Symmetric Bernoulli distributions and minimal dependence copulas

Journal of Multivariate Analysis 212 (2026) 105545

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Symmetric Bernoulli distributions and minimal dependence copulas

Alessandro Mutti *, Patrizia Semeraro

Department of Mathematical Sciences "G.L. Lagrange", Politecnico di Torino, Turin, Italy

ARTICLE INFO

AMS 2020 subject classifications:
primary 62E10
secondary 62H20

Keywords:
Convex order
Extremal mixture copulas
FGM copulas
Negative dependence
Symmetric Bernoulli distributions

ABSTRACT

The key result of this paper is to characterize all multivariate symmetric Bernoulli distributions whose sum is minimal under the convex order. In doing so, we automatically characterize extremal negative dependence among Bernoulli variables, since multivariate distributions with minimal convex sums are known to be strongly negative dependent. Moreover, beyond its interest per se, this result provides insight into negative dependence within the class of copulas. In particular, two classes of copulas can be built from multivariate symmetric Bernoulli distributions: extremal mixture copulas and FGM copulas. We analyze the extremal negative dependence structures of copulas constructed from symmetric Bernoulli vectors with minimal convex sums and explicitly find a class of minimal dependence copulas. This analysis is completed by investigating minimal pairwise dependence measures and correlations. Our main results derive from the geometric and algebraic representations of multivariate symmetric Bernoulli distributions, which effectively encode key statistical properties.

1. Introduction

A problem extensively studied in applied probability is finding bounds for sums $S = X_1 + \cdots + X_d$ of random variables with joint distribution in a given Fréchet class $\mathcal{F}_d(F_1, \ldots, F_d)$, i.e., the class of all joint distributions with one-dimensional marginals F_j (see, e.g., [1–5]). We consider bounds in the convex order, which are important in fields such as insurance and finance, where $\boldsymbol{X} = (X_1, \ldots, X_d)$ represents a portfolio of d risks and S represents its aggregate risk. Roughly speaking, the convex order indicates which of two aggregate risks has the lower variability. The problem of finding the upper bound has been solved: the upper bound is reached when the risks are comonotonic and their joint distribution is the upper Fréchet bound, i.e., the maximum element of $\mathcal{F}_d(F_1, \ldots, F_d)$ in concordance order [6]. Finding the lower bound is less straightforward: in dimension 2, the solution is the lower Fréchet bound, but for $d \geq 3$, the lower Fréchet bound generally fails to be a distribution; see [7]. The problem of finding distributions in $\mathcal{F}_d(F_1, \ldots, F_d)$ corresponding to minimal aggregate risk is as yet unsolved, and this is the problem we focus on. Following [8], we call the random vectors with minimal convex sums and their distributions the Σ_{cx}-smallest elements in $\mathcal{F}_d(F_1, \ldots, F_d)$, when they exist.

We consider two Fréchet classes. The first is the class $\mathcal{SB}_d$ of d-dimensional distributions with one-dimensional Bernoulli marginals of mean $p = 1/2$, called multivariate symmetric Bernoulli distributions. Multivariate Bernoulli distributions and their properties have been widely investigated in the statistical literature due to the importance of binary data in applications; see, e.g., [9–12]. The other class is the class of all copulas, i.e., multivariate distribution functions with one-dimensional uniform marginals [13]. Copulas are widely used to represent the dependence among risks in insurance and finance. Usually, marginal risks and their dependence structure are modeled separately, since using Sklar's theorem, it is possible to model dependence among risks with any given distribution using copulas; see, e.g., [14–16]. An application of copulas to financial risk analysis is provided

* Corresponding author.
E-mail address: alessandro.mutti@polito.it (A. Mutti).

https://doi.org/10.1016/j.jmva.2025.105545
Received 11 June 2025; Received in revised form 15 November 2025; Accepted 16 November 2025
Available online 17 November 2025

in [17]. Therefore, characterizing the class of copulas corresponding to minimal aggregate risk is an important step in understanding the dependence structures associated with low aggregate risk.

Although not all Fréchet classes admit a Σ_{cx}-smallest element (see Example 3.1 of [18]), there always exists a multivariate Bernoulli random vector with marginal means $p_1, \dots, p_d \in (0,1)$ that is a Σ_{cx}-smallest element in its Fréchet class. For symmetric Bernoulli distributions, the probability mass function (pmf) of the minimal sum in the convex order has support on the two adjacent points $(d-1)/2$ and $(d+1)/2$ if d is odd, or is degenerate at $d/2$ if d is even. In the literature, several approaches exist to find a Σ_{cx}-smallest element: one may consider the unique exchangeable solution (e.g., [19]), or non-exchangeable solutions as in Theorem 5.2 in [20] and Lemma 3.1 in [21]. However, the above approaches yield trivial solutions in the case of symmetric Bernoulli distributions, such as multivariate pmfs with support on only two points.

Our contribution is to find and characterize all Σ_{cx}-smallest elements in the Fréchet class $\mathcal{SB}_d$. This also provides a complete characterization of a class of extremal negative dependent symmetric Bernoulli random vectors. Extremal negative dependence is well-defined in dimension two as the maximal negative dependence between two random variables, and it is called countermonotonicity. Two variables are countermonotonic if a high value of one variable almost surely corresponds to a low value of the other; see [22] for the formal definition. The generalization to dimensions higher than 2 is not trivial and not unique. The authors of [8] review the main notions of extremal negative dependence and introduce a new concept, Σ-countermonotonicity. Theorem 3.8 in [8] shows that any Σ_{cx}-smallest element is necessarily Σ-countermonotonic. In the Bernoulli setting, [23] proves the converse implication: if a Bernoulli random vector $\boldsymbol{X}$ is Σ-countermonotonic, then it is a Σ_{cx}-smallest element in its Fréchet class. Therefore, by characterizing the class of Σ_{cx}-smallest elements in $\mathcal{SB}_d$, we also characterize Σ-countermonotonicity.

Although these results are of interest per se, they also contribute to the study of negative dependence within a more general framework. In fact, they allow us to characterize explicitly a class of Σ-countermonotonic copulas, i.e., minimal dependence copulas. Extreme negative dependence and its relationship with minimal risk have been extensively studied in the context of insurance and finance; see, among others, [15,22,24,25]. In this framework, the theory of copulas provides a useful tool to model dependence and find distributional bounds for dependent risks [5,15,26]. We consider two classes of copulas that can be built from multivariate symmetric Bernoulli distributions: the extremal mixture copulas [27,28] and the Farlie–Gumbel–Morgenstern (FGM) copulas [29]. While FGM copulas are in a one-to-one relationship with the elements of the class $\mathcal{SB}_d$, the extremal mixture copulas are in a one-to-one relationship with the subclass of palindromic Bernoulli distributions. We study the dependence structure of the copulas constructed from Σ_{cx}-smallest Bernoulli distributions in these two classes.

It is proved in this paper that it is always possible to find a class of extremal copulas — a subclass of the extremal mixture copulas — that are Σ-countermonotonic. This result can be improved when the dimension of the Fréchet class d is even: in this case, the extremal mixture copulas defined by Σ_{cx}-smallest Bernoulli distributions are Σ_{cx}-smallest elements in the Fréchet class of copulas. Therefore, if d is even, multivariate uniform variables have a minimum risk element. In [30], the authors prove that the FGM copulas constructed from the Σ_{cx}-smallest elements in $\mathcal{SB}_d$ are Σ_{cx}-smallest elements within the class of FGM copulas, although not within their Fréchet class. For this reason, we investigate the negative dependence of Σ_{cx}-smallest FGM copulas employing three widely used measures of dependence, viz. Pearson's correlation, Spearman's rho, and Kendall's tau.

Our results follow from the geometrical and algebraic representations of the class $\mathcal{SB}_d$. This class can be represented as a convex polytope [31] whose extremal points encode relevant statistical properties, such as extremal dependence or distributional bounds for relevant risk measures. Although extremal points can be found in closed form in special classes [32] and analytically in low dimension [31], finding them in high dimension becomes computationally infeasible. For this reason, in [20], the authors find a way around this limitation and map the class of multivariate Bernoulli distributions with given means $p \in (0,1)$ into an ideal of points in the ring of polynomials with rational coefficients. Using the results in [20], we find an analytical set of polynomials that generate the class $\mathcal{SB}_d$ and an analytical set of polynomials that generate the class of palindromic distributions. The latter generators are extremal points of the polytope and are associated with the extremal copulas. These connections allow us to find the Σ-countermonotonic extremal copulas. Indeed, the effectiveness of the algebraic representation is that the polynomial coefficients can be used to construct multivariate Bernoulli distributions with given statistical properties.

Summing up, the algebraic representation constitutes the basis for characterizing extremal negative dependence within the class $\mathcal{SB}_d$, which in turn leads to the characterization of extremal negative dependence in the class of copulas. Therefore, we introduce the algebraic representation in Section 2, as it is preliminary to all other results. The proofs of this section are technical and are postponed to Appendix A, while some examples are given in Appendix C. The rest of the paper is organized as follows. The notions of negative dependence and our main results are presented in Section 3. Specifically, in Section 3.1, we characterize the Σ_{cx}-smallest elements in $\mathcal{SB}_d$ using the algebraic representation. The proofs in Section 3.1 rely on the algebraic approach and are therefore presented in Appendix B. In Section 3.2, we study both the extremal mixture copulas and the FGM copulas, characterizing those corresponding to the Σ_{cx}-smallest symmetric Bernoulli distributions. We also identify a family of Σ-countermonotonic copulas. The proofs of this section, based on a probabilistic approach, are included in the main text. Section 4 is devoted to the study of minimal pairwise dependence measures and minimal correlations in $\mathcal{SB}_d$ and in the two classes of copulas. Concluding remarks are given in Section 5.

2. Algebraic representation

In this section, we present the algebraic representation of $\mathcal{SB}_d$. This representation was introduced in [20] for any Fréchet class of joint Bernoulli distributions with identical one-dimensional marginals of mean $p \in (0,1) \cap \mathbb{Q}$. We consider the case $p = 1/2$, corresponding to the class $\mathcal{SB}_d$, and prove new results required for the study of extremal negative dependence.

Since the proofs in this section are based on the algebraic approach, they are provided in Appendix A for clarity and readability.

We assume that vectors $\boldsymbol{x} = (x_1, \dots, x_d)$ are column vectors and we denote by $A^\top$ the transpose of a matrix A. Given two matrices $A \in \mathcal{M}(n \times m)$ and $B \in \mathcal{M}(d \times \ell)$, we write $A \| B$ for the row concatenation of A and B when $n = d$, and $A//B$ for their column concatenation when $m = \ell$. Given a Bernoulli random vector $\boldsymbol{X} = (X_1, \dots, X_d)$ with pmf $f : \{0,1\}^d \to [0,1]$, $f \in \mathcal{SB}_d$, we denote by $\boldsymbol{f} = (f_1, \dots, f_{2^d})$ the vector that contains the values and f over $\mathcal{X}_d = \{0,1\}^d$, i.e., $\boldsymbol{f} := (f(\boldsymbol{x}) : \boldsymbol{x} \in \mathcal{X}_d)$. We make the non-restrictive hypothesis that the set $\mathcal{X}_d$ of 2^d binary d-dimensional vectors is ordered according to the reverse-lexicographical criterion. For example, for $d = 3$, we have $\mathcal{X}_3 = \{000, 100, 010, 110, 001, 101, 011, 111\}$. Given that a pmf $f \in \mathcal{SB}_d$ uniquely determines a vector $\boldsymbol{f}$ (and vice versa), we also use the term pmfs to denote the vectors $\boldsymbol{f}$. By $\boldsymbol{X} \in \mathcal{SB}_d$ and $\boldsymbol{f} \in \mathcal{SB}_d$, we mean that the random vector $\boldsymbol{X}$ has pmf $f \in \mathcal{SB}_d$.

Finally, we denote by $P(\boldsymbol{z}) = \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}} a_{\boldsymbol{i}} \boldsymbol{z}^{\boldsymbol{i}}$ a polynomial in the ring $\mathbb{Q}[\boldsymbol{z}]$ of polynomials with rational coefficients $a_{\boldsymbol{i}}$ in the variables $\boldsymbol{z} = (z_1, \dots, z_{d-1})$, where $\boldsymbol{z}^{\boldsymbol{i}} = \prod_{j=1}^{d-1} z_j^{i_j}$. To simplify the notation, we write $a_{i_1 \dots i_{d-1}} := a_{(i_1, \dots, i_{d-1})} = a_{\boldsymbol{i}}$.

In [31], the authors show that $\mathcal{SB}_d$ is a convex polytope, viz.

$$\mathcal{SB}_d = \left\{ \boldsymbol{f} \in \mathbb{R}^{2^d} : H_d \boldsymbol{f} = \boldsymbol{0}, f_j \geq 0, \sum_{j=1}^{2^d} f_j = 1 \right\},$$

where H_d is a $d \times 2^d$ matrix whose rows are $(\boldsymbol{1}_{2^d} - 2\boldsymbol{x}_h)^\top$, $h \in \{1, \dots, d\}$, where $\boldsymbol{1}_{2^d}$ is the 2^d-vector with all elements equal to 1, and $\boldsymbol{x}_h$ is the 2^d-vector that contains the hth components of all the d-vectors $\boldsymbol{x} \in \mathcal{X}_d$. Therefore, $\mathcal{SB}_d$ is the convex hull of a finite set of points $\boldsymbol{r}_k \in \mathcal{SB}_d$, $k \in \{1, \dots, n_d\}$, called extremal points or extremal pmfs. In other terms, for any $\boldsymbol{f} \in \mathcal{SB}_d$, there exist n_d positive weights $\lambda_1, \dots, \lambda_{n_d}$ summing up to one such that

$$\boldsymbol{f} = \sum_{i=1}^{n_d} \lambda_i \boldsymbol{r}_i.$$

When the dimension d is sufficiently small, the extremal points of the convex polytope can be found using, for example, `4ti2`; see [31]. However, this representation has computational limitations. When the dimension d increases, due to the growth of the number n_d, finding all the extremal pmfs becomes computationally infeasible. For example, for the middle-size case $d = 6$, the class $\mathcal{SB}_6$ has $n_6 = 707{,}264$ extremal points. To overcome this limitation, the authors of [20] introduce a new algebraic representation of any Fréchet class of joint Bernoulli distributions with the same one-dimensional marginals with common mean $p \in (0,1) \cap \mathbb{Q}$, that proves to be extremely effective in the study of the case $p = 1/2$, i.e., the class $\mathcal{SB}_d$.

Following [20], we define the linear map $\mathcal{H}$ from the class $\mathcal{SB}_d$ to the polynomial ring with rational coefficients $\mathbb{Q}[z_1, \dots, z_{d-1}]$ as

$$\mathcal{H} : \mathcal{SB}_d \to \mathbb{Q}[z_1, \dots, z_{d-1}], \quad \boldsymbol{f} \mapsto \mathcal{H}(\boldsymbol{f}) = \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}} a_{\boldsymbol{i}} \boldsymbol{z}^{\boldsymbol{i}}, \tag{1}$$

where, for every $\boldsymbol{f} \in \mathcal{SB}_d$, the vector of coefficients $\boldsymbol{a} = (a_{\boldsymbol{i}})_{\boldsymbol{i} \in \mathcal{X}_{d-1}}$ in (1) is given by

$$\boldsymbol{a} = Q\boldsymbol{f}, \tag{2}$$

with $Q = (I(2^{d-1}) \| \tilde{I}(2^{d-1}))$, where $I(2^{d-1})$ is the identity matrix of order 2^{d-1} and $\tilde{I}(2^{d-1})$ is the square matrix of order 2^{d-1} with -1 on the anti-diagonal and 0 elsewhere. For every $\boldsymbol{i} \in \mathcal{X}_{d-1}$, set $\boldsymbol{s}_{\boldsymbol{i}} := (\boldsymbol{i}//0) = (i_1, \dots, i_{d-1}, 0)$. Because of the form of the matrix Q in (2), we can write the image of $\boldsymbol{f}$ through $\mathcal{H}$ as

$$\mathcal{H}(\boldsymbol{f}) = \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}} \{ f(\boldsymbol{s}_{\boldsymbol{i}}) - f(\boldsymbol{1}_d - \boldsymbol{s}_{\boldsymbol{i}}) \} \boldsymbol{z}^{\boldsymbol{i}}. \tag{3}$$

We call $\mathcal{C}_{\mathcal{H}}$ the image of $\mathcal{SB}_d$ through $\mathcal{H}$. From Theorem 3.1 in [20], $\mathcal{C}_{\mathcal{H}} \subseteq \mathcal{I}_{\mathcal{P}}$, where $\mathcal{I}_{\mathcal{P}} \subseteq \mathbb{Q}[\boldsymbol{z}]$ is the ideal of polynomials that vanish at points $\mathcal{P} = \{\boldsymbol{1}_{d-1}, \boldsymbol{1}_{d-1}^{-j}, j \in \{1, \dots, d-1\}\}$, where $\boldsymbol{1}_{d-1}^{-j}$ is a vector of length $d-1$ with -1 in position j and 1 elsewhere.

