ALEA, Lat. Am. J. Probab. Math. Stat. 20, 1077–1089 (2023)
DOI: 10.30757/ALEA.v20-39

Characterizations of multivariate distributions with limited memory revisited: An analytical approach

Natalia Shenkman

Department of Mathematics, Technische Universität München, 85748 Garching bei München, Germany
E-mail address: shenkman@tum.de, natalia.shenkman@hotmail.com

Abstract. Alternative proofs of the characterizations of the wide-sense geometric and of the Marshall-Olkin exponential distributions via monotone set functions are provided. In contrast to the ones presented in Shenkman (2017), which rely on the generative constructions of Arnold (1975) or Marshall and Olkin (1967) to establish that certain functions equipped with monotone parameters are proper survival functions, we aim herein to check that these candidates satisfy a set of well known necessary and sufficient analytical conditions. The major difficulty in such an approach consists in verifying that they do not infringe any of the so-called rectangle inequalities. Fortunately, a factorization shows that compliance is guaranteed as long as a finite number of very specific "basis" rectangle inequalities are not violated: a condition which is, by the very definition of the monotone parameters, trivially met.

1. Introduction

Distribution theory, as a main pillar of science, enjoys a very special status within the realm of mathematics. While univariate distributions have and continue to contribute heavily to this success, the transition from difficult probabilistic situations in real life to mathematical problems often require the involvement of multivariate distributions. In the sequel, we will focus our attention on two intimately related particular types of such distributions. On the one hand, we have the most famous multidimensional extension of the exponential distribution, which at the time, thanks to its characterization via both intuitive generative constructions based on shocks and its fulfillment of the so-called lack of memory property, bridged a relevant knowledge gap for the modeling of risk profiles of systems constituted of many parts interacting with each other; see Marshall and Olkin (1967). More precisely, denoting $\mathscr{S}_d := \{1,\dots,d\}$, let $\{\lambda_I\}_{\emptyset\neq I\subseteq\mathscr{S}_d}$ be a collection of non-negative parameters satisfying $\sum_{I:i\in I}\lambda_I > 0$ for all $i\in\mathscr{S}_d$, and let $\{E_I\}_{\emptyset\neq I\subseteq\mathscr{S}_d}$ be independent exponentially distributed random variables, where $E_I \sim Exp(\lambda_I)$. Then, the random vector $\boldsymbol{\tau} = (\tau_1,\dots,\tau_d)$ taking values in $\mathbb{R}^d_{>0}$ defined by

$$\tau_i = \min\big\{E_I :\ i\in I\big\},\quad i\in\mathscr{S}_d, \tag{1.1}$$

Received by the editors March 22nd, 2020; accepted July 15th, 2023.
2000 *Mathematics Subject Classification.* 60E05, 62H05, 62E10.
Key words and phrases. Marshall-Olkin exponential distribution, multivariate geometric distribution, lack of memory, rectangle probabilities.

is said to follow the d-variate Marshall-Olkin exponential ($\mathcal{MO}_d$) distribution with parameters $\{\lambda_I\}_{\emptyset\neq I\subseteq\mathscr{S}_d}$. Moreover, a random vector $\boldsymbol{\tau}$ supported on $\mathbb{R}^d_{>0}$ follows a $\mathcal{MO}_d$ distribution if and only if it satisfies the lack of memory (LM) property

$$\mathbb{P}\big(\tau_i > t_i + s,\, i \in I\big) = \mathbb{P}\big(\tau_i > t_i,\, i \in I\big)\, \mathbb{P}\big(\tau_i > s,\, i \in I\big) \tag{1.2}$$

for all $I \subseteq \mathscr{S}_d$ and all $\mathbf{t} \in \mathbb{R}^d_+$, $s > 0$, where $\mathbb{R}_+ = [0,\infty)$. On the other hand, we have the more complicated discrete analogue of the Marshall-Olkin distribution; see Esary and Marshall (1973) and Arnold (1975). Both articles propose their own discrete probabilistic model to account for the richer structure of the distribution. Taking the shock model of Arnold (1975) and borrowing the terminology related to the equivalent generative construction presented in Esary and Marshall (1973) leads us to a formal definition. Namely, let $\{\boldsymbol{\chi}^{(n)}\}_{n\in\mathbb{N}}$ be a sequence of d-variate independent and identically Bernoulli distributed random vectors with associated probabilities $\{p_I\}_{I\subseteq\mathscr{S}_d}$ satisfying $\sum_{I\subseteq\mathscr{S}_d} p_I = 1$ and $\sum_{I:i\notin I} p_I < 1$ for all $i \in \mathscr{S}_d$. Then, the random vector $\widetilde{\boldsymbol{\tau}} = (\widetilde{\tau}_1, \ldots, \widetilde{\tau}_d)$ with values in $\mathbb{N}^d$ defined by

$$\widetilde{\tau}_i = \min\big\{n \in \mathbb{N} :\, \chi_i^{(n)} = 0\big\}, \quad i \in \mathscr{S}_d, \tag{1.3}$$

