^2}$  for some normalizing constant  $\beta > 0$  one always finds a distribution  $\nu$  fulfilling the requirements of the proof of Theorem 4.2. However, apart from the Cauchy distribution, we do not know of any natural example of such  $\nu$ . This is due to the convexity of  $\sqrt{1/g}$ , which is a quite restrictive requirement. For instance, suppose that  $g$  has the power form  $g(x) = \frac{\beta}{1+|x|^\alpha}$  for some constants  $\alpha, \beta > 0$ . In this case,  $\alpha = 2$  leads to the standard Cauchy distribution,  $\alpha > 2$  implies that the mean of  $\nu$  is finite (so that 0 is the unique center), and  $\alpha < 2$  implies that  $\sqrt{1/g}$  is not convex.

## References

- [1] Bellini, F., O. E. Karışan, and M. Ç. Pınar (2016). Joint mixability of some integer matrices. *Discrete Optim.* 20, 90–104.
- [2] Bernard, C., X. Jiang, and R. Wang (2014). Risk aggregation with dependence uncertainty. *Insurance Math. Econom.* 54, 93–108.
- [3] Bernard, C., L. Rüschendorf, and S. Vanduffel (2017). Value-at-risk bounds with variance constraints. *J. Risk Insurance* 84(3), 923–959.
- [4] Boudt, K., E. Jakobsons, and S. Vanduffel (2017). Block rearranging elements within matrix columns to minimize the variability of the row sums. Forthcoming in *4OR*, DOI: 10.1007/s10288-017-0344-4.
- [5] Chen, R. and L. A. Shepp (1983). On the sum of symmetric random variables. *Amer. Statist.* 37(3), 237.
- [6] Coffman, Jr., E. G. and M. Yannakakis (1984). Permuting elements within columns of a matrix in order to minimize maximum row sum. *Math. Oper. Res.* 9(3), 384–390.
- [7] Embrechts, P., H. Liu, and R. Wang (2016). Quantile-based risk sharing. Preprint, available at <https://ssrn.com/abstract=2744142>.
- [8] Embrechts, P. and G. Puccetti (2006). Bounds for functions of dependent risks. *Finance Stoch.* 10(3), 341–352.
- [9] Embrechts, P., G. Puccetti, and L. Rüschendorf (2013). Model uncertainty and VaR aggregation. *J. Bank. Financ.* 37(8), 2750–2764.
- [10] Gaffke, N. and L. Rüschendorf (1981). On a class of extremal problems in statistics. *Math. Operationsforsch. Statist. Ser. Optim.* 12(1), 123–135.
- [11] Graham, R. L. (1966). Bounds for certain multiprocessing anomalies. *Bell System Tech. J.* 45(9), 1563–1581.
- [12] Haus, U.-U. (2015). Bounding stochastic dependence, joint mixability of matrices, and multidimensional bottleneck assignment problems. *Oper. Res. Lett.* 43(1), 74–79.
- [13] Lee, W. and J. Y. Ahn (2014). On the multidimensional extension of countermonotonicity and its applications. *Insurance Math. Econ.* 56, 68–79.

- [14] Pillai, N. S. and X.-L. Meng (2016). An unexpected encounter with Cauchy and Lévy. *Ann. Statist.* 44(5), 2089–2097.
- [15] Puccetti, G. and L. Rüschendorf (2012). Computation of sharp bounds on the distribution of a function of dependent risks. *J. Comput. Appl. Math.* 236(7), 1833–1840.
- [16] Puccetti, G., B. Wang, and R. Wang (2012). Advances in complete mixability. *J. Appl. Probab.* 49(2), 430–440.
- [17] Puccetti, G. and R. Wang (2015). Extremal dependence concepts. *Stat. Science* 30(4), 485–517.
- [18] Ramachandran, D. (1996). The marginal problem in arbitrary product spaces. In *Distribution with Fixed Marginals and Related Topics*, pp. 260–272. Hayward, CA: Inst. Math. Stat.
- [19] Rüschendorf, L. (1981). Sharpness of Fréchet-bounds. *Z. Wahrsch. Verw. Gebiete* 57(2), 293–302.
- [20] Rüschendorf, L. (2013). *Mathematical risk analysis. Dependence, risk bounds, optimal allocations and portfolios*. Heidelberg: Springer.
- [21] Rüschendorf, L. and L. Uckelmann (2002). Variance minimization and random variables with constant sum. In *Distributions with given marginals and statistical modelling*, pp. 211–222. Dordrecht: Kluwer Acad. Publ.
- [22] Simons, G. (1977). An unexpected expectation. *Ann. Probability* 5(1), 157–158.
- [23] Wang, B. and R. Wang (2011). The complete mixability and convex minimization problems with monotone marginal densities. *J. Multivariate Anal.* 102(10), 1344–1360.
- [24] Wang, B. and R. Wang (2016). Joint mixability. *Mathematics of Operations Research* 41(3), 808–826.
- [25] Wang, R. (2015). Current open questions in complete mixability. *Probab. Surv.* 12, 13–32.
- [26] Wang, R., L. Peng, and J. Yang (2013). Bounds for the sum of dependent risks and worst value-at-risk with monotone marginal densities. *Finance Stoch.* 17(2), 395–417.