Example 8 in Appendix C provides an example of the polynomial representation for $d = 3$ and shows that the map $\mathcal{H}$ is not injective. Indeed, the authors of [20] find a basis of the kernel $\mathcal{K}(\mathcal{H})$, the set of pmfs $\boldsymbol{f}$ such that $\mathcal{H}(\boldsymbol{f}) = 0$. A basis of $\mathcal{K}(\mathcal{H})$ is the set

$$\begin{aligned} \mathcal{B}_K &= \left\{ f \in \mathcal{SB}_d : \exists \boldsymbol{x} \in \mathcal{X}_d \text{ such that } f(\boldsymbol{x}) = f(\boldsymbol{1}_d - \boldsymbol{x}) = \tfrac{1}{2} \right\} \\ &= \left\{ \left(\tfrac{1}{2}, 0, 0, \dots, 0, 0, \tfrac{1}{2}\right), \left(0, \tfrac{1}{2}, 0, \dots, 0, \tfrac{1}{2}, 0\right), \left(0, 0, \tfrac{1}{2}, \dots, \tfrac{1}{2}, 0, 0\right), \dots \right\}. \end{aligned} \tag{4}$$

Note that if $\boldsymbol{f} \in \mathcal{B}_K$ it has support on two points. We denote by $\mathcal{PB}_d$ the class of d-dimensional palindromic Bernoulli pmfs, i.e., the pmfs f of Bernoulli random vectors such that $f(\boldsymbol{x}) = f(\boldsymbol{1}_d - \boldsymbol{x})$, for every $\boldsymbol{x} \in \mathcal{X}_d$; see [12] as a reference for palindromic distributions. The proofs of the following propositions are straightforward, yet the results are important for our purposes, because palindromic Bernoulli distributions generate the class of extremal mixture copulas [27], which are one of our objects of study. The proof of Proposition 1 is trivial and omitted.

Proposition 1. *The kernel of the map $\mathcal{H}$ coincides with the set of palindromic Bernoulli distributions, $\mathcal{K}(\mathcal{H}) \equiv \mathcal{PB}_d$.*

Proposition 2. *The pmfs of the basis $\mathcal{B}_K$ in (4) are extremal points of the polytope $\mathcal{SB}_d$.*

The basis $\mathcal{B}_K$ has 2^{d-1} pmfs; therefore, there are 2^{d-1} extremal points of $\mathcal{SB}_d$ that have null polynomial. The kernel $\mathcal{K}(\mathcal{H})$ is now fully characterized. It is more challenging to characterize the counter-image of a non-null polynomial. In [20], the authors suggest an algorithm to find a particular distribution from a given polynomial, which they call the type-0 pmf. Algorithm 1 is the version of their algorithm adapted to the class $\mathcal{SB}_d$.

Algorithm 1

Input: A polynomial $P(z) = \sum_{i\in\mathcal{X}_{d-1}} a_i z^i \in \mathcal{I}_{\mathcal{P}}$, $P(z) \not\equiv 0$.

For each $i \in \mathcal{X}_{d-1}$:
if $a_i \geq 0$, then $f^P(s_i) = a_i$ and $f^P(\mathbf{1}_d - s_i) = 0$;
if $a_i < 0$, then $f^P(s_i) = 0$ and $f^P(\mathbf{1}_d - s_i) = -a_i$.
Normalize f^P getting, with a small abuse of notation, $f^P = f^P/(\sum_{x\in\mathcal{X}_d} f^P(x))$.

Output: The type-0 pmf $f^P = (f_1^P, \dots, f_{2^d}^P) \in \mathcal{SB}_d$ corresponding to $P(z)$.

Propositions 3 and 4 are necessary to prove Theorem 2, one of our main results on negative dependence in the class $\mathcal{SB}_d$. A preliminary definition is necessary to introduce the concept of equivalence between polynomials of the ideal $\mathcal{I}_{\mathcal{P}}$.

Definition 1. Two polynomials $P(z)$ and $Q(z)$ of the ideal $\mathcal{I}_{\mathcal{P}}$ are equivalent, denoted by $P(z) \simeq Q(z)$, if there exists a constant $\mu > 0$, $\mu \in \mathbb{Q}$, such that $P(z) = \mu Q(z)$. We denote by $[P(z)] = \{Q(z) \in \mathcal{I}_{\mathcal{P}} : Q(z) \simeq P(z)\}$ the set of all the polynomials equivalent to $P(z)$.

Proposition 3. *Two equivalent polynomials generate the same type-0 pmf.*

Proposition 4 characterizes the set $\mathcal{H}^{-1}[P(z)] := \{ f \in \mathcal{SB}_d : \mathcal{H}(f) \in [P(z)] \}$, the collection of pmfs that $\mathcal{H}$ maps to a polynomial equivalent to $P(z) \in \mathcal{I}_{\mathcal{P}}$. This proposition is crucial for identifying all extremal negative dependent Bernoulli random vectors, which will be characterized through their polynomials.

Proposition 4. *Consider a polynomial $P(z) = \sum_{i\in\mathcal{X}_{d-1}} a_i z^i \in \mathcal{I}_{\mathcal{P}}$, such that $P(z) \not\equiv 0$. Then,*

$$\mathcal{H}^{-1}[P(z)] = \left\{ f \in \mathcal{SB}_d : f = \lambda f^P + (1-\lambda) f^K, \text{ for } f^K \in \mathcal{PB}_d, \lambda \in (0,1] \cap \mathbb{Q} \right\},$$

where f^P is the type-0 pmf of $P(z)$.

We conclude this section with the following proposition that highlights the importance of the type-0 pmfs and their link with the generators of $\mathcal{SB}_d$ as a convex polytope.

Proposition 5. *Every extremal point of $\mathcal{SB}_d$ is either a type-0 pmf or an element of $\mathcal{B}_K$.*

Remark 1. Proposition 3 holds for any Fréchet class of joint distributions with Bernoulli marginals of common mean $p \in [0,1] \cap \mathbb{Q}$. For $p \neq 1/2$, the argument is analogous, although the proof relies on the algorithm to find the type-0 pmfs in [20]. Example 9 in Appendix C shows, instead, that Propositions 4 and 5 hold only for $p = 1/2$, i.e., in the class $\mathcal{SB}_d$.

3. Minimal convex sums and extremal negative dependence

In this section, we recall the main ingredients of negative dependence and study the links between extremal negative dependence and minimality in the convex order. When studying negative dependence and, in particular, extremal negative dependence, the starting point is the definition of countermonotonicity.

Definition 2. A bivariate random vector (X, Y) is said to be countermonotonic if

$$\Pr\left\{ (X_1 - X_2)(Y_1 - Y_2) \leq 0 \right\} = 1,$$

where (X_1, Y_1) and (X_2, Y_2) are two independent copies of (X, Y).

Although this definition provides a clear characterization of extremal negative dependence for two-dimensional Fréchet classes, there is no unique and straightforward generalization to higher dimensions. Several approaches have been proposed to define notions of minimal dependence in Fréchet classes of dimension higher than 2. These notions are known as extremal negative dependence concepts; see [8]. An intuitive generalization of countermonotonicity is the notion of pairwise countermonotonicity, recently studied in [33].

Definition 3. A random vector $Y = (Y_1, \dots, Y_d)$ is pairwise countermonotonic if the pair (Y_{j_1}, Y_{j_2}) is countermonotonic, for every $j_1, j_2 \in \{1, \dots, d\}$, with $j_1 \neq j_2$.

The distribution of a pairwise countermonotonic random vector in a Fréchet class $\mathcal{F}_d(F_1,\dots,F_d)$ is the lower Fréchet bound $F_d^L(x_1,\dots,x_d)=\max\{F_1(x_1)+\cdots+F_d(x_d)-d+1,0\}$. However, as discussed in [8], a Fréchet class $\mathcal{F}_d(F_1,\dots,F_d)$ admits a pairwise countermonotonic random vector only under very restrictive assumptions on the marginal distributions. These requirements were first studied in [34] and are reported in Proposition 3.2 in [8]. Within the framework of Bernoulli distributions, as discussed in Section 4.1 of [23], these conditions imply that a pairwise countermonotonic Bernoulli random vector has marginal means $p_1,\dots,p_d$ such that $p_1+\cdots+p_d\le 1$ or $p_1+\cdots+p_d\ge d-1$. Also, if $F_1,\dots,F_d$ are continuous distributions, then the Fréchet class $\mathcal{F}_d(F_1,\dots,F_d)$ does not admit any pairwise countermonotonic random vector. Therefore, the two Fréchet classes we focus on in this paper, i.e., $\mathcal{SB}_d$ and the Fréchet class of distributions with standard uniform marginals, do not admit a pairwise countermonotonic random vector, in any dimension $d>2$. For this reason, we turn our attention to different notions of extremal negative dependence that are based on less restrictive assumptions.

We consider three notions of extremal negative dependence: minimality in convex sums, joint mixability, and Σ-countermonotonicity. Minimality in convex sums consists in finding vectors $\boldsymbol{Y}$ such that $Y_1+\cdots+Y_d$ is minimal in the convex order within a given class of distributions. The convex order is a variability order; thus, a random variable that is minimal in the convex order is a minimal risk random variable. Therefore, the purpose of this extremal negative dependence is to minimize the aggregate risk. We formally introduce the convex order.

Definition 4. Given two random variables Y_1 and Y_2 with finite means, Y_1 is said to be smaller than Y_2 under the convex order (denoted $Y_1\le_{cx}Y_2$) if $\mathrm{E}\{\phi(Y_1)\}\le\mathrm{E}\{\phi(Y_2)\}$, for all real-valued convex functions ϕ for which the expectations are finite.

We now define a class of vectors whose sum is minimal in the convex order, and we call them Σ_{cx}-smallest elements.

Definition 5. A Σ_{cx}-smallest element in a class of distributions $\mathcal{F}$ is a random vector $\boldsymbol{Y}=(Y_1,\dots,Y_d)$ with distribution in $\mathcal{F}$ such that

$$\sum_{j=1}^{d}Y_j\le_{cx}\sum_{j=1}^{d}Y_j',$$

for any random vector $\boldsymbol{Y}'$ with distribution in $\mathcal{F}$.

Remark 2. A desirable property of extremal negative dependence is to minimize a dependence order. Indeed, a pairwise countermonotonic random vector $\boldsymbol{Y}$ is minimal in the supermodular order, i.e., it is such that $\mathrm{E}\{\psi(\boldsymbol{Y})\}\le\mathrm{E}\{\psi(\boldsymbol{Y}')\}$, for any random vector $\boldsymbol{Y}'$ with the same marginal distributions and for all supermodular functions ψ such that the expectations are finite. We recall that a supermodular function is a function $\psi:\mathbb{R}^d\to\mathbb{R}$ such that $\psi(\boldsymbol{x})+\psi(\boldsymbol{y})\le\psi(\boldsymbol{x}\wedge\boldsymbol{y})+\psi(\boldsymbol{x}\vee\boldsymbol{y})$, for all $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^d$, where $\wedge$ and $\vee$ denote the component-wise minimum and maximum operators, respectively. To define the class of Σ_{cx}-smallest elements, it is sufficient to consider the subclass of supermodular functions such that $\psi(\boldsymbol{x})=\phi(x_1+\cdots+x_d)$, for some convex function $\phi:\mathbb{R}\to\mathbb{R}$. Note that in general Σ_{cx}-smallest elements are not minimal in the supermodular order. If we restrict to exchangeable Bernoulli random vectors, we have a particular case, where Σ_{cx}-smallest elements are minimal in supermodular order, as proved in [35].

The next notion we present is closely related to the previous definition of Σ_{cx}-smallest elements. It is the joint mixability property and it has been introduced in [36].

Definition 6. A d-dimensional random vector $\boldsymbol{Y}=(Y_1,\dots,Y_d)$ is said to be a joint mix if

$$\Pr\left(\sum_{j=1}^{d}Y_j=k\right)=1,$$

for some $k\in\mathbb{R}$, called joint center.

Since any joint mix has variance of the sum of its components equal to 0, it is necessarily a Σ_{cx}-smallest element of its Fréchet class, under the assumption that the marginals have finite mean.

However, there exist Fréchet classes that do not admit Σ_{cx}-smallest elements or joint mixes; see Example 3.1 in [18]. Therefore, we conclude this section with the last notion we consider, the Σ-countermonotonicity property, introduced in [8]. This definition is significant because every Fréchet class admits a Σ-countermonotonic random vector.

Definition 7. A d-dimensional random vector $\boldsymbol{Y}=(Y_1,\dots,Y_d)$ is said to be Σ-countermonotonic if, for every subset $J\subseteq\{1,\dots,d\}$, the pair $(\sum_{j\in J}Y_j,\sum_{j\notin J}Y_j)$ is countermonotonic.

We use the convention $\sum_{j\in\emptyset}Y_j=0$. In [8], the authors show that, when admissible, pairwise countermonotonicity coincides with Σ-countermonotonicity. Moreover, they also show that if a Fréchet class admits a joint mix or a Σ_{cx}-smallest pmf, they are necessarily Σ-countermonotonic.

In Section 3.1, we develop the study of extremal negative dependence within the class $\mathcal{SB}_d$, while the discussions in the class of extremal mixture copulas and in the class of FGM copulas are presented in Section 3.2.1 and Section 3.2.2, respectively.

3.1. Symmetric Bernoulli distributions

Our main result is to completely characterize the class of Σ_{cx}-smallest elements in $\mathcal{SB}_d$. This also yields a complete characterization of Σ-countermonotonic random vectors in $\mathcal{SB}_d$. Indeed, Theorem 4.1 in [23] states that a Bernoulli random vector is Σ-countermonotonic if and only if it is a Σ_{cx}-smallest element in its Fréchet class. Since the characterization of Σ_{cx}-smallest elements is given in algebraic terms, this section mainly builds on Section 2; the proofs, based on the algebraic approach, are presented in Appendix B.

The problem of finding Σ_{cx}-smallest elements is trivial if we restrict the analysis to exchangeable Bernoulli random vectors with marginal mean p, for any $p \in (0,1)$. In this case, there is only one Σ_{cx}-smallest element in the class, and, as already mentioned in Remark 2, it is also minimal in the supermodular order. The general problem, even with a common marginal mean p, is still open. In [20], Theorem 5.2, based on the algebraic representation of multivariate pmfs of Bernoulli random vectors with common mean p, provides an algorithm to construct a non-exchangeable Σ_{cx}-smallest element in the class. If $p = 1/2$, we establish a stronger result by providing an explicit characterization of all such elements. Due to their technical nature, the proofs of this section are given in Appendix B.

For every $k \in \{0,1,\dots,d\}$, we denote by $\mathcal{X}_d^k$ the set of d-dimensional binary vectors for which the sum equals k, i.e., $\mathcal{X}_d^k = \{\boldsymbol{x} \in \mathcal{X}_d : x_1 + \cdots + x_d = k\}$. Then, we define the sets

$$\mathcal{X}_d^* = \begin{cases} \mathcal{X}_d^{d/2}, & \text{if } d \text{ is even,} \\ \mathcal{X}_d^{(d-1)/2} \cup \mathcal{X}_d^{(d+1)/2}, & \text{if } d \text{ is odd,} \end{cases} \qquad \mathscr{I}_{d-1}^* = \begin{cases} \mathcal{X}_{d-1}^{d/2}, & \text{if } d \text{ is even,} \\ \mathcal{X}_{d-1}^{(d-1)/2} \cup \mathcal{X}_{d-1}^{(d+1)/2}, & \text{if } d \text{ is odd,} \end{cases}$$

and we denote by M_d the largest integer smaller than or equal to $d/2$, i.e.,

$$M_d = \begin{cases} \frac{d}{2}, & \text{if } d \text{ is even,} \\ \frac{d-1}{2}, & \text{if } d \text{ is odd.} \end{cases}$$

The following proposition is a restatement of Proposition 5.2 in [20] and identifies the distribution of the sum of a Bernoulli random vector that is minimal in the convex order.

Proposition 6. *A Bernoulli random vector $\boldsymbol{X} \in \mathcal{SB}_d$ is a Σ_{cx}-smallest element in $\mathcal{SB}_d$ if and only if the sum $S_{\boldsymbol{X}} = X_1 + \cdots + X_d$ has distribution given by* $\Pr\{S_{\boldsymbol{X}} = (d-1)/2\} = \Pr\{S_{\boldsymbol{X}} = (d+1)/2\} = 1/2$ *if d is odd, and by* $\Pr(S_{\boldsymbol{X}} = d/2) = 1$ *if d is even.*

Proposition 7 characterizes the support of the Σ_{cx}-smallest pmfs in $\mathcal{SB}_d$. Since the proof is based on probabilistic arguments, we leave it in the main text.