is said to follow the d-variate wide-sense geometric ($\mathcal{G}_d^{\mathcal{W}}$) distribution with parameters $\{p_I\}_{I\subseteq\mathscr{S}_d}$. While in the continuous case the solution of the functional equation (1.2) can be found inductively by inspection as it involves the fundamental multiplicative identity of the exponential function, a completely different concept is needed to make up for the loss of the exponential ansatz when the LM property is only discrete, i.e., when (1.2) only holds for all $I \subseteq \mathscr{S}_d$ and all $\mathbf{t} \in \mathbb{N}_0^d$, $s \in \mathbb{N}$, where $\mathbb{N} = \{1,2,\ldots\}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. Surprisingly, although Azlarov and Volodin (1983) came up with the right idea to solve the problem in dimension 2, it took many decades before it became mathematical knowledge that a random vector $\widetilde{\boldsymbol{\tau}}$ with values in $\mathbb{N}^d$ satisfies the discrete LM property if and only if it follows a $\mathcal{G}_d^{\mathcal{W}}$ distribution; see Mai et al. (2013a,b). Rather than pointing in the direction of a notable open problem, we believe that such a time scale highlights the nonchalance of experts towards a priori purely theoretical questions on complicated distributions without any concrete published applications; cf. Esary and Marshall (1973), Arnold (1975), Azlarov and Volodin (1983), Marshall and Olkin (1991) and Balakrishnan and Lai (2009, p. 452f.). However, sparked off by Mai and Scherer (2009a,b, 2011, 2012), who studied, though initially without realizing, the exchangeable Marshall-Olkin copula reparametrized with some novel monotone parameters, Mai et al. (2013a,b) have set the ball rolling as the firsts of a series of articles devoted to elevating the reparametrizations as relevant unconventional sophisticated technical tools. Indeed, when properly harnessed, they may offer a strong guidance through the multidimensional maze arising from any attempt to shed light on intrinsic properties of the $\mathcal{MO}_d$ and $\mathcal{G}_d^{\mathcal{W}}$ distributions: compare, e.g., Shenkman (2017) with Esary and Marshall (1974), Langberg et al. (1977), and Balakrishnan and Lai (2009, p. 215f.), or Shenkman (2020) with Mai and Scherer (2009a,b, 2011, 2012). This brings us to the prime concern of the present paper. In the literature, all characterizations of the $\mathcal{MO}_d$ and $\mathcal{G}_d^{\mathcal{W}}$ distributions via their survival functions are achieved using a probabilistic model equivalent to (1.1) or (1.3) seemingly corroborating the statement of Joe (1997, p. 11) that a purely analytical approach is often difficult to pursue if not hopeless; see Marshall and Olkin (1967), Esary and Marshall (1973), Galambos and Kotz (1978, Ch. 5), Mai et al. (2013b), Shenkman (2017). However, we will demonstrate that, quite remarkably, embarking on such an enterprise in the context of the distributions with limited memory is by no means doomed thanks to their monotone parameters.

2. Characterizations of the $\mathcal{G}_d^{\mathcal{W}}$ and $\mathcal{MO}_d$ survival functions

Before stating the characterization theorem, we need to recall from Shenkman (2017) some results on monotone set functions which will serve as parameters for the $\mathcal{G}_d^{\mathcal{W}}$ and $\mathcal{MO}_d$ distributions.

2.1. *Prerequisites on monotone set functions.* Let μ be a real-valued function defined on the subsets of $\mathscr{S}_d$. For any $T \subseteq \mathscr{S}_d$ and $i \in \overline{T}$, where $\overline{T}$ denotes the complement of T in $\mathscr{S}_d$, we introduce the forward difference

$$\nabla_i \, \mu(T) := \mu(T) - \mu\big(T \cup \{i\}\big).$$

For each non-empty subset I of $\overline{T}$, say $I = \{i_1, \ldots, i_{|I|}\}$, define recursively

$$\nabla_I \, \mu(T) := \nabla_{i_{|I|}} \cdots \nabla_{i_1} \mu(T) = \sum_{J \subseteq I} (-1)^{|J|} \, \mu\big(T \cup J\big)$$

and set $\nabla_\emptyset \, \mu(T) := \mu(T)$. The function μ is said to be d-monotone if it satisfies

$$\nabla_I \, \mu(T) \geq 0 \quad \forall \, T \subseteq \mathscr{S}_d \text{ and } \forall \, I \subseteq \overline{T}, \; I \neq \emptyset, \tag{2.1}$$

or, equivalently, it is d-monotone if $\nabla_{\overline{T}} \, \mu(T) \geq 0$ for all $T \subsetneq \mathscr{S}_d$; see Lemma 2.2 in Shenkman (2017).

Definition 2.1 ($\mathcal{M}_d$ and $\mathcal{LM}_d$)**.** We define $\mathcal{M}_d$ to be the set of all d-monotone functions $\mu : \mathcal{P}(\mathscr{S}_d) \to [0, 1]$ which satisfy

$$\mu(\emptyset) = 1, \tag{2.2a}$$

$$\mu\big(\{i\}\big) < 1 \quad \forall i \in \mathscr{S}_d. \tag{2.2b}$$

Moreover, we let $\mathcal{LM}_d$ be the set of all functions $\mu : \mathcal{P}(\mathscr{S}_d) \to (0, 1]$ which satisfy (2.2a) and (2.2b), and for which the composite function $\ln \circ \mu$ is d-monotone.

History shows that it was difficult to relate the sets $\mathcal{M}_d$ and $\mathcal{LM}_d$ to the $\mathcal{G}_d^{\mathcal{W}}$ and $\mathcal{MO}_d$ distributions. While the links are implicitly present in Esary and Marshall (1973, p. 20), in Formulas (2.1) and (2.8) in Langberg et al. (1977), in Theorems 4.2 and 3.1 in Mai and Scherer (2009b,a), respectively, and in Theorem 4.1 in Sun et al. (2017), an explicit assignment was, despite a decisive impetus given by Gnedin and Pitman (2008), still only reached in stages; see Mai and Scherer (2009a, 2011), Mai et al. (2013b), and Shenkman (2017).

The relationship between the sets $\mathcal{M}_d$ and $\mathcal{LM}_d$ is made clear by our next proposition. It builds on the work of Bennett (1992) and Rhoades (1963) on totally monotone sequences; see Theorem 8 in Bennett (1992).

Proposition 2.2. *Let $\mu : \mathcal{P}(\mathscr{S}_d) \to (0, 1]$. Then, $\mu \in \mathcal{LM}_d$ if and only if $\mu^s \in \mathcal{M}_d$ for all $s > 0$.*

Proof: Refer to Appendix A. □

Recalling that a random vector $\boldsymbol{\tau}$ with survival function S is said to be min-infinitely divisible (min-id) if $S^{1/n}$ is itself a survival function for all $n \in \mathbb{N}$, Proposition 2.3 shows that $\mathcal{M}_d$ and $\mathcal{LM}_d$ are intimately related to the d-variate (min-id) Bernoulli distributions.

Proposition 2.3. *Let $\mu : \mathcal{P}(\mathscr{S}_d) \to [0, 1]$. Then, $\mu \in \mathcal{M}_d$ if and only if there exists a random vector $\boldsymbol{\chi}$ with values in $\{0, 1\}^d$ such that $\mathbb{P}(\chi_i = 1) < 1$ for all $i \in \mathscr{S}_d$ and*

$$\mu(T) = \mathbb{P}\big(\chi_i = 1, \, i \in T\big), \quad T \subseteq \mathscr{S}_d. \tag{2.3}$$

Additionally, $\mu \in \mathcal{LM}_d$ if and only if $\boldsymbol{\chi}$ is min-id with $\mathbb{P}(\boldsymbol{\chi} = \mathbf{1}) > 0$.

Proof: Combining Proposition 2.2 with Propositions 2.4 and 3.7 in Shenkman (2017) yields the claim. □

2.2. *Analytical derivation of the $\mathcal{G}_d^{\mathcal{W}}$ and $\mathcal{MO}_d$ survival functions.* Shenkman (2017) showed that the survival function of the $\mathcal{G}_d^{\mathcal{W}}$, resp. $\mathcal{MO}_d$, distribution can be derived directly from the LM property (1.2). For the convenience of the reader, we go over the argumentation for the $\mathcal{MO}_d$ distribution and remind that it covers implicitly the case of the $\mathcal{G}_d^{\mathcal{W}}$ distribution. Let $\boldsymbol{\tau}$ follow a $\mathcal{MO}_d$ distribution and denote the ordered list of $t_1, \dots, t_d \in \mathbb{R}_+$ by $t_{(1)} \leq \dots \leq t_{(d)}$ with the convention $t_{(0)} := 0$. Further, for each $\mathbf{t} \in \mathbb{R}_+^d$, let $\pi_{\mathbf{t}}\colon \mathscr{S}_d \to \mathscr{S}_d$ be a permutation depending on $\mathbf{t}$ such that $t_{\pi_{\mathbf{t}}(1)} \leq \dots \leq t_{\pi_{\mathbf{t}}(d)}$. Then, relying solely on (1.2), the survival function of $\boldsymbol{\tau}$ can be written for all $\mathbf{t} \in \mathbb{R}_+^d$ as

$$
\begin{aligned}
S(\mathbf{t}) &= \mathbb{P}\big(\tau_1 > t_1, \dots, \tau_d > t_d\big) = \mathbb{P}\big(\tau_{\pi_{\mathbf{t}}(1)} > t_{(1)}, \dots, \tau_{\pi_{\mathbf{t}}(d)} > t_{(d)}\big) \\
&= \mathbb{P}\big(\tau_{\pi_{\mathbf{t}}(1)} > t_{(1)}, \dots, \tau_{\pi_{\mathbf{t}}(d)} > t_{(1)}\big)\mathbb{P}\big(\tau_{\pi_{\mathbf{t}}(2)} > t_{(2)} - t_{(1)}, \dots, \tau_{\pi_{\mathbf{t}}(d)} > t_{(d)} - t_{(1)}\big) \\
&= \mathbb{P}\big(\tau_{\pi_{\mathbf{t}}(1)} > t_{(1)}, \dots, \tau_{\pi_{\mathbf{t}}(d)} > t_{(1)}\big)\mathbb{P}\big(\tau_{\pi_{\mathbf{t}}(2)} > t_{(2)} - t_{(1)}, \dots, \tau_{\pi_{\mathbf{t}}(d)} > t_{(2)} - t_{(1)}\big) \\
&\quad \times \mathbb{P}\big(\tau_{\pi_{\mathbf{t}}(3)} > t_{(3)} - t_{(2)}, \dots, \tau_{\pi_{\mathbf{t}}(d)} > t_{(d)} - t_{(2)}\big) \\
&= \prod_{i=1}^{d} \mathbb{P}\big(\tau_{\pi_{\mathbf{t}}(i)} > t_{(i)} - t_{(i-1)}, \dots, \tau_{\pi_{\mathbf{t}}(d)} > t_{(i)} - t_{(i-1)}\big),
\end{aligned}
$$