Proposition 7. *A pmf $\boldsymbol{f} \in \mathcal{SB}_d$ is a Σ_{cx}-smallest element in $\mathcal{SB}_d$ if and only if the support of $\boldsymbol{f}$ is contained in $\mathcal{X}_d^*$.*

Proof. Consider a Bernoulli random vector $\boldsymbol{X}$ with pmf $\boldsymbol{f} \in \mathcal{SB}_d$. If $\boldsymbol{f}$ is a Σ_{cx}-smallest element in $\mathcal{SB}_d$, the distribution of the sum $S_{\boldsymbol{X}} = X_1 + \cdots + X_d$ is specified by Proposition 6, and we can easily prove that $f(\boldsymbol{x}) = 0$, for all $\boldsymbol{x} \notin \mathcal{X}_d^*$. Conversely, if the support of f is contained in $\mathcal{X}_d^*$, $S_{\boldsymbol{X}}$ is equal to $(d-1)/2$ or $(d+1)/2$ with probability one, when d is odd, and $S_{\boldsymbol{X}}$ is equal to $d/2$ with probability one, when d is even. Moreover, $\boldsymbol{X} \in \mathcal{SB}_d$ implies that $\mathrm{E}(S_{\boldsymbol{X}}) = d/2$. Therefore, the distribution of $S_{\boldsymbol{X}}$ is the distribution specified in Proposition 6, and we conclude that $\boldsymbol{f}$ is a Σ_{cx}-smallest element in $\mathcal{SB}_d$. $\square$

Remark 3. When d is odd and $\mathcal{X}_d^* = \mathcal{X}_d^{(d-1)/2} \cup \mathcal{X}_d^{(d+1)/2}$, if the support of $\boldsymbol{f} \in \mathcal{SB}_d$ is contained in $\mathcal{X}_d^*$, Proposition 6 imply that there exist $\boldsymbol{x}_1 \in \mathcal{X}_d^{(d-1)/2}$ and $\boldsymbol{x}_2 \in \mathcal{X}_d^{(d+1)/2}$, such that $f(\boldsymbol{x}_1) > 0$ and $f(\boldsymbol{x}_2) > 0$.

It is worth noting that there does not exist any joint mix in $\mathcal{SB}_d$ when d is odd. When d is even, instead, a Bernoulli random vector is a Σ_{cx}-smallest element in $\mathcal{SB}_d$ if and only if it is a joint mix; in this case, the definitions of Σ-countermonotonic random vector, Σ_{cx}-smallest element, and joint mix coincide in $\mathcal{SB}_d$. Therefore, building on Proposition 7, we can characterize extremal negative dependence by considering the class of Σ_{cx}-smallest elements. We first identify the Σ_{cx}-smallest palindromic pmfs.

Proposition 8. *A pmf $\boldsymbol{f}^{K*} \in \mathcal{PB}_d$ is a Σ_{cx}-smallest element in $\mathcal{SB}_d$ if and only if it is a convex linear combination of the Σ_{cx}-smallest elements of the basis of the kernel $\mathcal{B}_K$ in (4).*

We denote by $\mathcal{PB}_d^* \subseteq \mathcal{PB}_d$ the set of Σ_{cx}-smallest palindromic Bernoulli distributions. Proposition 8 states that $\mathcal{PB}_d^*$ is a convex polytope whose extremal points are the Σ_{cx}-smallest pmfs in $\mathcal{B}_K$, which are easy to identify. Indeed, they have support on exactly two points, $\boldsymbol{x}$ and $\boldsymbol{1}_d - \boldsymbol{x}$, where $\boldsymbol{x}$ is such that $x_1 + \cdots + x_d = M_d$. We now consider the entire class $\mathcal{SB}_d$. The following theorem characterizes the coefficients of the polynomials corresponding to the Σ_{cx}-smallest pmfs of the class $\mathcal{SB}_d$.

Theorem 1. *Let $\boldsymbol{f}^* \in \mathcal{SB}_d$ be a Σ_{cx}-smallest element in $\mathcal{SB}_d$. Then, the coefficients of the polynomial $\mathcal{H}(\boldsymbol{f}^*) = P^*(\boldsymbol{z}) = \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}} a_{\boldsymbol{i}} \boldsymbol{z}^{\boldsymbol{i}} \in \mathcal{I}_{\mathcal{P}}$ are such that:*

(i) *$a_{\boldsymbol{i}} = 0$, for every $\boldsymbol{i} \notin \mathscr{I}_{d-1}^*$,*

(ii) *the sum of the coefficients of the monomials of the same order is equal to 0, i.e.,*

$$\sum_{i \in \mathcal{X}_{d-1}^k} a_i = 0, \quad \textit{for every } k \in \{0, \dots, d-1\},$$

(iii) *for each* $j \in \{1, \dots, d-1\}$*, the sum of the coefficients* a_i *involving* z_j *is zero, i.e.,*

$$\sum_{i \in \mathcal{X}_{d-1} : i_j = 1} a_i = 0, \quad \textit{for each } j \in \{1, \dots, d-1\}.$$

From Theorem 1, all the polynomials $P^*(\boldsymbol{z})$ with a Σ_{cx}-smallest pmf in their counter-image $\mathcal{H}^{-1}[P^*(\boldsymbol{z})]$ are of the form

$$P^*(\boldsymbol{z}) = \sum_{\boldsymbol{i} \in \mathscr{I}_{d-1}^*} a_{\boldsymbol{i}} \boldsymbol{z}^{\boldsymbol{i}}, \tag{5}$$

where the coefficients $a_{\boldsymbol{i}}$, for $\boldsymbol{i} \in \mathscr{I}_{d-1}^*$, verify the conditions (ii) and (iii) of the above theorem. The next corollary to Theorem 1 states that the coefficients of the polynomials of the Σ_{cx}-smallest pmfs in the class $\mathcal{SB}_d$ are the solutions of a homogeneous linear system. The number n_d^* vectors in $\mathscr{I}_{d-1}^*$ is given by

$$n_d^* = \begin{cases} \binom{d-1}{M_d} + \binom{d-1}{M_d+1}, & \text{if } d \text{ is odd,} \\ \binom{d-1}{M_d}, & \text{if } d \text{ is even.} \end{cases}$$

Corollary 1. *Let* $\boldsymbol{f} \in \mathcal{SB}_d$ *be a* Σ_{cx}*-smallest element in* $\mathcal{SB}_d$*. Then, the coefficients of the polynomial* $\mathcal{H}(\boldsymbol{f}) = P(\boldsymbol{z}) = \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}} a_{\boldsymbol{i}} \boldsymbol{z}^{\boldsymbol{i}} \in \mathcal{I}_{\mathcal{P}}$ *are the solutions of*

$$A_d \boldsymbol{a} = \boldsymbol{0}, \tag{6}$$

where $\boldsymbol{a} = (a_{\boldsymbol{i}} : \boldsymbol{i} \in \mathscr{I}_{d-1}^*)$ *and* A_d *is obtained from the matrix* $A_{\mathscr{I}_{d-1}^*} = (\boldsymbol{i} : \boldsymbol{i} \in \mathscr{I}_{d-1}^*) \in \mathcal{M}((d-1) \times n_d^*)$*, whose columns are the vectors* $\boldsymbol{i} \in \mathscr{I}_{d-1}^*$*. In particular,*

- *if* d *is even,* $A_d = (\boldsymbol{1}_{n_d^*}^\top // A_{\mathscr{I}_{d-1}^*}) \in \mathcal{M}(d \times n_d^*)$*;*
- *if* d *is odd,* $A_d = (R_1 // R_2 // A_{\mathscr{I}_{d-1}^*}) \in \mathcal{M}((d+1) \times n_d^*)$*, where* $R_1 \in \mathcal{M}(1 \times n_d^*)$ *is a row vector with 1s in correspondence of the indexes* $\boldsymbol{i}$ *with sum* M_d *and zeros elsewhere, and* $R_2 \in \mathcal{M}(1 \times n_d^*)$ *is a row vector with 1s in correspondence of the indexes* $\boldsymbol{i}$ *with sum* $M_d + 1$ *and zeros elsewhere.*

The following two examples characterize the polynomials of Σ_{cx}-smallest pmfs in dimensions $d = 3$ and $d = 4$, respectively.

Example 1. We consider $d = 3$. Since d is odd, we have $M_d = (d-1)/2 = 1$ and $\mathscr{I}_2^* = \{(1,0),(0,1),(1,1)\}$. For $(i_1, i_2) \in \mathscr{I}_2^*$, the first row of A_3 is equal to 1 if $i_1 + i_2 = M_d = 1$ and 0 otherwise, and the second row is the opposite, viz.

$$A_3 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

The matrix $A_3 \in \mathcal{M}(4 \times 3)$ has rank 3. Consequently, the linear system in (6) admits only the trivial solution $a_{10} = a_{01} = a_{11} = 0$, and all Σ_{cx}-smallest pmfs in $\mathcal{SB}_3$ have null polynomials.

Example 2. We consider $d = 4$. Since d is even, we have $M_d = d/2 = 2$ and $\mathscr{I}_3^* = \{(1,1,0),(1,0,1),(0,1,1)\}$. The first row of A_4 is a vector of all 1s, viz.

$$A_4 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

The matrix $A_4 \in \mathcal{M}(4 \times 3)$ has rank 3. Consequently, the linear system in (6) admits only the trivial solution $a_{110} = a_{101} = a_{011} = 0$, and all Σ_{cx}-smallest pmfs in $\mathcal{SB}_4$ have null polynomials.

Remark 4. As shown in Examples 1 and 2, in the cases $d = 3$ and $d = 4$, all Σ_{cx}-smallest pmfs have null polynomial, i.e., $\mathcal{H}(\boldsymbol{f}) = 0$, if $\boldsymbol{f}$ is Σ_{cx}-smallest. Therefore, the set of Σ_{cx}-smallest pmfs is included in $\mathcal{K}(\mathcal{H})$. Thus, for $d \le 4$, both the Σ_{cx}-smallest pmfs and the "Σ_{cx}-maximal" pmf (the upper Fréchet bound) are palindromic Bernoulli distributions.

Corollary 1 states that the coefficients of the polynomials of all Σ_{cx}-smallest pmfs of $\mathcal{SB}_d$ are solutions of the homogeneous linear system in (6). However, there exist pmfs in $\mathcal{SB}_d$ that are not Σ_{cx}-smallest elements yet generate a polynomial of the form in (5). This is a consequence of the fact that the map $\mathcal{H}$ is not injective. For example, let $f^* \in \mathcal{SB}_d$ be a Σ_{cx}-smallest Bernoulli pmf, and denote by $P^*(\boldsymbol{z})$ its corresponding polynomial. We know that $P^*(\boldsymbol{z})$ verifies the three properties of Theorem 1. Consider also $g=(f^*+\bar{f})/2$, where $\bar{f}(0,\dots,0)=\bar{f}(1,\dots,1)=1/2$. Since $\mathcal{H}(\bar{f})\equiv 0$, it follows by linearity of the map $\mathcal{H}$ that $\mathcal{H}(g)=P^*(\boldsymbol{z})$, although g is not a Σ_{cx}-smallest element in $\mathcal{SB}_d$. Theorem 2 states the key result of this section, because it completely characterizes the class of Σ_{cx}-smallest pmfs.

Theorem 2. *Let $P^*(\boldsymbol{z})=\sum_{\boldsymbol{i}\in\mathcal{X}_{d-1}} a_{\boldsymbol{i}}\boldsymbol{z}^{\boldsymbol{i}}\in\mathbb{Q}[z_1,\dots,z_{d-1}]$ be a non-null polynomial that verifies the three properties of Theorem 1. Then, $P^*(\boldsymbol{z})\in\mathcal{I}_{\mathcal{P}}$ and the type-0 pmf $\boldsymbol{f}^*$ corresponding to $P^*(\boldsymbol{z})$ is a Σ_{cx}-smallest pmf of $\mathcal{SB}_d$. Moreover, the family*

$$\left\{\boldsymbol{f}=\lambda\boldsymbol{f}^*+(1-\lambda)\boldsymbol{f}^{K*}:\boldsymbol{f}^{K*}\in\mathcal{PB}_d^*,\ \lambda\in(0,1]\cap\mathbb{Q}\right\},$$

constitutes the set of all Σ_{cx}-smallest pmfs corresponding to polynomials equivalent to $P^*(\boldsymbol{z})$.

The algebraic representation and the proofs of the main results are technical; however, their strength lies in their simplicity of use. In what follows, we illustrate how to apply these results to find a Σ_{cx}-smallest element in $\mathcal{SB}_d$ and, at least in principle, how they allow us to determine all Σ_{cx}-smallest elements. To find a Σ_{cx}-smallest element in the class $\mathcal{SB}_d$, we proceed as follows:

1. Choose a polynomial $P(\boldsymbol{z})=\sum_{\boldsymbol{i}\in\mathcal{X}_{d-1}} a_{\boldsymbol{i}}\boldsymbol{z}^{\boldsymbol{i}}$, with coefficients that satisfies the conditions in Theorem 1.
2. Apply Algorithm 1 to find the type-0 pmf $\boldsymbol{f}^*$.

The type-0 pmf $\boldsymbol{f}^*$ is a Σ_{cx}-smallest element in $\mathcal{SB}_d$. To find all Σ_{cx}-smallest elements, we rely on Corollary 1, which provides a method to determine the coefficients of all polynomials satisfying the conditions of Theorem 1. Then, for each polynomial, we construct the type-0 pmf $\boldsymbol{f}^*$. By Theorem 2, all Σ_{cx}-smallest pmfs can be written as

$$\boldsymbol{f}=\lambda\boldsymbol{f}^*+(1-\lambda)\boldsymbol{f}^{K*},\quad \lambda\in(0,1],$$

where $\boldsymbol{f}^{K*}$ is a convex combination of pmfs in $\mathcal{B}_K$ with support in $\mathcal{X}_d^*$. We recall that this is equivalent to finding all Σ-countermonotonic elements in $\mathcal{SB}_d$, which also satisfy the joint mixability property, when d is even.

We conclude this section with Examples 3 and 4, which characterize the Σ_{cx}-smallest elements of the classes $\mathcal{SB}_5$ and $\mathcal{SB}_6$, respectively.

Example 3. In this example, we show how to find all Σ_{cx}-smallest elements in the class $\mathcal{SB}_5$. By Corollary 1, we build the matrix

$$A_5=\begin{pmatrix}1&1&1&0&1&1&0&1&0&0\\0&0&0&1&0&0&1&0&1&1\\1&1&0&1&1&0&1&0&1&0\\1&0&1&1&0&1&1&0&0&1\\0&1&1&1&0&0&0&1&1&1\\0&0&0&0&1&1&1&1&1&1\end{pmatrix},$$

that is given by $A_5=(R_1//R_2//A_{I_4^*})$. Since $\mathrm{rank}(A_5)=5$, the solution space of the system in (6) has dimension $n_5^*-\mathrm{rank}(A_5)=5$. A basis of the space of the solutions of $A_5\boldsymbol{a}=\boldsymbol{0}$ is

$$\begin{aligned}\mathcal{A}=\big\{&\boldsymbol{a}^{(1)}=(0,1,-1,0,-1,1,0,0,0,0),\boldsymbol{a}^{(2)}=(0,1,0,-1,-1,0,1,0,0,0),\boldsymbol{a}^{(3)}=(1,0,-1,0,-1,0,0,1,0,0),\\&\boldsymbol{a}^{(4)}=(1,0,0,-1,-1,0,0,0,1,0),\boldsymbol{a}^{(5)}=(1,1,-1,-1,-1,0,0,0,0,1)\big\}.\end{aligned}$$

Thus, every polynomial whose coefficients

$$\boldsymbol{a}=(a_{1100},a_{1010},a_{0110},a_{1110},a_{1001},a_{0101},a_{1101},a_{0011},a_{1011},a_{0111})$$

are a linear combination of the vectors of the basis $\mathcal{A}$ verify the three assumptions of Theorem 1 and its type-0 pmf is a Σ_{cx}-smallest element in $\mathcal{SB}_5$. For example, the polynomial corresponding to the vector $\boldsymbol{a}^{(1)}$ is $P_1(\boldsymbol{z})=z_1z_3-z_2z_3-z_1z_4+z_2z_4$ and the corresponding type-0 pmf $\boldsymbol{f}^{(1)}$ is such that $f^{(1)}((1,0,1,0,0))=f^{(1)}((1,0,0,1,1))=f^{(1)}((0,1,1,0,1))=f^{(1)}((0,1,0,1,0))=1/4$ and it is zero elsewhere. Following Lemma 2.3 in [37], which gives the conditions for a pmf to be an extremal point, it can be proved that $\boldsymbol{f}^{(1)}$ is an extremal pmf of the polytope $\mathcal{SB}_5$. A general Σ_{cx}-smallest pmf in $\mathcal{H}^{-1}[P_1(\boldsymbol{z})]$ can be found as $\boldsymbol{f}=\lambda\boldsymbol{f}^{(1)}+(1-\lambda)\boldsymbol{f}^{K*}$, where $\boldsymbol{f}^{K*}\in\mathcal{PB}_5^*$.