i.e., the probability of the event that the component τ_i survives t_i units of time for all $i \in \mathscr{S}_d$ can be factored into the product of d probabilities: the probability that all the components survive the first $t_{(1)}$ units of time multiplied by the probability that the components with indices in $\{\pi_{\mathbf{t}}(2), \dots, \pi_{\mathbf{t}}(d)\}$ survive additional $t_{(2)} - t_{(1)}$ units of time, etc. Again thanks to (1.2), we obtain for all $I \subseteq \mathscr{S}_d$ and all $m, n \in \mathbb{N}$, that

$$
\mathbb{P}\Big(\tau_i > \frac{m}{n},\, i \in I\Big)^n = \mathbb{P}\big(\tau_i > 1,\, i \in I\big)^m.
$$

In fact, bearing in mind that $\mathbb{P}\big(\tau_i > 0,\, i \in \mathscr{S}_d\big) = 1$, the equality holds for all $m \in \mathbb{N}_0$. From the density of the rationals in $\mathbb{R}$ and the right-continuity of the survival function, it follows that

$$
\mathbb{P}\big(\tau_i > t,\, i \in I\big) = \mathbb{P}\big(\tau_i > 1,\, i \in I\big)^t \tag{2.4}
$$

for all $I \subseteq \mathscr{S}_d$ and all $t \in \mathbb{R}_+$. So, we have

$$
\begin{aligned}
S(\mathbf{t}) &= \prod_{i=1}^{d} \mathbb{P}\big(\tau_{\pi_{\mathbf{t}}(i)} > 1, \dots, \tau_{\pi_{\mathbf{t}}(d)} > 1\big)^{t_{(i)} - t_{(i-1)}} \\
&= \prod_{i=1}^{d} \mu\big(\{\pi_{\mathbf{t}}(i), \dots, \pi_{\mathbf{t}}(d)\}\big)^{t_{(i)} - t_{(i-1)}},
\end{aligned} \tag{2.5}
$$

where μ is a function defined on the subsets of $\mathscr{S}_d$ by

$$
\mu(T) = \mathbb{P}\big(\tau_i > 1,\, i \in T\big), \quad T \subseteq \mathscr{S}_d.
$$

The $2^d - 1$ values $\{\mu(T)\}_{\emptyset \neq T \subseteq \mathscr{S}_d}$ uniquely determine S via (2.5), which yields that at any point in time, the probability that a group of components of $\boldsymbol{\tau}$ survives one additional unit of time is governed by a unique d-variate Bernoulli distribution of the random vector $\boldsymbol{\chi}$ given by

$$
\chi_i := \begin{cases} 0, & \text{if } \tau_i \leq 1, \\ 1, & \text{if } \tau_i > 1, \end{cases} \quad i \in \mathscr{S}_d.
$$

The survival probabilities of $\boldsymbol{\chi}$ are completely determined by μ via (2.3). Since for all $n \in \mathbb{N}$, the random vector $n\boldsymbol{\tau}$ has survival function $S^{1/n}$, we see that for each $n \in \mathbb{N}$, $\mu^{1/n}$ itself defines the survival probabilities of a d-variate Bernoulli distribution, which, in turn, implies that $\boldsymbol{\chi}$ is min-id. As $\boldsymbol{\tau}$ is almost surely finite, we obtain for all $i \in \mathscr{S}_d$, that $\mathbb{P}(\chi_i = 1) < 1$. Moreover, from (2.4) together with the fact that $S(\mathbf{0}) = 1$ and the right-continuity of the survival function, we deduce that $\mathbb{P}(\boldsymbol{\chi} = \mathbf{1}) > 0$. Thus, in view of Proposition 2.3, $\mu \in \mathcal{LM}_d$.

2.3. *Characterization theorem.* The characterizations of the $\mathcal{G}_d^{\mathcal{W}}$ and $\mathcal{MO}_d$ survival functions with the classical parameters of their shock models (1.3) and (1.1) are without any doubt essential. However, the alternative characterizations given in terms of parameters belonging to the sets $\mathcal{M}_d$ and $\mathcal{LM}_d$, already partially introduced in Section 2.2 and stated formally in Theorem 2.4, have a lot to offer.

Theorem 2.4 (Characterizations of the $\mathcal{G}_d^{\mathcal{W}}$ and $\mathcal{MO}_d$ survival functions)**.**

(*i*) *A function $S : \mathbb{R}_+^d \to [0,1]$ defines a $\mathcal{MO}_d$ survival function if and only if there exists $\mu \in \mathcal{LM}_d$ such that for all $\mathbf{t} \in \mathbb{R}_+^d$, it holds that*

$$S(\mathbf{t}) = \prod_{i=1}^{d} \Big\{ \mu\big(\{\pi_{\mathbf{t}}(i), \ldots, \pi_{\mathbf{t}}(d)\}\big) \Big\}^{t_{(i)} - t_{(i-1)}}. \tag{2.6}$$

(*ii*) *A function $S : \mathbb{N}_0^d \to [0,1]$ defines a $\mathcal{G}_d^{\mathcal{W}}$ survival function if and only if there exists $\mu \in \mathcal{M}_d$ such that (2.6) holds for all $\mathbf{t} \in \mathbb{N}_0^d$.*

Proof: The proofs of necessity for the $\mathcal{G}_d^{\mathcal{W}}$ and $\mathcal{MO}_d$ survival functions were already given in Shenkman (2017) but repeated in Section 2.2 for completeness. Sufficiency is shown in Section 3.2 for $\mathcal{MO}_d$ and in Section 3.3 for $\mathcal{G}_d^{\mathcal{W}}$. As opposed to the standard practice, we will not rely on any probabilistic model: The relationships between the parameters in the sets $\mathcal{M}_d$ and $\mathcal{LM}_d$ and the ones in the constructions (1.1) and (1.3) can be found in Shenkman (2017). □

Theorem 3.5 given in Mai et al. (2013b) was identified as the tool of choice to carry out a thorough study of the exchangeable $\mathcal{G}_d^{\mathcal{W}}$ distribution; see Mai et al. (2013b, p. 463). Theorem 2.4 extends its restricted scope of application since the indisputable benefit of this technical tool, originally specialized for the exchangeable $\mathcal{G}_d^{\mathcal{W}}$ distribution, is brought to bear on all the distributions with limited memory. This way, many known results, prominent or not, previously perceived to be tedious to derive, have been relegated to the class of low-hanging fruits. For instance, the second part of the proof of Lemma 2.1 in Shenkman (2020) illustrates how straightforward Lemmas 2.1 and 3.1 in Mai and Scherer (2013) and Mai et al. (2013b) really are when considered in the context of the above reparametrizations. Additionally, bringing this tool into play yields quite a few new results and proof techniques such as the one we are about to discover in Section 3.

3. Analytical proofs of sufficiency in Theorem 2.4

McNeil and Nešlehová (2009) characterized the multivariate Archimedean copulas using an approach which relies on an analogue of the characterization of cumulative distribution functions via their analytical properties for survival functions; see, e.g., Joe (1997, p. 11). Accordingly, finding out whether a given function is a survival function requires, among other things, to check that it satisfies the d-decreasingness condition on hyperrectangles of all possible sizes; see Lemma 1 in McNeil and Nešlehová (2009). Lemma 3.1 below, an adaption of this result, allows to only consider hypercubes of certain sizes as its proof takes into account the σ-additivity of probability measures. Whereas the authors were not particularly affected by the seemingly stronger condition, it would have been an insurmountable hurdle for the proof technique we present in Section 3.2.

3.1. *A technical lemma.* Before proceeding, let us introduce some notation: Inequalities between vectors are understood componentwise, $\mathbf{1}_I : \mathscr{S}_d \to \{0,1\}$ denotes the indicator function of a set $I \subseteq \mathscr{S}_d$, and for any $\mathbf{t} \in \mathbb{R}^d$, $\|\mathbf{t}\|_1 := \sum_{i=1}^d |t_i|$.

Lemma 3.1. *A function $S : \mathbb{R}_+^d \to [0,1]$ is a survival function of a probability measure on $\mathbb{R}_{>0}^d$ if and only if*

(i) $S(\mathbf{0}) = 1$ *and* $\lim_{t_i \to \infty} S(\mathbf{t}) = 0$ *for each* $i \in \mathscr{S}_d$ *and all* $t_j \geq 0$, $j \in \mathscr{S}_d \backslash \{i\}$,

(ii) S *is right-continuous:* $\forall\, \mathbf{t} \in \mathbb{R}_+^d$ *one has that* $\forall\, \epsilon \in (0,1)\ \exists\, \delta > 0$ *s.t.* $\forall\, \mathbf{s} \geq \mathbf{t}$

$$\|\mathbf{s} - \mathbf{t}\|_1 < \delta \quad \Rightarrow \quad |S(\mathbf{s}) - S(\mathbf{t})| < \epsilon,$$

(iii) (rectangle inequality) for any $\mathbf{t} \in \mathbb{R}_+^d$, *there exists* $\tilde{s} > 0$ *s.t. for all* $s \in (0, \tilde{s}]$

$$\sum_{I \subseteq \mathscr{S}_d} (-1)^{|I|} S\Big(t_1 + s \cdot \mathbf{1}_I(1), \ldots, t_d + s \cdot \mathbf{1}_I(d)\Big) \;\geq\; 0.$$