Example 4. In this example, we show how to find all Σ_{cx}-smallest elements of the class $\mathcal{SB}_6$. By Corollary 1, we build the matrix $A_6=(\boldsymbol{1}_{10}^{\top}//A_{I_5^*})$. We find that the basis $\mathcal{A}$ of the solution space in Example 3 is also a basis of the space of solutions of the system $A_6\boldsymbol{a}=\boldsymbol{0}$. Thus, every polynomial that have a Σ_{cx}-smallest pmf in its counter-image has coefficients

$$\boldsymbol{a}=(a_{11100},a_{11010},a_{10110},a_{01110},a_{11001},a_{10101},a_{01101},a_{10011},a_{01011},a_{00111})$$

that are a linear combination of $\boldsymbol{a}^{(1)},\boldsymbol{a}^{(2)},\boldsymbol{a}^{(3)},\boldsymbol{a}^{(4)}$, and $\boldsymbol{a}^{(5)}$. For example, the polynomial with coefficients $\boldsymbol{a}^{(1)}$ is $P_1(\boldsymbol{z})=z_1z_2z_4-z_1z_3z_4-z_1z_2z_5+z_1z_3z_5$ and the corresponding type-0 pmf $\boldsymbol{f}^{(1)}$ is such that $f^{(1)}((1,1,0,1,0,0))=f^{(1)}((0,1,0,0,1,1))=f^{(1)}((0,0,1,1,0,1))=f^{(1)}((1,0,1,0,1,0))=1/4$ and it is zero elsewhere. As in Example 3, it can be proved that $\boldsymbol{f}^{(1)}$ is an extremal pmf of the polytope $\mathcal{SB}_6$. Finally, we consider the linear combination $\tilde{\boldsymbol{a}}=\boldsymbol{a}^{(1)}-\boldsymbol{a}^{(2)}-\boldsymbol{a}^{(3)}+\boldsymbol{a}^{(4)}$. The resulting polynomial is $\tilde{P}(\boldsymbol{z})=P_1(\boldsymbol{z})-P_2(\boldsymbol{z})-P_3(\boldsymbol{z})+P_4(\boldsymbol{z})=z_1z_3z_5-z_2z_3z_5-z_1z_4z_5+z_2z_4z_5$ and its type-0 pmf $\tilde{\boldsymbol{f}}$ is such that $\tilde{f}((1,0,1,0,1,0))=\tilde{f}((1,0,0,1,0,1))=\tilde{f}((0,1,1,0,0,1))=\tilde{f}((0,1,0,1,1,0))=1/4$ and zero elsewhere. It can be proved that also $\tilde{\boldsymbol{f}}$ is an extremal point of $\mathcal{SB}_6$.

3.2. Extremal negative dependent copulas

We now turn to the study of copulas. We recall that a d-dimensional copula is the cumulative distribution function (cdf) on $[0,1]^d$ of a d-dimensional random vector $\boldsymbol{U}$ whose univariate marginals are uniformly distributed on $[0,1]$. Two classes of copulas can be constructed from symmetric Bernoulli distributions: extremal mixture copulas and Farlie–Gumbel–Morgenstern (FGM) copulas. Both of these classes inherit some dependence properties from $\mathcal{SB}_d$. In particular, the results on extremal negative dependence within the Bernoulli class allow us to find a sub-family of the class of extremal copulas that are Σ-countermonotonic, i.e., they represent extremal negative dependence in the entire class of copulas, and a class of copulas with the joint mixability property.

3.2.1. Extremal mixture copulas

In this section, we study the class of extremal mixture copulas. These copulas are in a one-to-one correspondence with the palindromic Bernoulli distributions (see [27]) that coincide with the kernel of the map $\mathcal{H}$ (Proposition 1).

Definition 8. Given a standard uniform random variable U, an extremal copula with index set $J \subseteq \{1,\dots,d\}$ is the distribution function of the d-dimensional random vector $\boldsymbol{V} = (V_1,\dots,V_d)$ where $V_j \overset{d}{=} U$ if $j \in J$, and $V_j \overset{d}{=} 1-U$ if $j \notin J$, for every $j \in \{1,\dots,d\}$.

There exist 2^{d-1} different d-dimensional extremal copulas. For $\boldsymbol{i} \in \mathcal{X}_{d-1}$, let $J_{\boldsymbol{i}} = \{j \in \{1,\dots,d-1\} : \boldsymbol{i}_j = 1\}$ be the set of indexes corresponding to 1s in $\boldsymbol{i}$. It is possible to infer the explicit form of the copulas, i.e., for every $\boldsymbol{i} \in \mathcal{X}_{d-1}$, one has

$$C_{\boldsymbol{i}}(\boldsymbol{u}) = (\min_{j\in J_{\boldsymbol{i}}} u_j + \min_{j\in \bar{J}_{\boldsymbol{i}}} u_j - 1)^+, \quad \boldsymbol{u} \in [0,1]^d,$$

where $\bar{J}_{\boldsymbol{i}} = \{1,\dots,d\} \setminus J_{\boldsymbol{i}}$ and $y^+ = \max(0,y)$; we use the convention $\min_{j\in\emptyset} u_j = 1$.

One may consider a wider class of copulas by taking mixtures of extremal copulas; see [28].

Definition 9. An extremal mixture copula C is a copula of the form

$$C = \sum_{\boldsymbol{i}\in\mathcal{X}_{d-1}} w_{\boldsymbol{i}} C_{\boldsymbol{i}},$$

where, for every $\boldsymbol{i} \in \mathcal{X}_{d-1}$, $C_{\boldsymbol{i}}$ is the extremal copula with index set $J_{\boldsymbol{i}}$ and the weights $w_{\boldsymbol{i}}$ are such that $w_{\boldsymbol{i}} \geq 0$, for every $\boldsymbol{i} \in \mathcal{X}_{d-1}$, and $\sum_{\boldsymbol{i}\in\mathcal{X}_{d-1}} w_{\boldsymbol{i}} = 1$.

We denote by $\mathcal{C}_d^{\mathrm{EM}}$ the class of extremal mixture copulas. The following proposition has been proved in [27] and states that there exists a non-injective map between the class of multivariate Bernoulli distributions and the class of extremal mixture copulas.

Proposition 9. *Let U be a standard uniform random variable and $\boldsymbol{X}$ a d-dimensional multivariate Bernoulli random vector with pmf f. Let $\boldsymbol{X}$ and U be independent. Then the cdf of the uniform random vector*

$$\boldsymbol{V} = U\boldsymbol{X} + (1-U)(\mathbf{1}_d - \boldsymbol{X}) \tag{7}$$

is an extremal mixture copula with weights given, for each $\boldsymbol{i} \in \mathcal{X}_{d-1}$, by

$$w_{\boldsymbol{i}} = f(\boldsymbol{s}_{\boldsymbol{i}}) + f(\mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}}), \tag{8}$$

where $\boldsymbol{s}_{\boldsymbol{i}} = (\boldsymbol{i}//0)$.

Given an extremal mixture copula with weights $w_{\boldsymbol{i}}$, for every $\boldsymbol{i} \in \mathcal{X}_{d-1}$, there exist infinitely many Bernoulli distributions satisfying (8). However, it is possible to identify a unique Bernoulli distribution by considering the class of palindromic Bernoulli distributions $\mathcal{PB}_d$, characterized by the constraint $f(\boldsymbol{s}_{\boldsymbol{i}}) = f(\mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}})$, for every $\boldsymbol{i} \in \mathcal{X}_{d-1}$. Therefore, the class $\mathcal{PB}_d$ is in a one-to-one correspondence with the family of extremal mixture copulas $\mathcal{C}_d^{\mathrm{EM}}$, see [27]:

$$\mathcal{PB}_d \longleftrightarrow \mathcal{C}_d^{\mathrm{EM}}. \tag{9}$$

In particular, the extremal copulas correspond to the pmfs of the basis $\mathcal{B}_K$ in (4).

The results of Section 3.1 are useful to explore the concept of negative dependence in the class of extremal mixture copulas. We conclude this section with three results within the class $\mathcal{C}_d^{\mathrm{EM}}$.

Proposition 10. *Let $\boldsymbol{X}, \boldsymbol{X}' \in \mathcal{SB}_d$ and let $\boldsymbol{V}$ and $\boldsymbol{V}'$ be the uniform random vectors with extremal mixture copulas constructed from $\boldsymbol{X}$ and $\boldsymbol{X}'$, respectively, via the stochastic representation in (7). Then,*

$$\sum_{j=1}^d X_j \leq_{cx} \sum_{j=1}^d X'_j \iff \sum_{j=1}^d V_j \leq_{cx} \sum_{j=1}^d V'_j.$$

Proof. Given a Bernoulli random vector $\boldsymbol{X} \in \mathcal{SB}_d$, from the stochastic representation in (7), we have

$$\sum_{j=1}^d V_j \Big| (U=u) = \sum_{j=1}^d uX_j + (1-u)(1-X_j) = d(1-u) + (2u-1)\sum_{j=1}^d X_j.$$

Since $f(ax+b)$, with $a,b\in\mathbb{R}$, is convex if f is a convex function, it follows that

$$\sum_{j=1}^{d} X_j \leq_{\mathrm{cx}} \sum_{j=1}^{d} X'_j \implies \sum_{j=1}^{d} V_j|(U=u) \leq_{\mathrm{cx}} \sum_{j=1}^{d} V'_j|(U=u) \implies \sum_{j=1}^{d} V_j \leq_{\mathrm{cx}} \sum_{j=1}^{d} V'_j,$$

where the last implication follows from Theorem 3.A.12(b) in [38]. The converse follows from (7) by observing that

$$\sum_{j=1}^{d} X_j \Big|(U=u) = \frac{d(1-u)+\sum_{j=1}^{d} V_j}{(2u-1)},$$

and applying the same arguments. □

Proposition 10 implies that if $\boldsymbol{X}\in\mathcal{SB}_d$ is a Σ_{cx}-smallest element, then $\boldsymbol{V}$ in (7) is a Σ_{cx}-smallest element in $\mathcal{C}_d^{\mathrm{EM}}$. An important consequence of the one-to-one map in (9) is that we can construct all the extremal mixture copulas from $\mathcal{PB}_d$. Therefore, the Σ_{cx}-smallest elements in $\mathcal{C}_d^{\mathrm{EM}}$ can be constructed from Σ_{cx}-smallest palindromic Bernoulli random vectors. We recall that Proposition 8 identifies all Σ_{cx}-smallest pmfs in $\mathcal{PB}_d$.

The last two results of this section characterize the extremal negative dependence in the entire class of copulas, not only in $\mathcal{C}_d^{\mathrm{EM}}$. Proposition 11 states that the extremal copulas built from Σ_{cx}-smallest Bernoulli random vectors in $\mathcal{B}_K$ are Σ-countermonotonic, but, in general, are not Σ_{cx}-smallest in the entire Fréchet class of copulas. Proposition 12, instead, implies that, when d is even, the extremal mixture copulas constructed from Σ_{cx}-smallest pmfs of $\mathcal{PB}_d$ have the joint mixability property, hence they are Σ_{cx}-smallest elements in the entire class of copulas.

Proposition 11. *Let $\boldsymbol{X}$ be a Bernoulli random vector with pmf $\boldsymbol{f}\in\mathcal{B}_K$, where $\mathcal{B}_K$ is the basis of the kernel of $\mathcal{H}$, given in (4). If $\boldsymbol{X}$ is a Σ_{cx}-smallest element in $\mathcal{SB}_d$, then the uniform random vector $\boldsymbol{V}$ from (7) is a Σ_{cx}-smallest in $\mathcal{C}_d^{EM}$ and Σ-countermonotonic.*

Proof. We assume d odd, as the case d even is analogous. Since $\boldsymbol{X}$ is a Σ_{cx}-smallest element in $\mathcal{SB}_d$ (and also in $\mathcal{PB}_d$), it follows from Proposition 10 that the random vector $\boldsymbol{V}$ is a Σ_{cx}-smallest element in $\mathcal{C}_d^{\mathrm{EM}}$. We now prove that $\boldsymbol{V}$ is Σ-countermonotonic. Let $J\subset\{1,\dots,d\}$ be a set of indexes, $J\neq\emptyset$. Since $\boldsymbol{f}\in\mathcal{B}_K$, there exists $\boldsymbol{x}\in\mathcal{X}_d$ such that the random vector $\boldsymbol{X}$ can only take the values $\boldsymbol{x}$ or $\mathbf{1}_d-\boldsymbol{x}$. Moreover, since $\boldsymbol{X}$ is a Σ_{cx}-smallest element, Proposition 7 implies that $\boldsymbol{x}\in\mathcal{X}_d^*$. Recall that the sum of the components of $\boldsymbol{x}\in\mathcal{X}_d^*$ is equal to $M_d=(d-1)/2$ or M_d+1. Therefore, we have two options: either $\sum_{j=1}^d x_j=M_d$ and $\sum_{j=1}^d(1-x_j)=M_d+1$, or $\sum_{j=1}^d x_j=M_d+1$ and $\sum_{j=1}^d(1-x_j)=M_d$. We consider the case $\sum_{j=1}^d x_j=M_d$; the other case follows by taking $\boldsymbol{y}=\mathbf{1}_d-\boldsymbol{x}$. Let $k:=\sum_{j\in J}x_j$. We have

$$\sum_{j\in\bar{J}} x_j=M_d-k,\quad \sum_{j\in J}(1-x_j)=|J|-k,\quad \sum_{j\in\bar{J}}(1-x_j)=M_d+1-(|J|-k),$$

where $\bar{J}$ denotes the complement of J and $|J|$ its cardinality. Let us define two random variables

$$A=\sum_{j\in J}V_j=U\sum_{j\in J}X_j+(1-U)\sum_{j\in J}(1-X_j),$$
$$B=\sum_{j\in\bar{J}}V_j=U\sum_{j\in\bar{J}}X_j+(1-U)\sum_{j\in\bar{J}}(1-X_j).$$

By conditioning on the two possible outcomes of the random variable $\boldsymbol{X}$, we have

$$A|(\boldsymbol{X}=\boldsymbol{x})=|J|-k+(2k-|J|)U,\quad B|(\boldsymbol{X}=\boldsymbol{x})=M_d+1-(|J|-k)-(1+2k-|J|)U,$$
$$A|(\boldsymbol{X}=\mathbf{1}_d-\boldsymbol{x})=k-(2k-|J|)U,\quad B|(\boldsymbol{X}=\mathbf{1}_d-\boldsymbol{x})=M_d-k+(1+2k-|J|)U.$$

Let (A_1,B_1) and (A_2,B_2) be two independent copies of (A,B). We have, for $h\in\{1,2\}$,

$$A_h=U_h\sum_{j\in J}X_j^{(h)}+(1-U_h)\sum_{j\in J}(1-X_j^{(h)}),\quad B_h=U_h\sum_{j\in\bar{J}}X_j^{(h)}+(1-U_h)\sum_{j\in\bar{J}}(1-X_j^{(h)}),$$

where U_1 and U_2 are independent standard uniform random variables, and $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$ are iid copies of $\boldsymbol{X}$, independent of U_1 and U_2. We want to prove that (A,B) is countermonotonic, viz.

$$\Pr\{(A_1-A_2)(B_1-B_2)\leq 0\}=\sum_{\boldsymbol{x}_1}\sum_{\boldsymbol{x}_2}\Pr\left\{(A_1-A_2)(B_1-B_2)\leq 0|\boldsymbol{X}^{(1)}=\boldsymbol{x}_1,\boldsymbol{X}^{(2)}=\boldsymbol{x}_2\right\}\Pr(\boldsymbol{X}^{(1)}=\boldsymbol{x}_1)\Pr(\boldsymbol{X}^{(2)}=\boldsymbol{x}_2). \tag{10}$$

The pair $(\boldsymbol{x}_1,\boldsymbol{x}_2)$ can take one of four possible values: $(\boldsymbol{x},\boldsymbol{x}),(\boldsymbol{x},\mathbf{1}_d-\boldsymbol{x}),(\mathbf{1}_d-\boldsymbol{x},\boldsymbol{x})$, or $(\mathbf{1}_d-\boldsymbol{x},\mathbf{1}_d-\boldsymbol{x})$.

Case 1. Let $(\boldsymbol{x}_1,\boldsymbol{x}_2)=(\boldsymbol{x},\boldsymbol{x})$. We have

$$\begin{aligned}\Pr\{(A_1-A_2)(B_1-B_2)\leq 0|\boldsymbol{X}^{(1)}=\boldsymbol{x},\boldsymbol{X}^{(2)}=\boldsymbol{x}\}&=\Pr\{(2k-|J|)(U_1-U_2)\cdot(-1)(1+2k-|J|)(U_1-U_2)\leq 0\}\\&=\Pr\{-(2k-|J|)(1+2k-|J|)(U_1-U_2)^2\leq 0\}=1,\end{aligned}$$

where the last equality follows because $2k-|J|\in\mathbb{Z}$ and, if $2k-|J|\geq 0$ then $1+2k-|J|>0$, while if $2k-|J|<0$ then $1+2k-|J|\leq 0$.

Case 2. Let $(\boldsymbol{x}_1,\boldsymbol{x}_2)=(\boldsymbol{x},\mathbf{1}_d-\boldsymbol{x})$. We have

$$\Pr\{(A_1-A_2)(B_1-B_2)\leq 0|\boldsymbol{X}^{(1)}=\boldsymbol{x},\boldsymbol{X}^{(2)}=\mathbf{1}_d-\boldsymbol{x}\}=\Pr\{-(2k-|J|)(1+2k-|J|)(U_1+U_2-1)^2\leq 0\}=1,$$

for the same argument in Case 1.

Cases 3 and 4 are analogous. From (10), we have

$$\Pr\{(A_1 - A_2)(B_1 - B_2) \leq 0\} = \sum_{\boldsymbol{x}_1} \sum_{\boldsymbol{x}_2} 1 \cdot \Pr(\boldsymbol{X}^{(1)} = \boldsymbol{x}_1)\Pr(\boldsymbol{X}^{(2)} = \boldsymbol{x}_2) = 1.$$

Therefore, $\sum_{j\in J} V_j$ and $\sum_{j\in \bar{J}} V_j$ are countermonotonic, and $\boldsymbol{V}$ is Σ-countermonotonic. □

The following example shows that Proposition 11 holds only for extremal copulas and not for extremal mixture copulas.

Example 5. Let $d = 5$. Consider $\boldsymbol{X} \in \mathcal{PB}_5$ such that $\Pr(\boldsymbol{X} = \boldsymbol{x}_1) = \Pr(\boldsymbol{X} = \boldsymbol{x}_2) = \Pr(\boldsymbol{X} = \boldsymbol{1}_5 - \boldsymbol{x}_1) = \Pr(\boldsymbol{X} = \boldsymbol{1}_5 - \boldsymbol{x}_2) = 1/4$, with $\boldsymbol{x}_1 = (1,0,0,0,1)$ and $\boldsymbol{x}_2 = (0,0,0,1,1)$. Since $\boldsymbol{X}$ has support in $\mathcal{X}_5^*$, by Proposition 7, $\boldsymbol{X}$ is a Σ_{cx}-smallest element in $\mathcal{SB}_5$. Let $\boldsymbol{V}$ be the uniform random vector with extremal mixture copula obtained from (7), and define $A = \sum_{j\in J} V_j$ and $B = \sum_{j\in \bar{J}} V_j$, with $J = \{1,5\}$. Let (A_1, B_1) and (A_2, B_2) be two independent copies of (A, B). We have, for $h \in \{1,2\}$,

$$A_h = U_h\left(X_1^{(h)} + X_5^{(h)}\right) + (1-U_h)\left(2 - X_1^{(h)} - X_5^{(h)}\right),$$
$$B_h = U_h\left(X_2^{(h)} + X_3^{(h)} + X_4^{(h)}\right) + (1-U_h)\left(3 - X_2^{(h)} - X_3^{(h)} - X_4^{(h)}\right),$$

where U_1 and U_2 are independent standard uniform random variables, and $\boldsymbol{X}^{(1)}$ and $\boldsymbol{X}^{(2)}$ are iid copies of $\boldsymbol{X}$, independent of U_1 and U_2. It holds that,

$$\Pr\{(A_1 - A_2)(B_1 - B_2) \leq 0 | \boldsymbol{X}^{(1)} = \boldsymbol{x}_1, \boldsymbol{X}^{(2)} = \boldsymbol{x}_2\} = \Pr\{(2U_1 - 1)(1 - 3U_1 + U_2) \leq 0\} < 1.$$

Therefore, $\Pr\{(A_1 - A_2)(B_1 - B_2) \leq 0\} < 1$ and $\boldsymbol{V}$ is not Σ-countermonotonic.

Proposition 12. *Let d be even. If $\boldsymbol{X}^* \in \mathcal{PB}_d$ is a Σ_{cx}-smallest element in $\mathcal{SB}_d$, then $\boldsymbol{V}^* = U\boldsymbol{X}^* + (1-U)(\boldsymbol{1} - \boldsymbol{X}^*)$ is a joint mix.*

Proof. By Proposition 6, the sum $X_1^* + \cdots + X_d^*$ takes value $d/2$ with probability 1. We have

$$\Pr\left(\sum_{j=1}^d V_j^* = \frac{d}{2}\right) = \Pr\left\{d(1-U) + (2U-1)\sum_{j=1}^d X_j^* = \frac{d}{2}\right\} = \int_0^1 \Pr\left\{\sum_{j=1}^d X_j^* = \frac{d/2 - d(1-u)}{2u-1}\right\} \mathrm{d}u$$
$$= \int_0^1 \Pr\left(\sum_{j=1}^d X_j^* = \frac{d}{2}\right) \mathrm{d}u = 1.$$
□

We conclude this section with an example of a Σ_{cx}-smallest (thus Σ-countermonotonic) copula and an example of a Σ-countermonotonic copula that is not Σ_{cx}-smallest.

Example 6. Let $d = 100$. Let $\boldsymbol{V}^{(1)}$ be a random vector with uniform marginals corresponding to the copula

$$C_1(u_1, \ldots, u_{100}) = (\min_{j\leq 50} u_j + \min_{j\geq 51} u_j - 1)^+. \tag{11}$$

The copula C_1 in (11) is build from the symmetric Σ_{cx}-smallest Bernoulli distribution with support on $\boldsymbol{x}^{(1)} = (\boldsymbol{1}_{50}//\boldsymbol{0}_{50})$, where $\boldsymbol{0}_{50}$ is a vector of length 50 with all zeros, and $\boldsymbol{1}_{100} - \boldsymbol{x}^{(1)}$. By Proposition 12, $\boldsymbol{V}^{(1)}$ is a joint mix; therefore, it is Σ-countermonotonic and a Σ_{cx}-smallest element in its Fréchet class.

Let $d = 103$. Let $\boldsymbol{V}^{(2)}$ be a random vector with uniform marginals corresponding to the copula

$$C_2(u_1, \ldots, u_{103}) = (\min_{j\leq 51} u_j + \min_{j\geq 52} u_j - 1)^+. \tag{12}$$

The copula C_2 in (12) corresponds to the symmetric Σ_{cx}-smallest Bernoulli distribution with support on $\boldsymbol{x}^{(2)} = (\boldsymbol{1}_{51}//\boldsymbol{0}_{52})$ and $\boldsymbol{1}_{103} - \boldsymbol{x}^{(2)}$. According to Proposition 11, the random vector $\boldsymbol{V}^{(2)}$ is Σ-countermonotonic and a Σ_{cx}-smallest element in $\mathcal{C}_{103}^{\mathrm{EM}}$. However, $\boldsymbol{V}^{(2)}$ is not a Σ_{cx}-smallest element in its Fréchet class. In fact, we can see that $\boldsymbol{V}^{(2)}$ is not a joint mix, as the sum of its components varies in the interval $(51, 52)$; yet, its Fréchet class admits a joint mix. Consider, for example, $\boldsymbol{V}^{(JM)} := (\boldsymbol{V}^{(1)}//\boldsymbol{V}^{(3)})$, where $\boldsymbol{V}^{(3)}$ is a 3-dimensional joint mix with uniform marginals, independent of $\boldsymbol{V}^{(1)}$, and with the dependence structure specified in Example 3 in [39]. $\boldsymbol{V}^{(1)}$ and $\boldsymbol{V}^{(3)}$ are joint mixes, thus $\boldsymbol{V}^{(JM)}$ is a joint mix and

$$\sum_{j=1}^d V_j^{(JM)} \leq_{cx} \sum_{j=1}^d V_j^{(2)}.$$

3.2.2. FGM copulas

Another class of copulas that can be built from Bernoulli random vectors via a stochastic representation is the class $\mathcal{C}_d^{\mathrm{FGM}}$ of multivariate FGM copulas. In this section, we recall their definition and the stochastic representation introduced in [29], which provides a one-to-one correspondence with the class $\mathcal{SB}_d$. In the class of FGM copulas, the elements corresponding to Σ_{cx}-smallest Bernoulli random vectors are Σ_{cx}-smallest in the class $\mathcal{C}_d^{\mathrm{FGM}}$, but not in the whole class of copulas. It is therefore interesting to compare the minimal negative dependence in the class of FGM copulas, with the minimal dependence in the whole class of copulas. This is the focus of Section 4. Here, we present some known, but necessary, results on FGM copulas.

Definition 10. A multivariate copula C belongs to the class of FGM copulas if it has the following expression:

$$C(\boldsymbol{u}) = u_1 \cdots u_d \left(1 + \sum_{k=2}^{d} \sum_{1 \le j_1 < \cdots < j_k \le d} \theta_{j_1 \ldots j_k} \bar{u}_{j_1} \cdots \bar{u}_{j_k} \right), \quad \boldsymbol{u} \in [0,1]^d,$$

where, for each $j \in \{1, \ldots, d\}$, $\bar{u}_j = 1 - u_j$.

There exist 2^d constraints on the parameters for the existence of a FGM copula, which are

$$1 + \sum_{k=2}^{d} \sum_{1 \le j_1 < \cdots < j_k \le d} \theta_{j_1 \ldots j_k} \epsilon_{j_1} \epsilon_{j_2} \cdots \epsilon_{j_k} \ge 0,$$

for every $(\epsilon_1, \ldots, \epsilon_d) \in \{-1, 1\}^d$; see [40]. When $d = 2$, the admissible set for the unique parameter is the interval $[-1, 1]$. However, as the dimension increases, the shape of the set of admissible parameters becomes more and more complex. A useful tool to address this problem is the stochastic representation provided in [29]. Let $\boldsymbol{Z}_0 = (Z_{1,0}, \ldots, Z_{d,0})$ be a vector of independent exponential random variables with mean $1/2$ and let $\boldsymbol{Z}_1 = (Z_{1,1}, \ldots, Z_{d,1})$ be a vector of independent exponential random variables with mean 1. Let $\boldsymbol{Z}_0$ and $\boldsymbol{Z}_1$ be independent. The following theorem, proved in [29], shows the existence of a one-to-one correspondence between the class $\mathcal{SB}_d$ and the class $\mathcal{C}_d^{\mathrm{FGM}}$.

Theorem 3. *Let $\boldsymbol{X} \in \mathcal{SB}_d$ be a d-dimensional symmetric Bernoulli random vector. Let $\boldsymbol{U} = (U_1, \ldots, U_d)$ be a random vector such that, for each $j \in \{1, \ldots, d\}$,*

$$U_j = 1 - \exp\{-(Z_{j,0} + X_j Z_{j,1})\}. \tag{13}$$

Then, $\boldsymbol{U}$ has a d-variate distribution with standard uniform marginals, and its cdf is a FGM copula given, for each $\boldsymbol{u} \in [0,1]^d$, by

$$C(\boldsymbol{u}) = \sum_{\boldsymbol{x} \in \mathcal{X}_d} f_{\boldsymbol{X}}(\boldsymbol{x}) \prod_{h=1}^{d} u_h \left(1 + (-1)^{x_h} (1 - u_h) \right).$$

In [29], the authors derive the parameters of the FGM copula in terms of the centered moments of its corresponding symmetric Bernoulli distribution, viz.

$$\theta_{j_1 \ldots j_k} = (-2)^k \, \mathrm{E}_{\boldsymbol{X}} \left\{ \prod_{\ell=1}^{k} \left(X_{j_\ell} - \frac{1}{2} \right) \right\}, \tag{14}$$

for $k \in \{2, \ldots, d\}$ and $1 \le j_1 < \cdots < j_k \le d$.

Regarding extremal negative dependence in $\mathcal{C}_d^{\mathrm{FGM}}$, in [41], the authors explicitly find the FGM copula that corresponds to the Σ_{cx}-smallest exchangeable Bernoulli distribution. Moreover, in a slightly more general context, the authors of [30] show that the one-to-one map between the classes $\mathcal{SB}_d$ and $\mathcal{C}_d^{\mathrm{FGM}}$ preserves the convex order of the sums of the components. We restate Theorem 4.2 in [30] using the notation adopted in this paper.

Theorem 4. *Let $\boldsymbol{U}$ and $\boldsymbol{U}'$ be the uniform random vectors with FGM copula constructed from two Bernoulli random vectors $\boldsymbol{X}, \boldsymbol{X}' \in \mathcal{SB}_d$, respectively, via the stochastic representation in (13). Then*

$$\sum_{j=1}^{d} X_j \le_{cx} \sum_{j=1}^{d} X'_j \implies \sum_{j=1}^{d} U_j \le_{cx} \sum_{j=1}^{d} U'_j.$$

Theorem 4 implies that FGM copulas constructed from the Σ_{cx}-smallest element of $\mathcal{SB}_d$ are Σ_{cx}-smallest in the class of FGM copulas. Using the characterization of all Σ_{cx}-smallest elements of $\mathcal{SB}_d$ in Section 3.1, we can investigate some properties of a Σ_{cx}-smallest FGM copula.

Remark 5. We have already seen in Examples 1 and 2 that the Σ_{cx}-smallest pmfs are palindromic in dimension $d \le 4$. By Proposition 3.4 in [29], we know that the FGM copulas constructed from the class $\mathcal{PB}_d$ are radially symmetric and, in particular, have $\theta_{j_1 \ldots j_k} = 0$ for $1 \le j_1 < \cdots < j_k \le d$, for every odd value of k.

4. Minimal pairwise dependence measures

The most common dependence measures are Pearson's correlation ρ_P, Spearman's rho ρ_S, and Kendall's tau τ_K, defined by

$$\rho_P(X, Y) = \frac{\mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}, \quad \rho_S(X, Y) = \rho_P\{F_X(X), F_Y(Y)\},$$
$$\tau_K(X, Y) = \Pr\{(X - X')(Y - Y') \ge 0\} - \Pr\{(X - X')(Y - Y') \le 0\}.$$

When the marginals are continuous, Spearman's rho and Kendall's tau are measures that depend only on the copula of the two random variables and not on their marginals; see [13]. This is not true when the marginals are discrete; see Section 4.2 in [42]. It is evident that $\rho_P(U,V)=\rho_S(U,V)$ for uniform random variables U and V. It is also easy to verify that Spearman's rho and Pearson's correlation are equal for Bernoulli random variables; therefore, we consider only Pearson's correlation ρ_P.

In this section, we compare these measures for the Σ-countermonotonic random vectors in the class $\mathcal{SB}_d$, the Σ-countermonotonic copulas in the family of extremal mixture copulas, and the Σ_{cx}-smallest elements in the class of FGM copulas.

The following proposition follows from direct computations.

Proposition 13. *Let X_i and X_j be two Bernoulli random variables of means p_i and p_j, respectively. Then, we have*

$$\rho_P(X_i,X_j)=\frac{\tau_K(X_i,X_j)}{2\sqrt{p_i(1-p_i)p_j(1-p_j)}}.$$

In [31], the authors prove that the upper and lower bounds for correlations in a Fréchet class of multivariate Bernoulli pmfs are attained at the extremal points of the Fréchet class, making it possible to determine the range of admissible correlations. From Proposition 13, it follows that Kendall's tau also reaches its bounds on the extremal points.

For dimensions higher than 2, we adopt a simple approach to generalize these dependence measures, namely by averaging over all pairwise measures; see Section 3.1 in [43]. In particular, we denote the mean of Pearson's correlation and the mean of Kendall's tau of a d-dimensional random vector $\boldsymbol{Y}=(Y_1,\dots,Y_d)$ as follows

$$\bar\rho_P(\boldsymbol{Y})=\frac{2}{d(d-1)}\sum_{1\le i<j\le d}\rho_P(Y_i,Y_j),\quad \bar\tau_K(\boldsymbol{Y})=\frac{2}{d(d-1)}\sum_{1\le i<j\le d}\tau_K(Y_i,Y_j).$$

Proposition 14. *Let $\boldsymbol{X}\in\mathcal{SB}_d$ and let $\boldsymbol{V}\in C_d^{EM}$ and $\boldsymbol{U}\in C_d^{FGM}$ be the uniform random vectors built from $\boldsymbol{X}$ with extremal mixture copula and FGM copula, respectively. We have*

$$\rho_P(X_{j_1},X_{j_2})=\rho_P(V_{j_1},V_{j_2})=3\rho_P(U_{j_1},U_{j_2})=2\tau_K(X_{j_1},X_{j_2})=2\tau_K(V_{j_1},V_{j_2})=\frac{9}{2}\tau_K(U_{j_1},U_{j_2}),$$

for every $j_1,j_2\in\{1,\dots,d\}$, $j_1\neq j_2$, and

$$\bar\rho_P(\boldsymbol{X})=\bar\rho_P(\boldsymbol{V})=3\bar\rho_P(\boldsymbol{U})=2\bar\tau_K(\boldsymbol{X})=2\bar\tau_K(\boldsymbol{V})=\frac{9}{2}\bar\tau_K(\boldsymbol{U}).$$

Proof. Let $j_1,j_2\in\{1,\dots,d\}$, $j_1\neq j_2$. From Proposition 13, we have $\rho_P(X_{j_1},X_{j_2})=2\tau_K(X_{j_1},X_{j_2})$. Moreover, from the stochastic representation in (7), standard computations give $\rho_P(V_{j_1},V_{j_2})=\rho_P(X_{j_1},X_{j_2})$ and $\tau_K(V_{j_1},V_{j_2})=\tau_K(X_{j_1},X_{j_2})$. Regarding the FGM copula, Point 4 of Corollary 3.1 in [30] implies that $\rho_P(X_{j_1},X_{j_2})=3\rho_P(U_{j_1},U_{j_2})$. Since $\boldsymbol{U}$ has distribution in C_d^{FGM}, the cdf of the pair (U_{j_1},U_{j_2}) is a bivariate FGM copula with parameter $\theta_{j_1j_2}$, given by (14). It is well known that $\rho_P(U_{j_1},U_{j_2})=\theta_{j_1j_2}/3$ and $\tau_K(U_{j_1},U_{j_2})=2\theta_{j_1j_2}/9$; see [44]. Therefore, $\rho_P(U_{j_1},U_{j_2})=3\tau_K(U_{j_1},U_{j_2})/2$ and the conclusion follows since, from (14), $\theta_{j_1j_2}=\rho_P(X_{j_1},X_{j_2})$. □

As a consequence of Proposition 14, the analysis of the pairwise dependence measures of extremal mixture copulas and FGM copulas is fully described by the pairwise dependence measures of the corresponding symmetric Bernoulli distributions. Moreover, Corollary 2 shows that it is sufficient to consider palindromic Bernoulli distributions, and the corresponding copulas, to describe all the possible structures of pairwise dependence measures. Indeed, the authors of [45] proved that for every $\boldsymbol{X}\in\mathcal{SB}_d$ there exists $\boldsymbol{X}'\in\mathcal{PB}_d$ with the same bivariate Pearson's correlation structure.