Proof: Lemma 1 in McNeil and Nešlehová (2009) asserts that the d-decreasingness of S, i.e., for all $\mathbf{t} \in \mathbb{R}_+^d$ and all $\mathbf{s} \in \mathbb{R}_{>0}^d$,

$$\sum_{I \subseteq \mathscr{S}_d} (-1)^{|I|} S\Big(t_1 + s_1 \cdot \mathbf{1}_I(1), \ldots, t_d + s_d \cdot \mathbf{1}_I(d)\Big) \;\geq\; 0,$$

together with the conditions (i) and (ii) are necessary and sufficient in order for S to define a survival function of a probability measure on $\mathbb{R}_{>0}^d$. To conclude the proof, we will show that conditions (ii) and (iii) imply the d-decreasingness of S. Thanks to the right-continuity of S, it follows for all $\mathbf{t} \in \mathbb{R}_+^d$ and all $\mathbf{s} \in \mathbb{R}_{>0}^d$, that

$$S\Big(t_1 + s_1, \ldots, t_d + s_d\Big) = \lim_{n \to \infty} S\left(t_1 + \frac{\lceil n s_1 \rceil}{n}, \ldots, t_d + \frac{\lceil n s_d \rceil}{n}\right),$$

where $\lceil \cdot \rceil$ denotes the ceiling function. Hence, we obtain

$$\begin{aligned}
&\sum_{I \subseteq \mathscr{S}_d} (-1)^{|I|} S\Big(t_1 + s_1 \cdot \mathbf{1}_I(1), \ldots, t_d + s_d \cdot \mathbf{1}_I(d)\Big) \\
&= \lim_{n \to \infty} \sum_{I \subseteq \mathscr{S}_d} (-1)^{|I|} S\left(t_1 + \frac{\lceil n s_1 \rceil}{n} \mathbf{1}_I(1), \ldots, t_d + \frac{\lceil n s_d \rceil}{n} \mathbf{1}_I(d)\right) \\
&= \lim_{n \to \infty} \sum_{i_1=0}^{\lceil n s_1 \rceil - 1} \cdots \sum_{i_d=0}^{\lceil n s_d \rceil - 1} \sum_{I \subseteq \mathscr{S}_d} (-1)^{|I|} S\left(t_1 + \frac{i_1}{n} + \frac{1}{n} \mathbf{1}_I(1), \ldots, t_d + \frac{i_d}{n} + \frac{1}{n} \mathbf{1}_I(d)\right) \geq 0,
\end{aligned}$$

where the last equality holds since we have decomposed the hyperrectangle

$$\left(t_1, t_1 + \frac{\lceil n s_1 \rceil}{n}\right] \times \ldots \times \left(t_d, t_d + \frac{\lceil n s_d \rceil}{n}\right]$$

into $\prod_{i=1}^{d} \lceil n s_i \rceil$ hypercubes of length $\frac{1}{n}$. □

3.2. *Main result: Proof of sufficiency in Theorem 2.4* (i). We show that $S : \mathbb{R}_+^d \to [0,1]$ given as in (2.6) satisfies conditions $(i) - (iii)$ of Lemma 3.1 if $\mu \in \mathcal{LM}_d$, whereby checking the non-infringement of condition (iii) is the difficult part.

$S(\mathbf{0}) = 1$ is easily seen while $\lim_{t_i \to \infty} S(\mathbf{t}) = 0$ for all $i \in \mathscr{S}_d$ simply follows from the fact that

$$\mu(I) < 1, \quad \forall\, I \subseteq \mathscr{S}_d,\ I \neq \emptyset. \tag{3.1}$$

To see why (3.1) holds, note first that $\mu\big(\{i\}\big) < 1$ for all $i \in \mathscr{S}_d$ by (2.2b). Striving for a contradiction, assume that (3.1) is not true. Then, there exists $I \subseteq \mathscr{S}_d$ with $|I| > 1$ such that $\mu(I) \geq 1$ and $\mu(T) < 1$ for all $T \subsetneq I$. For any $i \in I$, it follows from the d-monotonicity of $\ln \circ \mu$ that $\ln \mu\big(I \backslash \{i\}\big) - \ln \mu(I) \geq 0$. Therefore, $\mu(I) \leq \mu\big(I \backslash \{i\}\big) < 1$, which is a contradiction. This gives us (i).

We now verify (ii). Let $\mathbf{t} \in \mathbb{R}_+^d$ and $\epsilon \in (0,1)$ be given. Order the elements of $\mathbf{t}$ such that $t_{(1)} = \ldots = t_{(l_1)} < t_{(l_1+1)} = \ldots = t_{(l_2)} < \ldots < t_{(l_{m-1}+1)} = \ldots = t_{(l_m)}$, where $l_m = d$, and define

$$M_{\mathbf{t}} := \begin{cases} \min\limits_{i \in \{2,\ldots,m\}} \big\{ t_{(l_i)} - t_{(l_{i-1})} \big\}, & \text{if } l_1 < d, \\ 1, & \text{if } l_1 = d. \end{cases}$$

Introduce

$$I_i := \big\{ \pi_{\mathbf{t}}(i), \ldots, \pi_{\mathbf{t}}(d) \big\}, \quad i \in \mathscr{S}_d,$$

and set $I_{d+1} := \emptyset$. Next, let

$$L := \min_{i \in \{1,\ldots,d\}} \left\{ \frac{\mu\big(I_i\big)}{\mu\big(I_{i+1}\big)} \right\}.$$