Corollary 2. *Let $\boldsymbol{X}\in\mathcal{SB}_d$ and let $\boldsymbol{U}$ be a uniform random vector with the FGM copula constructed from $\boldsymbol{X}$. Then there exists $\boldsymbol{X}'\in\mathcal{PB}_d$ such that $\rho_P(U_{j_1},U_{j_2})=\rho_P(U'_{j_1},U'_{j_2})$ and $\tau_K(U_{j_1},U_{j_2})=\tau_K(U'_{j_1},U'_{j_2})$, for every $j_1,j_2\in\{1,\dots,d\}$, with $j_1\neq j_2$, where $\boldsymbol{U}'$ is a uniform random vector with FGM copula constructed from $\boldsymbol{X}'$. Moreover, the extremal mixture copulas constructed from $\boldsymbol{X}$ and $\boldsymbol{X}'$ via the stochastic representation in (7) coincide.*

Proof. Given $\boldsymbol{X}\in\mathcal{SB}_d$, from Theorem 1 in [45], there exists $\boldsymbol{X}'\in\mathcal{PB}_d$ such that $\rho_P(X_{j_1},X_{j_2})=\rho_P(X'_{j_1},X'_{j_2})$, for every $j_1,j_2\in\{1,\dots,d\}$, with $j_1\neq j_2$. Moreover, by Proposition 14, one has $\rho_P(U_{j_1},U_{j_2})=\rho_P(U'_{j_1},U'_{j_2})$ and $\tau_K(U_{j_1},U_{j_2})=\tau_K(U'_{j_1},U'_{j_2})$, for every $j_1,j_2\in\{1,\dots,d\}$, with $j_1\neq j_2$. From the construction of $\boldsymbol{X}'\in\mathcal{PB}_d$ in the proof of Theorem 1 in [45], $\boldsymbol{X}$ and $\boldsymbol{X}'$ are such that

$$f(\boldsymbol{s}_i)+f(\boldsymbol{1}_d-\boldsymbol{s}_i)=f'(\boldsymbol{s}_i)+f'(\boldsymbol{1}_d-\boldsymbol{s}_i),$$

for every $\boldsymbol{i}\in\mathcal{X}_{d-1}$. Therefore, the weights w_i, given in (8), of the corresponding extremal mixture copulas are equal and these copulas coincide. □

The following corollary to Proposition 14 states that extremal mixture copulas and FGM copulas built from Σ_{cx}-smallest Bernoulli random vectors have minimal mean correlation and Kendall's tau. Its proof relies on the following known fact: the mean Pearson correlation of a Bernoulli random vector $\boldsymbol{X}\in\mathcal{SB}_d$ can be expressed as the expectation of a convex function of the sum $S_{\boldsymbol{X}}=X_1+\dots+X_d$, i.e.,

$$\bar\rho_P(\boldsymbol{X})=\mathrm{E}\{\phi(S_{\boldsymbol{X}})\},\tag{15}$$

where

$$\phi(y) = \begin{cases} \frac{8}{d(d-1)}\binom{y}{2} - 1, & \text{if } y \geq 2, \\ -1, & \text{otherwise;} \end{cases} \tag{16}$$

see [46]. Since ϕ is a convex function, if $\boldsymbol{X}$ is a Σ_{cx}-smallest element in $\mathcal{SB}_d$, then $\bar{\rho}_P(\boldsymbol{X})$ is minimal and, by Proposition 6, we have

$$\bar{\rho}_P(\boldsymbol{X}) = \begin{cases} -\frac{1}{d-1}, & \text{if } d \text{ is even}, \\ -\frac{1}{d}, & \text{if } d \text{ is odd}. \end{cases} \tag{17}$$

We observe that if d is odd, the minimal mean Pearson correlation in the class $\mathcal{SB}_d$ is equal to the minimal mean Pearson correlation in the class $\mathcal{SB}_{d+1}$.

Corollary 3. *Let $\boldsymbol{X} \in \mathcal{SB}_d$ be a Σ_{cx}-smallest element in $\mathcal{SB}_d$ and let $\boldsymbol{V} \in \mathcal{C}_d^{EM}$ and $\boldsymbol{U} \in \mathcal{C}_d^{FGM}$ be the uniform random vectors built from $\boldsymbol{X}$ with extremal mixture copula and FGM copula, respectively. We have*

$$\bar{\rho}_P(\boldsymbol{V}) \leq \bar{\rho}_P(\boldsymbol{V}'), \quad \bar{\tau}_K(\boldsymbol{V}) \leq \bar{\tau}_K(\boldsymbol{V}'),$$

for any $\boldsymbol{V}' \in \mathcal{C}_d^{EM}$, and

$$\bar{\rho}_P(\boldsymbol{U}) \leq \bar{\rho}_P(\boldsymbol{U}'), \quad \bar{\tau}_K(\boldsymbol{U}) \leq \bar{\tau}_K(\boldsymbol{U}'),$$

for any $\boldsymbol{U}' \in \mathcal{C}_d^{FGM}$. Moreover,

$$\bar{\rho}_P(\boldsymbol{V}) = \begin{cases} -\frac{1}{d-1}, & \text{if } d \text{ is even}, \\ -\frac{1}{d}, & \text{if } d \text{ is odd}, \end{cases} \quad \bar{\tau}_K(\boldsymbol{V}) = \begin{cases} -\frac{1}{2(d-1)}, & \text{if } d \text{ is even}, \\ -\frac{1}{2d}, & \text{if } d \text{ is odd}, \end{cases}$$

and

$$\bar{\rho}_P(\boldsymbol{U}) = \begin{cases} -\frac{1}{3(d-1)}, & \text{if } d \text{ is even}, \\ -\frac{1}{3d}, & \text{if } d \text{ is odd}, \end{cases} \quad \bar{\tau}_K(\boldsymbol{U}) = \begin{cases} -\frac{2}{9(d-1)}, & \text{if } d \text{ is even}, \\ -\frac{2}{9d}, & \text{if } d \text{ is odd}. \end{cases}$$

Proof. Let $\boldsymbol{X} \in \mathcal{SB}_d$ be a Σ_{cx}-smallest element in $\mathcal{SB}_d$. From (15) and (16), since ϕ in (16) is a convex function, $\bar{\rho}_P(\boldsymbol{X}) \leq \bar{\rho}_P(\boldsymbol{X}')$, for any $\boldsymbol{X}' \in \mathcal{SB}_d$. Since every extremal mixture copula and every FGM copula can be built from a Bernoulli random vector $\boldsymbol{X}' \in \mathcal{SB}_d$, the thesis follows from Proposition 14 and from (17). □

The mean Pearson correlation of a Σ_{cx}-smallest Bernoulli random vector, or of a copula built from it, depends only on the dimension of the class. However, Σ_{cx}-smallest Bernoulli random vectors in the same class $\mathcal{SB}_d$ have different dependence structures. We now study pairwise dependence measures of different Σ_{cx}-smallest Bernoulli random vectors in the class $\mathcal{SB}_d$. We first consider a Σ_{cx}-smallest Bernoulli random vector $\boldsymbol{X}^K$ with pmf $f^K \in \mathcal{B}_K$. We know that there exists $\boldsymbol{x} \in \mathcal{X}_d^*$ such that $f^K(\boldsymbol{x}) = f^K(\boldsymbol{1}_d - \boldsymbol{x}) = 1/2$. It follows that there are n_d^+ comonotonic pairs (a pair is comonotonic if one random variable is a deterministic non-decreasing transformation of the other; see [8]), where

$$n_d^+ = \begin{cases} \frac{d(d-2)}{4}, & \text{if } d \text{ is even}, \\ \frac{(d-1)^2}{4}, & \text{if } d \text{ is odd}, \end{cases}$$

and n_d^- countermonotonic pairs, where

$$n_d^- = \binom{d}{2} - n_d^+ = \begin{cases} \frac{d^2}{4}, & \text{if } d \text{ is even}, \\ \frac{(d-1)(d+1)}{4}, & \text{if } d \text{ is odd}. \end{cases}$$

Obviously, countermonotonic pairs have correlation $\rho_P = -1$ and comonotonic pairs have correlation $\rho_P = 1$, and the mean Pearson correlation of a random vector $\boldsymbol{X}^K \in \mathcal{B}_K$ is given by $2(n_d^+ - n_d^-)/\{d(d-1)\}$, which is equal to (17). Since the extremal copulas are in a one-to-one relationship with the elements in $\mathcal{B}_K$, and $\rho_P(X_{j_1}, X_{j_2}) = \rho_P(V_{j_1}, V_{j_2})$, we can conclude that extremal copulas that are built from Σ_{cx}-smallest Bernoulli random vectors have n_d^+ comonotonic pairs and n_d^- countermonotonic pairs, and they are Σ-countermonotonic by Proposition 11.

We can then consider the unique exchangeable Σ_{cx}-smallest Bernoulli random vector $\boldsymbol{X}^e \in \mathcal{SB}_d$. In this case, for every $j_1, j_2 \in \{1, \dots, d\}$, with $j_1 \neq j_2$, we have $\rho_P(X_{j_1}^e, X_{j_2}^e) = \bar{\rho}_P(\boldsymbol{X}^e)$, where $\bar{\rho}_P(\boldsymbol{X}^e)$ is given by (17). The Bernoulli random vector $\boldsymbol{X}^e$ and the uniform random vectors $\boldsymbol{V}^e$ and $\boldsymbol{U}^e$, associated with the extremal mixture copula and FGM copula defined by $\boldsymbol{X}^e$, are pairwise negatively correlated (all pairwise Pearson's correlations are non-positive). See [23] for a more detailed analysis of the properties of $\boldsymbol{X}^e$.

Corollary 2 implies that palindromic Bernoulli distributions are sufficient to construct FGM copulas with any admissible bivariate dependence measures ρ_P and τ_K. Let $\boldsymbol{X}$ be a Bernoulli random vector with pmf $f \in \mathcal{SB}_d$, but $f \notin \mathcal{PB}_d$, and let $\boldsymbol{V}$ and $\boldsymbol{U}$ be the uniform random vectors built from $\boldsymbol{X}$ with extremal mixture copula and FGM copula, respectively. By Corollary 2, there exists $\boldsymbol{X}'$ with pmf $f' \in \mathcal{PB}_d$ such that $\boldsymbol{V}$ has the same distribution of $\boldsymbol{V}'$ and $\rho_P(U_{j_1}, U_{j_2}) = \rho_P(U'_{j_1}, U'_{j_2})$, for every $j_1, j_2 \in \{1, \dots, d\}$, with $j_1 \neq j_2$, where $\boldsymbol{V}'$ and $\boldsymbol{U}'$ are the uniform random vectors with the extremal mixture copula and FGM copula constructed from $\boldsymbol{X}'$, respectively. To compare the dependence structures of $\boldsymbol{X}$, $\boldsymbol{X}'$, and their corresponding FGM copulas, we need to consider higher order moments. We define the centered cross moments of order three of a d-dimensional random vector $\boldsymbol{Y} = (Y_1, \dots, Y_d)$ as

$$\tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{Y}) = \mathrm{E}\left[\prod_{h=1}^{3}\left\{\frac{Y_{j_h} - \mathrm{E}(Y_{j_h})}{\sqrt{\mathrm{Var}(Y_{j_h})}}\right\}\right],$$

for $j_1, j_2, j_3 \in \{1, \dots, d\}$, with j_1, j_2, j_3 all distinct.

Proposition 15. *Let $\boldsymbol{X}' \in \mathcal{PB}_d$ and let $\boldsymbol{V}'$ and $\boldsymbol{U}'$ be the uniform random vectors constructed from $\boldsymbol{X}' \in \mathcal{PB}_d$ with extremal mixture copula and FGM copula, respectively. Then, $\tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{X}') = \tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{V}') = \tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{U}') = 0$, for every $j_1, j_2, j_3 \in \{1, \dots, d\}$, with j_1, j_2, j_3 all distinct.*

Proof. It is easy to show that if a random vector $\boldsymbol{Y}$ has the same distribution as $\mathbf{1}_d - \boldsymbol{Y}$, then $\tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{Y}) = 0$ for every $j_1, j_2, j_3 \in \{1, \dots, d\}$, with j_1, j_2, j_3 all distinct. A palindromic Bernoulli random vector is such that $\boldsymbol{X}'$ has the same distribution as $\mathbf{1}_d - \boldsymbol{X}'$, and every extremal mixture copula is radially symmetric, i.e., $\boldsymbol{V}'$ has the same distribution as $\mathbf{1}_d - \boldsymbol{V}'$. Finally, by Proposition 3.4 in [29], if $\boldsymbol{X}' \in \mathcal{PB}_d$, the corresponding FGM copula $\boldsymbol{U}'$ is radially symmetric. □

The above proposition shows that the palindromic Bernoulli random vectors and the copulas built from them have centered cross moments of order three equal to zero. This is not true for any $\boldsymbol{f} \in \mathcal{SB}_d$. Proposition 16 establishes the relation between centered cross moments of order three of a general Bernoulli random vector and of the copulas constructed from it. We conclude this section by providing an example of a symmetric Bernoulli random vector that is non-palindromic, and whose centered cross moments of order three are different from zero. The proof of the following proposition follows from standard computation and is therefore omitted.

Proposition 16. *Let $\boldsymbol{X} \in \mathcal{SB}_d$ and let $\boldsymbol{V}$ and $\boldsymbol{U}$ be the uniform random vectors constructed from $\boldsymbol{X}$ with extremal mixture copula and FGM copula, respectively. Then, $\tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{X}) = \tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{V}) = -\sqrt{3}/9\, \tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{U})$, for every $j_1, j_2, j_3 \in \{1, \dots, d\}$, with j_1, j_2, j_3 all distinct.*

We conclude this section with an example in dimension $d = 6$.

Example 7. Consider $\boldsymbol{f}^{(1)} \in \mathcal{SB}_6$ from Example 4, which is clearly not palindromic. Let $\boldsymbol{X}$ be a Bernoulli random vector with pmf $\boldsymbol{f}^{(1)}$, and let $\boldsymbol{V}$ and $\boldsymbol{U}$ be the uniform random vectors built from $\boldsymbol{X}$ with extremal mixture copula and FGM copula, respectively. From Corollary 2, there exists a Bernoulli random vector $\boldsymbol{X}'$, with pmf $\boldsymbol{f}' \in \mathcal{PB}_6$, such that $\rho_P(U_{j_1}, U_{j_2}) = \rho_P(U'_{j_1}, U'_{j_2})$, for every $j_1, j_2 \in \{1, \dots, d\}$, with $j_1 \neq j_2$, where $\boldsymbol{U}'$ is a uniform random vector with the FGM copula defined by $\boldsymbol{X}'$. Moreover, $\boldsymbol{V}$ has the same distribution as $\boldsymbol{V}'$, where $\boldsymbol{V}'$ is a uniform random vector with the extremal mixture copula constructed from $\boldsymbol{X}'$. However, it is not true that $\boldsymbol{U}$ has the same distribution as $\boldsymbol{U}'$. Indeed, by Proposition 15, we have $\tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{X}') = \tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{U}') = 0$ for every $j_1, j_2, j_3 \in \{1, \dots, 6\}$, with j_1, j_2, j_3 all distinct, while

$$\tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{X}) = \begin{cases} 1, & \text{if } (j_1, j_2, j_3) \in \{(1,2,4), (1,3,5), (2,5,6), (3,4,6)\}, \\ -1, & \text{if } (j_1, j_2, j_3) \in \{(3,5,6), (2,4,6), (1,3,4), (1,2,5)\}, \\ 0, & \text{otherwise}, \end{cases}$$

and, by Proposition 16, $\tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{U}) = -3\sqrt{3}\, \tilde{\mu}_{j_1,j_2,j_3}(\boldsymbol{X})$, for every $j_1, j_2, j_3 \in \{1, \dots, 6\}$, with j_1, j_2, j_3 all distinct.

5. Conclusion

Some classes of copulas can be constructed from multivariate symmetric Bernoulli distributions, inheriting certain dependence properties. We study the minimal risk and extremal negative dependence distributions of multivariate symmetric Bernoulli distributions and characterize the dependence properties of the corresponding copulas. In particular, we explicitly identify a class of Σ-countermonotonic copulas. The connection between copulas and Bernoulli distributions proves effective in deriving statistical properties of copula families, including minimal correlation. In this context, the recent article [30] investigates the characterization of extremal negative dependence within the class of FGM copulas and some of their generalizations. A key role in our findings is played by the geometric and algebraic structure of multivariate Bernoulli distributions, which has its own theoretical interest and warrants further investigation in our future research.

Acknowledgments

The authors sincerely thank the Associate Editor and two anonymous referees for their careful reading of the paper and for their valuable suggestions and comments, which have significantly improved the manuscript. Patrizia Semeraro gratefully acknowledges financial support from the INdAM-GNAMPA project CUP E53C22001930001.

Appendix A. Proofs of Section 2

Proof of Proposition 2. Assume that $\boldsymbol{f}^K \in \mathcal{B}_K \subseteq \mathcal{SB}_d$ is not an extremal point of $\mathcal{SB}_d$. Then, there exist $\boldsymbol{f}_1, \boldsymbol{f}_2 \in \mathcal{SB}_d$, with $\boldsymbol{f}_1 \neq \boldsymbol{f}_2$, and $\lambda \in (0,1)$ such that

$$\boldsymbol{f}^K = \lambda \boldsymbol{f}_1 + (1-\lambda)\boldsymbol{f}_2. \tag{A.1}$$

Since $\boldsymbol{f}^K \in \mathcal{B}_K$, there exists $\boldsymbol{x} \in \mathcal{X}_d$ such that $f^K(\boldsymbol{x}) = f^K(\mathbf{1}_d - \boldsymbol{x}) = 1/2$ and $f^K(\boldsymbol{y}) = 0$, for every $\boldsymbol{y} \in \mathcal{X}_d \setminus \{\boldsymbol{x}, \mathbf{1}_d - \boldsymbol{x}\}$. (A.1) implies that $f_1(\boldsymbol{y}) = f_2(\boldsymbol{y}) = 0$, for every $\boldsymbol{y} \in \mathcal{X}_d \setminus \{\boldsymbol{x}, \mathbf{1}_d - \boldsymbol{x}\}$. Since $\boldsymbol{f}_1, \boldsymbol{f}_2 \in \mathcal{SB}_d$, the constraints on the marginal means imply that $f_1(\boldsymbol{x}) = f_1(\mathbf{1}_d - \boldsymbol{x}) = f_2(\boldsymbol{x}) = f_2(\mathbf{1}_d - \boldsymbol{x}) = 1/2$. Hence $\boldsymbol{f}^K = \boldsymbol{f}_1 = \boldsymbol{f}_2$, which contradicts $\boldsymbol{f}_1 \neq \boldsymbol{f}_2$. We conclude that $\boldsymbol{f}^K \in \mathcal{B}_K$ is an extremal point of $\mathcal{SB}_d$. □

Proof of Proposition 3. Let $P(\boldsymbol{z})$ and $Q(\boldsymbol{z})$ be two equivalent polynomials and let $\boldsymbol{f}^P$ and $\boldsymbol{f}^Q$ be the two corresponding type-0 pmfs. By Definition 1, there exists $\mu \in (0, +\infty) \cap \mathbb{Q}$, such that $P(\boldsymbol{z}) = \mu Q(\boldsymbol{z})$. We apply Algorithm 1 to both $P(\boldsymbol{z})$ and $Q(\boldsymbol{z})$. The coefficients of the two polynomials have the same sign; therefore, after the first step of the algorithm, we have $\boldsymbol{f}^P = \mu \boldsymbol{f}^Q$. After normalization, the two type-0 pmfs are identical. □

Proof of Proposition 4. Let $A := \{\boldsymbol{f} \in \mathcal{SB}_d : \boldsymbol{f} = \lambda \boldsymbol{f}^P + (1-\lambda)\boldsymbol{f}^K, \text{ for } \boldsymbol{f}^K \in \mathcal{PB}_d, \lambda \in (0,1] \cap \mathbb{Q}\}$, where $\boldsymbol{f}^P$ is the type-0 pmf of $P(\boldsymbol{z})$ and, without restriction, suppose that $P(\boldsymbol{z})$ is such that $\mathcal{H}(\boldsymbol{f}^P) = P(\boldsymbol{z})$. The map $\mathcal{H}$ is linear, therefore $\mathcal{H}(\lambda \boldsymbol{f}^P + (1-\lambda)\boldsymbol{f}^K) = \lambda \mathcal{H}(\boldsymbol{f}^P) + (1-\lambda)\mathcal{H}(\boldsymbol{f}^K) = \lambda P(\boldsymbol{z})$, which implies $A \subseteq \mathcal{H}^{-1}[P(\boldsymbol{z})]$. Let $\boldsymbol{f} \in \mathcal{H}^{-1}[P(\boldsymbol{z})]$ and let $\mathcal{H}(\boldsymbol{f}) = Q(\boldsymbol{z})$. By definition of $\mathcal{H}^{-1}$, there exists $\mu \in (0, +\infty) \cap \mathbb{Q}$ such that $Q(\boldsymbol{z}) = \mu P(\boldsymbol{z})$ and, for every $\boldsymbol{i} \in \mathcal{X}_{d-1}$, we have

$$a_{\boldsymbol{i}}^Q = \mu a_{\boldsymbol{i}}^P, \tag{A.2}$$

where $a_{\boldsymbol{i}}^Q$ and $a_{\boldsymbol{i}}^P$ are the coefficients of the polynomials $Q(\boldsymbol{z})$ and $P(\boldsymbol{z})$, respectively. Our first goal is to show that $\mu \leq 1$. For every $\boldsymbol{x} \in \mathcal{X}_d$, there exists a unique $\boldsymbol{i} \in \mathcal{X}_{d-1}$ such that $\boldsymbol{x} = \boldsymbol{s}_{\boldsymbol{i}}$ or $\boldsymbol{x} = \mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}}$. By Algorithm 1, for each $\boldsymbol{i} \in \mathcal{X}_{d-1}$, we have either $|a_{\boldsymbol{i}}^P| = f^P(\boldsymbol{s}_{\boldsymbol{i}})$ or $|a_{\boldsymbol{i}}^P| = f^P(\mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}})$, so that

$$\sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}} |a_{\boldsymbol{i}}^P| = \sum_{\boldsymbol{x} \in \mathcal{X}_d} f^P(\boldsymbol{x}) = 1.$$

Moreover, (3) implies

$$\sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}} |a_{\boldsymbol{i}}^Q| \leq \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}} |f(\boldsymbol{s}_{\boldsymbol{i}})| + |f(\mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}})| = \sum_{\boldsymbol{x} \in \mathcal{X}_d} f(\boldsymbol{x}) = 1 = \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}} |a_{\boldsymbol{i}}^P|. \tag{A.3}$$

By (A.2) and (A.3), we conclude that $\mu \leq 1$. We now aim to prove that $\boldsymbol{f} \in A$, i.e., that there exist $\lambda \in (0,1] \cap \mathbb{Q}$ and a pmf $\boldsymbol{f}^K \in \mathcal{PB}_d$ such that $\boldsymbol{f} = \lambda \boldsymbol{f}^P + (1-\lambda)\boldsymbol{f}^K$. If we show that

$$\boldsymbol{f}^K = \frac{1}{1-\mu}\boldsymbol{f} - \frac{\mu}{1-\mu}\boldsymbol{f}^P$$

is a pmf with null polynomial, the claim follows by taking $\lambda = \mu$. First, we observe that

$$\mathcal{H}(\boldsymbol{f}^K) = \mathcal{H}\left(\frac{1}{1-\mu}\boldsymbol{f} - \frac{\mu}{1-\mu}\boldsymbol{f}^P\right) = \frac{1}{1-\mu}\mathcal{H}(\boldsymbol{f}) - \frac{\mu}{1-\mu}\mathcal{H}(\boldsymbol{f}^P) = 0.$$

Also, since $1/(1-\mu) - \mu/(1-\mu) = 1$, the components of $\boldsymbol{f}^K$ have sum equal to one. It remains to show that the components of $\boldsymbol{f}^K$ are nonnegative. From (3), $f(\boldsymbol{s}_{\boldsymbol{i}}) - f(\mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}}) = \mu f^P(\boldsymbol{s}_{\boldsymbol{i}}) - \mu f^P(\mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}})$. Thus, we have

$$f(\boldsymbol{s}_{\boldsymbol{i}}) - \mu f^P(\boldsymbol{s}_{\boldsymbol{i}}) = f(\mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}}) - \mu f^P(\mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}}). \tag{A.4}$$

By construction, the type-0 pmf satisfies either $f^P(\boldsymbol{s}_{\boldsymbol{i}}) = 0$ or $f^P(\mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}}) = 0$. In the first case, (A.4) becomes $f(\mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}}) - \mu f^P(\mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}}) = f(\boldsymbol{s}_{\boldsymbol{i}}) \geq 0$, while in the other case, (A.4) becomes $f(\boldsymbol{s}_{\boldsymbol{i}}) - \mu f^P(\boldsymbol{s}_{\boldsymbol{i}}) = f(\mathbf{1}_d - \boldsymbol{s}_{\boldsymbol{i}}) \geq 0$. Thus, for every $\boldsymbol{x} \in \mathcal{X}_d$, we have $f(\boldsymbol{x}) - \mu f^P(\boldsymbol{x}) \geq 0$. Since for every $\boldsymbol{x} \in \mathcal{X}_d$ we have $f^K(\boldsymbol{x}) = (f(\boldsymbol{x}) - \mu f^P(\boldsymbol{x}))/(1-\mu) \geq 0$, all the components of $\boldsymbol{f}^K$ are non-negative. Hence, $\boldsymbol{f}^K$ is a pmf of the kernel of $\mathcal{H}$ and we have $\mathcal{H}^{-1}[P(\boldsymbol{z})] \subseteq A$. It follows that $\mathcal{H}^{-1}[P(\boldsymbol{z})] \equiv A$. □

Proof of Proposition 5. Let $\boldsymbol{f}$ be an extremal pmf of $\mathcal{SB}_d$ and let $\mathcal{H}(\boldsymbol{f}) = P(\boldsymbol{z})$. If $P(\boldsymbol{z}) \equiv 0$, then $\boldsymbol{f} \in \mathcal{B}_K$, which is the set of extremal points that generate $\mathcal{K}(\mathcal{H})$. Suppose $P(\boldsymbol{z}) \not\equiv 0$. By Proposition 4, there exists an element of the kernel of $\mathcal{H}$, $\boldsymbol{f}^K \in \mathcal{K}(\mathcal{H})$ such that

$$\boldsymbol{f} = \lambda \boldsymbol{f}^P + (1-\lambda)\boldsymbol{f}^K,$$

for some $\lambda \in (0,1] \cap \mathbb{Q}$, where $\boldsymbol{f}^P$ is the type-0 pmf of $P(\boldsymbol{z})$. Since $\boldsymbol{f}$ is an extremal point, we have $\boldsymbol{f} = \boldsymbol{f}^P$. □

Appendix B. Proofs of Section 3

Proof of Proposition 8. Any pmf $\boldsymbol{f}^{K*} \in \mathcal{PB}_d \equiv \mathcal{K}(\mathcal{H})$ can be expressed as a convex linear combination of $\mathcal{B}_K$,

$$\boldsymbol{f}^{K*} = \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}} \lambda_i \boldsymbol{g}_i,$$

where, for $\boldsymbol{i} \in \mathcal{X}_{d-1}$, the extremal pmf $\boldsymbol{g}_i \in \mathcal{B}_K$ satisfies $g_i(\boldsymbol{s}_i) = g_i(\mathbf{1}_d - \boldsymbol{s}_i) = 1/2$. Let $\boldsymbol{f}^{K*}$ be a Σ_{cx}-smallest element in $\mathcal{SB}_d$. By Proposition 7, $\boldsymbol{f}^{K*}$ has support on $\mathcal{X}_d^*$, and it is easy to show that $\lambda_i = 0$ for every $\boldsymbol{g}_i$ that is not a Σ_{cx}-smallest element in $\mathcal{SB}_d$. The converse implication holds since any convex linear combination of Σ_{cx}-smallest elements is itself a Σ_{cx}-smallest element. □

Proof of Theorem 1. By construction, $\boldsymbol{x} \in \mathcal{X}_d^*$ if and only if $\mathbf{1}_d - \boldsymbol{x} \in \mathcal{X}_d^*$, and, by Proposition 7, if $\boldsymbol{f}^*$ is a Σ_{cx}-smallest pmf in $\mathcal{SB}_d$, then $f^*(\boldsymbol{x}) = 0$ for every $\boldsymbol{x} \notin \mathcal{X}_d^*$. Moreover, for $\boldsymbol{i} = (i_1, \dots, i_{d-1}) \notin \mathscr{I}_{d-1}^*$, we have $\boldsymbol{s}_i = (i_1, \dots, i_{d-1}, 0) \notin \mathcal{X}_d^*$ and $\mathbf{1}_d - \boldsymbol{s}_i = (1 - i_1, \dots, 1 - i_{d-1}, 1) \notin \mathcal{X}_d^*$. It follows that, for every $\boldsymbol{i} \notin \mathscr{I}_{d-1}^*$, $f^*(\boldsymbol{s}_i) = f^*(\mathbf{1}_d - \boldsymbol{s}_i) = 0$, and $a_i = f^*(\boldsymbol{s}_i) - f^*(\mathbf{1}_d - \boldsymbol{s}_i) = 0$. This proves point (i). We now prove the other two points of the theorem with the assumption that d is odd; the case where d is even follows analogously. By point (i), we can write

$$P^*(\boldsymbol{z}) = \sum_{\boldsymbol{i} \notin \mathscr{I}_{d-1}^*} a_i \boldsymbol{z}^{\boldsymbol{i}} = \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}^{M_d}} a_i \boldsymbol{z}^{\boldsymbol{i}} + \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}^{M_d+1}} a_i \boldsymbol{z}^{\boldsymbol{i}}. \tag{B.1}$$

Corollary 3.1 in [20] states that the image of any pmf through the map $\mathcal{H}$ is a linear combination of the fundamental polynomials. For each $\boldsymbol{i} \in \mathcal{X}_{d-1} \setminus \{\mathbf{0}\}$, where $\mathbf{0}$ denotes the $(d-1)$-dimensional zero vector, the fundamental polynomial $F_i(\boldsymbol{z})$ is given by

$$F_i(\boldsymbol{z}) = F_{j_1,\dots,j_{n_i}}(\boldsymbol{z}) = \prod_{h=1}^{n_i} z_{j_h} - \sum_{h=1}^{n_i} z_{j_h} + (n_i - 1) = \boldsymbol{z}^{\boldsymbol{i}} - \sum_{h=1}^{n_i} z_{j_h} + (n_i - 1),$$

where $(j_1, \dots, j_{n_i})$ are the position of the ones in $\boldsymbol{i}$, and $n_i := \sum_{j=1}^{d-1} i_j$ is their number; see [20] for further details. Note that $F_i(\boldsymbol{z}) \equiv 0$ for every $\boldsymbol{i} \in \mathcal{X}_{d-1}^1$. We can write $P^*(\boldsymbol{z}) = \mathcal{H}(\boldsymbol{f}^*)$ as a linear combination of the fundamental polynomials

$$P^*(\boldsymbol{z}) = \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1} \setminus \{\mathbf{0}\}} \gamma_i F_i(\boldsymbol{z}) = \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1} \setminus \{\mathbf{0}\}} \gamma_i \left\{ \boldsymbol{z}^{\boldsymbol{i}} - \sum_{h=1}^{n_i} z_{j_h} + (n_i - 1) \right\}, \tag{B.2}$$

with $\gamma_i \in \mathbb{Q}$, for every $\boldsymbol{i} \in \mathcal{X}_{d-1} \setminus \{\mathbf{0}\}$. Equating the two representation of $P^*(\boldsymbol{z})$ in (B.1) and (B.2), for every $\boldsymbol{i} \in \cup_{k=2}^{d-1} \mathcal{X}_{d-1}^k$, we have $\gamma_i = a_i$ if $\boldsymbol{i} \in \mathscr{I}_{d-1}^*$, and $\gamma_i = 0$ otherwise. Moreover, since $M_d \geq 1$, the expression in (B.1) has no constant term. Therefore, the constant term in (B.2) equals zero, i.e.,

$$\sum_{\boldsymbol{i} \in \mathcal{X}_{d-1} \setminus \{\mathbf{0}\}} a_i (n_i - 1) = (M_d - 1) \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}^{M_d}} a_i + M_d \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}^{M_d+1}} a_i = 0. \tag{B.3}$$

Recall that the polynomial $P^*(\boldsymbol{z}) \in \mathcal{I}_{\mathcal{P}}$ vanishes at points $\mathcal{P} = \{\mathbf{1}_{d-1}, \mathbf{1}_{d-1}^{-j}, j \in \{1, \dots, d-1\}\}$; in particular,

$$P^*(\mathbf{1}_{d-1}) = \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}^{M_d}} a_i + \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}^{M_d+1}} a_i = 0. \tag{B.4}$$

Point (ii) follows by plugging (B.4) in (B.3). In addition, for every $j \in \{1, \dots, d-1\}$, we can rewrite $P^*(\mathbf{1}_{d-1})$ as

$$P^*(\mathbf{1}_{d-1}) = \sum_{\boldsymbol{i} \in \mathscr{I}_{d-1}^* : i_j = 0} a_i + \sum_{\boldsymbol{i} \in \mathscr{I}_{d-1}^* : i_j = 1} a_i = 0, \tag{B.5}$$

and, for every $j \in \{1, \dots, d-1\}$, we also have

$$P^*(\mathbf{1}_{d-1}^{-j}) = \sum_{\boldsymbol{i} \in \mathscr{I}_{d-1}^* : i_j = 0} a_i - \sum_{\boldsymbol{i} \in \mathscr{I}_{d-1}^* : i_j = 1} a_i = 0. \tag{B.6}$$

Eqs. (B.5) and (B.6) imply point (iii). □

Proof of Corollary 1. By point (i), the condition in point (ii) of Theorem 1 simplifies to

$$\begin{cases} \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}^{M_d}} a_i = \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}^{M_d+1}} a_i = 0, & \text{if } d \text{ is odd,} \\ \sum_{\boldsymbol{i} \in \mathcal{X}_{d-1}^{M_d}} a_i = 0, & \text{if } d \text{ is even.} \end{cases} \tag{B.7}$$

Moreover, since

$$\sum_{\boldsymbol{i} \in \mathscr{I}_{d-1}^* : i_j = 1} a_i = \sum_{\boldsymbol{i} \in \mathscr{I}_{d-1}^*} i_j a_i,$$

point (iii) of Theorem 1 reads

$$\sum_{i \in \mathscr{I}^*_{d-1}} i_j a_i = 0, \quad j \in \{1, \dots, d-1\}. \tag{B.8}$$