Notice that $L \in (0,1)$ since μ is positive and $L \leq \mu\big(\{\pi_{\mathbf{t}}(d)\}\big) < 1$. Finally, define

$$\delta := \min \left\{ M_{\mathbf{t}}, \frac{\ln(1-\epsilon)}{\ln L} \right\}.$$

Then, because $\delta \leq M_{\mathbf{t}}$, we have for all $\mathbf{s} \geq \mathbf{t}$ with $\|\mathbf{s} - \mathbf{t}\|_1 < \delta$, that $s_i - t_i < M_{\mathbf{t}}$ for all $i \in \mathscr{S}_d$, and, hence, there exist $\pi_{\mathbf{s}}$ and $\pi_{\mathbf{t}}$ satisfying $\pi_{\mathbf{s}}(i) = \pi_{\mathbf{t}}(i)$ for all $i \in \mathscr{S}_d$. It follows that

$$\begin{aligned} |S(\mathbf{s}) - S(\mathbf{t})| &= \left| \prod_{i=1}^{d} \mu^{s_{(i)} - s_{(i-1)}}\big(I_i\big) - \prod_{j=1}^{d} \mu^{t_{(j)} - t_{(j-1)}}\big(I_j\big) \right| \\ &= \prod_{j=1}^{d} \mu^{t_{(j)} - t_{(j-1)}}\big(I_j\big) \left(1 - \prod_{i=1}^{d} \left(\frac{\mu\big(I_i\big)}{\mu\big(I_{i+1}\big)} \right)^{s_{(i)} - t_{(i)}} \right) \\ &\leq 1 - \prod_{i=1}^{d} \left(\frac{\mu\big(I_i\big)}{\mu\big(I_{i+1}\big)} \right)^{s_{(i)} - t_{(i)}} \leq 1 - L^{\|\mathbf{s}-\mathbf{t}\|_1} < 1 - L^{\delta} \leq \epsilon \end{aligned}$$

as required.

To show (iii), we will transform, for arbitrary given $\mathbf{t} \in \mathbb{R}_+^d$ and $s \in (0, M_{\mathbf{t}}]$, a sum of 2^d terms of the type $S\big(\mathbf{t} + s \cdot \mathbf{j}\big)$, $\mathbf{j} \in \{0,1\}^d$, into a product of non-negative terms. Keeping in mind the ordering of the elements of $\mathbf{t}$ given in the proof of (ii), observe that $t_{(l_i)} + s \leq t_{(l_{i+1})}$ for all $i = 1, \ldots, m-1$. Set $l_0 := 0$. For any $\mathbf{j} \in \{0,1\}^d$ and each $i = 0, 1, \ldots, m-1$, let $j_{[l_i+1]} \leq j_{[l_i+2]} \leq \ldots \leq j_{[l_{i+1}]}$ denote the ordered list of $j_{\pi_{\mathbf{t}}(l_i+1)}, j_{\pi_{\mathbf{t}}(l_i+2)}, \ldots, j_{\pi_{\mathbf{t}}(l_{i+1})}$. Notice that in this way we have sorted all the elements of the vector $(\mathbf{t} + s \cdot \mathbf{j})$ in ascending order and, in particular, for all $i = 0, 1, \ldots, m-1$, we have that

$$t_{(l_i+1)} + s \cdot j_{[l_i+1]} \;\leq\; t_{(l_i+2)} + s \cdot j_{[l_i+2]} \;\leq\; \ldots \;\leq\; t_{(l_{i+1})} + s \cdot j_{[l_{i+1}]}.$$

Setting $j_{[0]} := 0$ and remembering that an empty product is by convention equal to 1, we obtain from Formula (2.6) that

$$\begin{aligned} S\big(\mathbf{t} + s \cdot \mathbf{j}\big) = \prod_{i=0}^{m-1} \Bigg(& \mu^{t_{(l_i+1)} - t_{(l_i)} + s(j_{[l_i+1]} - j_{[l_i]})}\big(\{\pi_{\mathbf{t}+s\mathbf{j}}(l_i+1), \ldots, \pi_{\mathbf{t}+s\mathbf{j}}(d)\}\big) \\ & \times \prod_{k=l_i+2}^{l_{i+1}} \mu^{s(j_{[k]} - j_{[k-1]})}\big(\{\pi_{\mathbf{t}+s\mathbf{j}}(k), \ldots, \pi_{\mathbf{t}+s\mathbf{j}}(d)\}\big) \Bigg). \end{aligned} \tag{3.2}$$

Next, define

$$\kappa := \mu(\mathscr{S}_d)^{t_{(1)}} \prod_{i=1}^{m-1} \mu^{t_{(l_i+1)}-t_{(l_i)}-s}\big(\{\pi_{\mathbf{t}+s\mathbf{j}}(l_i+1),\ldots,\pi_{\mathbf{t}+s\mathbf{j}}(d)\}\big). \tag{3.3}$$