We consider two cases. If d is odd, (B.7) and (B.8) lead to the linear system

$$\begin{cases} \sum_{i \in \mathcal{X}^{M_d}_{d-1}} a_i = 0, \\ \sum_{i \in \mathcal{X}^{M_d+1}_{d-1}} a_i = 0, \\ \sum_{i \in \mathscr{I}^*_{d-1}} i_j a_i = 0, \quad j \in \{1, \dots, d-1\}, \end{cases}$$

whose coefficient matrix is $A_d = (R_1 // R_2 // A_{\mathscr{I}^*_{d-1}})$. If instead d is even, the linear system becomes

$$\begin{cases} \sum_{i \in \mathcal{X}^{M_d}_{d-1}} a_i = 0, \\ \sum_{i \in \mathscr{I}^*_{d-1}} i_j a_i = 0, \quad j \in \{1, \dots, d-1\}, \end{cases}$$

whose coefficient matrix is $A_d = (\mathbf{1}^\top_{n^*_d} // A_{\mathscr{I}^*_{d-1}})$. Hence, the assertion follows. □

Proof of Theorem 2. We first consider the case d odd. Our first goal is to show that $P^*(\boldsymbol{z}) \in \mathcal{I}_{\mathcal{P}}$, i.e., $P^*(\mathbf{1}_{d-1}) = 0$ and $P^*(\mathbf{1}^{-j}_{d-1}) = 0$, for every $j \in \{1, \dots, d\}$. From point (i) of Theorem 1, we can write

$$P^*(\boldsymbol{z}) = \sum_{i \in \mathscr{I}^*_{d-1}} a_i \boldsymbol{z}^i = \sum_{i \in \mathcal{X}^{M_d}_{d-1}} a_i \boldsymbol{z}^i + \sum_{i \in \mathcal{X}^{M_d+1}_{d-1}} a_i \boldsymbol{z}^i.$$

Therefore,

$$P^*(\mathbf{1}_{d-1}) = \sum_{i \in \mathscr{I}^*_{d-1}} a_i = \sum_{i \in \mathcal{X}^{M_d}_{d-1}} a_i + \sum_{i \in \mathcal{X}^{M_d+1}_{d-1}} a_i = 0,$$

where the last equality follows from point (ii) of Theorem 1. Moreover, for each $j \in \{1, \dots, d-1\}$, it holds

$$0 = \sum_{i \in \mathscr{I}^*_{d-1}} a_i = \sum_{i \in \mathscr{I}^*_{d-1} : i_j = 0} a_i + \sum_{i \in \mathscr{I}^*_{d-1} : i_j = 1} a_i. \tag{B.9}$$

Hence, by point (iii) of Theorem 1 and by (B.9), it follows that, for each $j \in \{1, \dots, d-1\}$,

$$P^*(\mathbf{1}^{-j}_{d-1}) = \sum_{i \in \mathscr{I}^*_{d-1} : i_j = 0} a_i - \sum_{i \in \mathscr{I}^*_{d-1} : i_j = 1} a_i = 0.$$

Then, $P^*(\boldsymbol{z}) \in \mathcal{I}_{\mathcal{P}}$, and we can apply Algorithm 1 to find the corresponding type-0 pmf $\boldsymbol{f}^*$. For every $\boldsymbol{i} \in \mathcal{X}_{d-1}$ such that $a_{\boldsymbol{i}} = 0$, the type-0 pmf has $f^*(\boldsymbol{s_i}) = f^*(\mathbf{1}_d - \boldsymbol{s_i}) = 0$. This implies that, given $k \in \{0, 1, \dots, d-1\}$, if $a_{\boldsymbol{i}} = 0$ for all $\boldsymbol{i} \in \mathcal{X}^k_{d-1}$, the sum of the components of a Bernoulli random vector with the type-0 as pmf does not have support on k or $d-k$. Therefore, the condition $a_{\boldsymbol{i}} = 0$ for every $\boldsymbol{i} \notin \mathscr{I}^*_{d-1}$ implies that the type-0 pmf $\boldsymbol{f}^*$ has support only on points with sum equal to M_d or $M_d + 1$. Thus, by Proposition 7, $\boldsymbol{f}^*$ is a Σ_{cx}-smallest element in $\mathcal{SB}_d$. It remains to prove the last part of the theorem. First, we note that if $\boldsymbol{f} = \lambda \boldsymbol{f}^* + (1-\lambda)\boldsymbol{f}^{K*}$, for $\lambda \in (0,1] \cap \mathbb{Q}$ and $\boldsymbol{f}^{K*} \in \mathcal{PB}^*_d$, then $\boldsymbol{f}$ is a Σ_{cx}-smallest pmf in $\mathcal{SB}_d$. Conversely, suppose that $\boldsymbol{f}$ is a Σ_{cx}-smallest pmf such that $\mathcal{H}(\boldsymbol{f}) = Q(\boldsymbol{z}) \in [P^*(\boldsymbol{z})]$. From Proposition 4, we have that $\boldsymbol{f} = \lambda \boldsymbol{f}^* + (1-\lambda)\boldsymbol{f}^K$, for some $\lambda \in (0,1] \cap \mathbb{Q}$ and $\boldsymbol{f}^K \in \mathcal{PB}_d$. Since $f(\boldsymbol{x}) = f^*(\boldsymbol{x}) = 0$ if $\boldsymbol{x} \notin \mathcal{X}^*_d$, we have that $\boldsymbol{f}^K$ is a Σ_{cx}-smallest pmf, i.e., $\boldsymbol{f}^K \in \mathcal{PB}^*_d$. The arguments for the case of even d are analogous. □

Appendix C. The class of symmetric Bernoulli distributions: Examples

Example 8. We consider the case $d = 3$, $\mathcal{SB}_3$. We have $\boldsymbol{a} = Q\boldsymbol{f}$, where

$$\boldsymbol{a} = \begin{pmatrix} a_{00} \\ a_{10} \\ a_{01} \\ a_{11} \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{f} = \left(f_{000}\ f_{100}\ f_{010}\ f_{110}\ f_{001}\ f_{101}\ f_{011}\ f_{111}\right)^\top.$$

Therefore, for $d = 3$, the polynomials in $\mathcal{C}_{\mathcal{H}} \subseteq \mathcal{I}_{\mathcal{P}}$ are of the form $P(\boldsymbol{z}) = a_{00} + a_{10}z_1 + a_{01}z_2 + a_{11}z_1z_2$, where

$$a_{00} = f_{000} - f_{111}, \quad a_{10} = f_{100} - f_{011}, \quad a_{01} = f_{010} - f_{101}, \quad a_{11} = f_{110} - f_{001}.$$

The polynomials in $\mathcal{I}_{\mathcal{P}}$ vanish at points $\mathbf{1}_{d-1} = (1,1)$, $\mathbf{1}_{d-1}^{-1} = (-1,1)$ and $\mathbf{1}_{d-1}^{-2} = (1,-1)$. For example, let us consider the following pmfs and their corresponding polynomials $\mathcal{H}(\boldsymbol{f})$:

$$\begin{aligned}
\boldsymbol{f}_1 &= (0.3, 0.1, 0.1, 0, 0.1, 0, 0, 0.4), & P_1(\boldsymbol{z}) &= -0.1 + 0.1z_1 + 0.1z_2 - 0.1z_1z_2;\\
\boldsymbol{f}_2 &= (0.1, 0.1, 0.1, 0.2, 0.3, 0, 0, 0.2), & P_2(\boldsymbol{z}) &= -0.1 + 0.1z_1 + 0.1z_2 - 0.1z_1z_2;\\
\boldsymbol{f}_3 &= (0, 0.25, 0.25, 0, 0.25, 0, 0, 0.25), & P_3(\boldsymbol{z}) &= -0.25 + 0.25z_1 + 0.25z_2 - 0.25z_1z_2;\\
\boldsymbol{f}_4 &= (0.25, 0, 0, 0.25, 0, 0.25, 0.25, 0), & P_4(\boldsymbol{z}) &= 0.25 - 0.25z_1 - 0.25z_2 + 0.25z_1z_2.
\end{aligned}$$

The polynomials $P_1(\boldsymbol{z})$ and $P_2(\boldsymbol{z})$ are identical, while $P_3(\boldsymbol{z}) = 5P_1(\boldsymbol{z})/2$ and $P_4(\boldsymbol{z}) = -P_3(\boldsymbol{z})$.

Example 9. Consider the class $\mathcal{B}_3(2/5)$ of 3-dimensional Bernoulli distributions with common marginal mean 2/5. The pmf $\boldsymbol{r} = (0, 1/5, 1/5, 1/5, 2/5, 0, 0, 0)$ is an extremal pmf of $\mathcal{B}_3(2/5)$ (see Example 1 in [20]) and it is not the type-0 pmf associated to its polynomial. Indeed, we have $\mathcal{H}(\boldsymbol{r}) = Q(\boldsymbol{z}) = -1/5 + z_1/5 + z_2/5 - z_1z_2/5$ and, by the algorithm to find type-0 pmfs proposed in [20], which holds for any $p \in [0,1] \cap \mathbb{Q}$, the type-0 pmf associated to $Q(\boldsymbol{z})$ is $\boldsymbol{f} = (0, 3/10, 3/10, 0, 3/10, 0, 0, 1/10)$. Thus, $\boldsymbol{r}$ is an extremal pmf, but it is not a type-0 pmf. It is straightforward to see that we can write

$$\boldsymbol{r} = \frac{2}{3}\boldsymbol{f} + \frac{1}{3}\boldsymbol{f}^K,$$

where $2\boldsymbol{f}/3$ is the non-normalized type-0 pmf associated to $Q(\boldsymbol{z})$ and $\boldsymbol{f}^K/3$ is an element of the kernel of the linear map defined in Section 3 in [20] between $\mathbb{R}^{2^d}$ and $\mathbb{R}^{2^{d-1}}$. However, it can be verified that $\boldsymbol{f}^K$ has negative components and, therefore, it is not a pmf.

References

[1] M. Denuit, C. Genest, É. Marceau, Stochastic bounds on sums of dependent risks, Insurance Math. Econom. 25 (1) (1999) 85–104.
[2] P. Embrechts, G. Puccetti, Bounds for functions of dependent risks, Finance Stoch. 10 (2006) 341–352.
[3] T. Lux, A. Papapantoleon, Improved Fréchet–Hoeffding bounds on d-copulas and applications in model-free finance, Ann. Appl. Probab. (2017) 3633–3671.
[4] G. Puccetti, L. Rüschendorf, Sharp bounds for sums of dependent risks, J. Appl. Probab. 50 (1) (2013) 42–53.
[5] R. Wang, L. Peng, J. Yang, Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities, Finance Stoch. 17 (2013) 395–417.
[6] R. Kaas, J. Dhaene, M.J. Goovaerts, Upper and lower bounds for sums of random variables, Insurance Math. Econom. 27 (2) (2000) 151–168.
[7] H. Joe, Multivariate Models and Multivariate Dependence Concepts, Chapman & Hall, London, 1997.
[8] G. Puccetti, R. Wang, Extremal dependence concepts, Statist. Sci. 30 (4) (2015) 485–517.
[9] N.R. Chaganty, H. Joe, Range of correlation matrices for dependent Bernoulli random variables, Biometrika 93 (1) (2006) 197–206.
[10] B. Dai, S. Ding, G. Wahba, Multivariate Bernoulli distribution, Bernoulli 19 (4) (2013) 1465–1483.
[11] T. Hu, C. Xie, L. Ruan, Dependence structures of multivariate Bernoulli random vectors, J. Multivariate Anal. 94 (1) (2005) 172–195.
[12] G.M. Marchetti, N. Wermuth, Palindromic Bernoulli distributions, Electron. J. Stat. 10 (2) (2016) 2435–2460.
[13] R.B. Nelsen, An Introduction to Copulas, second ed., Springer, New York, 2006.
[14] P. Embrechts, Copulas: A personal view, J. Risk Insur. 76 (3) (2009) 639–650.
[15] P. Embrechts, A. Höing, A. Juri, Using copulae to bound the value-at-risk for functions of dependent risks, Finance Stoch. 7 (2003) 145–167.
[16] B.V. de Melo Mendes, R.M. de Souza, Measuring financial risks with copulas, Int. Rev. Financ. Anal. 13 (1) (2004) 27–45.
[17] X. Zhang, H. Jiang, Application of copula function in financial risk analysis, Comput. Electr. Eng. 77 (2019) 376–388.
[18] C. Bernard, X. Jiang, R. Wang, Risk aggregation with dependence uncertainty, Insurance Math. Econom. 54 (2014) 93–108.
[19] T. Hu, Z. Wu, On dependence of risks and stop-loss premiums, Insurance Math. Econom. 24 (3) (1999) 323–332.
[20] R. Fontana, P. Semeraro, High dimensional Bernoulli distributions: Algebraic representation and applications, Bernoulli 30 (1) (2024) 825–850.
[21] C. Bernard, L. Rüschendorf, S. Vanduffel, J. Yao, How robust is the value-at-risk of credit risk portfolios? Eur. J. Financ. 23 (6) (2017) 507–534.
[22] W. Lee, K.C. Cheung, J.Y. Ahn, Multivariate countermonotonicity and the minimal copulas, J. Comput. Appl. Math. 317 (2017) 589–602.
[23] H. Cossette, E. Marceau, A. Mutti, P. Semeraro, Extremal negative dependence and the strongly Rayleigh property, 2025, arXiv preprint arXiv:2504.17679.
[24] J. Dhaene, M. Denuit, The safest dependence structure among risks, Insurance Math. Econom. 25 (1) (1999) 11–21.
[25] W. Lee, J.Y. Ahn, On the multidimensional extension of countermonotonicity and its applications, Insurance Math. Econom. 56 (2014) 68–79.
[26] L. Rüschendorf, Random variables with maximum sums, Adv. in Appl. Probab. 14 (3) (1982) 623–632.
[27] A.J. McNeil, J. Nešlehová, A.D. Smith, On attainability of Kendall's tau matrices and concordance signatures, J. Multivariate Anal. 191 (2022) 105033.
[28] E.-M. Tiit, Mixtures of multivariate quasi-extremal distributions having given marginals, Lect. Notes-Monogr. Ser. (1996) 337–357.
[29] C. Blier-Wong, H. Cossette, E. Marceau, Stochastic representation of FGM copulas using multivariate Bernoulli random variables, Comput. Statist. Data Anal. 173 (2022) 107506.
[30] H. Cossette, E. Marceau, A. Mutti, P. Semeraro, Generalized FGM dependence: Geometrical representation and convex bounds on sums, 2024, arXiv preprint arXiv:2406.10648.
[31] R. Fontana, P. Semeraro, Representation of multivariate Bernoulli distributions with a given set of specified moments, J. Multivariate Anal. 168 (2018) 290–303.
[32] R. Fontana, E. Luciano, P. Semeraro, Model risk in credit risk, Math. Finance 31 (1) (2021) 176–202.
[33] J.-G. Lauzier, L. Lin, R. Wang, Pairwise counter-monotonicity, Insurance Math. Econom. 111 (2023) 279–287.
[34] G. Dall'Aglio, Fréchet classes and compatibility of distribution functions, in: Symposia Mathematica, vol. 9, 1972, pp. 131–150.
[35] E. Frostig, Comparison of portfolios which depend on multivariate Bernoulli random variables with fixed marginals, Insurance Math. Econom. 29 (3) (2001) 319–331.
[36] B. Wang, R. Wang, The complete mixability and convex minimization problems with monotone marginal densities, J. Multivariate Anal. 102 (10) (2011) 1344–1360.
[37] M. Terzer, Large Scale Methods to Enumerate Extreme Rays and Elementary Modes (Ph.D. thesis), ETH Zürich, 2009.
[38] M. Shaked, J.G. Shanthikumar, Stochastic Orders, Springer, New York, 2007.
[39] N. Gaffke, L. Rüschendorf, On a class of extremal problems in statistics, Math. Oper. Stat. Ser. Optim. 12 (1) (1981) 123–135.
[40] S. Cambanis, Some properties and generalizations of multivariate Eyraud–Gumbel–Morgenstern distributions, J. Multivariate Anal. 7 (4) (1977) 551–559.
[41] C. Blier-Wong, H. Cossette, E. Marceau, Exchangeable FGM copulas, Adv. in Appl. Probab. 56 (1) (2024) 205–234.

[42] C. Genest, J. Nešlehová, A primer on copulas for count data, ASTIN Bullettin: J. IAA 37 (2) (2007) 475–515.
[43] I. Gijbels, V. Kika, M. Omelka, On the specification of multivariate association measures and their behaviour with increasing dimension, J. Multivariate Anal. 182 (2021) 104704.
[44] S. Saminger-Platz, A. Kolesárová, A. Šeliga, R. Mesiar, E.P. Klement, The impact on the properties of the EFGM copulas when extending this family, Fuzzy Sets and Systems 415 (2021) 1–26.
[45] M. Huber, N. Marić, Admissible Bernoulli correlations, J. Stat. Distrib. Appl. 6 (2019) 1–8.
[46] E. Boros, A. Prékopa, Closed form two-sided bounds for probabilities that at least r and exactly r out of n events occur, Math. Oper. Res. 14 (2) (1989) 317–342.