Making use of (3.2) together with (3.3), we now infer that

$$\begin{aligned}
D(\mathbf{t},s) &:= \sum_{I\subseteq\mathscr{S}_d} (-1)^{|I|} S\Big(t_1+s\cdot\mathbf{1}_I(1),\ldots,t_d+s\cdot\mathbf{1}_I(d)\Big)\\
&= \sum_{j_1=0}^{1}\cdots\sum_{j_d=0}^{1}(-1)^{j_1+\ldots+j_d} S\big(\mathbf{t}+s\cdot\mathbf{j}\big)\\
&= \kappa\sum_{j_1=0}^{1}\cdots\sum_{j_d=0}^{1}(-1)^{j_1+\ldots+j_d}\mu^{-s}(\mathscr{S}_d)\\
&\quad\times\prod_{i=0}^{m-1}\Bigg(\mu^{s\left(j_{[l_i+1]}-j_{[l_i]}+1\right)}\big(\{\pi_{\mathbf{t}+s\mathbf{j}}(l_i+1),\ldots,\pi_{\mathbf{t}+s\mathbf{j}}(d)\}\big)\\
&\quad\times\prod_{k=l_i+2}^{l_{i+1}}\mu^{s\left(j_{[k]}-j_{[k-1]}\right)}\big(\{\pi_{\mathbf{t}+s\mathbf{j}}(k),\ldots,\pi_{\mathbf{t}+s\mathbf{j}}(d)\}\big)\Bigg).
\end{aligned}$$

Applying the conventions $l_0=0$ and $j_{[0]}=0$ in the first equality, as well as $l_m=d$, $\{\pi_{\mathbf{t}+s\mathbf{j}}(d+1),\ldots,\pi_{\mathbf{t}+s\mathbf{j}}(d)\}=\emptyset$, and $\mu(\emptyset)=1$ in the last equality, we may write

$$\begin{aligned}
&\mu^{-s}(\mathscr{S}_d)\prod_{i=0}^{m-1}\mu^{s\left(1-j_{[l_i]}\right)}\big(\{\pi_{\mathbf{t}+s\mathbf{j}}(l_i+1),\ldots,\pi_{\mathbf{t}+s\mathbf{j}}(d)\}\big)\\
&\quad=\prod_{i=1}^{m-1}\mu^{s\left(1-j_{[l_i]}\right)}\big(\{\pi_{\mathbf{t}+s\mathbf{j}}(l_i+1),\ldots,\pi_{\mathbf{t}+s\mathbf{j}}(d)\}\big)\\
&\quad=\prod_{i=0}^{m-2}\mu^{s\left(1-j_{[l_{i+1}]}\right)}\big(\{\pi_{\mathbf{t}+s\mathbf{j}}(l_{i+1}+1),\ldots,\pi_{\mathbf{t}+s\mathbf{j}}(d)\}\big)\\
&\quad=\prod_{i=0}^{m-1}\mu^{s\left(1-j_{[l_{i+1}]}\right)}\big(\{\pi_{\mathbf{t}+s\mathbf{j}}(l_{i+1}+1),\ldots,\pi_{\mathbf{t}+s\mathbf{j}}(d)\}\big).
\end{aligned}$$

Hence, using the above computation together with the fact that for all $i = 0, 1, \ldots, m-1$, we have $\{\pi_{\mathbf{t}+s\mathbf{j}}(l_i+1), \ldots, \pi_{\mathbf{t}+s\mathbf{j}}(d)\} = \{\pi_{\mathbf{t}}(l_i+1), \ldots, \pi_{\mathbf{t}}(d)\}$, it follows that

$$
\begin{aligned}
D(\mathbf{t}, s) &= \kappa \sum_{j_1=0}^{1} \cdots \sum_{j_d=0}^{1} (-1)^{j_1+\ldots+j_d} \prod_{i=0}^{m-1} \Bigg(\mu^{s j_{[l_i+1]}}\big(\{\pi_{\mathbf{t}}(l_i+1), \ldots, \pi_{\mathbf{t}}(d)\}\big) \\
&\qquad \times \mu^{s\left(1-j_{[l_{i+1}]}\right)}\big(\{\pi_{\mathbf{t}}(l_{i+1}+1), \ldots, \pi_{\mathbf{t}}(d)\}\big) \\
&\qquad \times \prod_{k=l_i+2}^{l_{i+1}} \mu^{s\left(j_{[k]}-j_{[k-1]}\right)}\big(\{\pi_{\mathbf{t}+s\mathbf{j}}(k), \ldots, \pi_{\mathbf{t}+s\mathbf{j}}(d)\}\big) \Bigg) \\
&= \kappa \prod_{i=0}^{m-1} \sum_{j_{\pi_{\mathbf{t}}(l_i+1)}=0}^{1} \cdots \sum_{j_{\pi_{\mathbf{t}}(l_{i+1})}=0}^{1} (-1)^{j_{\pi_{\mathbf{t}}(l_i+1)}+\ldots+j_{\pi_{\mathbf{t}}(l_{i+1})}} \\
&\qquad \times \mu^{s j_{[l_i+1]}}\big(\{\pi_{\mathbf{t}}(l_i+1), \ldots, \pi_{\mathbf{t}}(d)\}\big) \mu^{s\left(1-j_{[l_{i+1}]}\right)}\big(\{\pi_{\mathbf{t}}(l_{i+1}+1), \ldots, \pi_{\mathbf{t}}(d)\}\big) \\
&\qquad \times \prod_{k=l_i+2}^{l_{i+1}} \mu^{s\left(j_{[k]}-j_{[k-1]}\right)}\big(\{\pi_{\mathbf{t}+s\mathbf{j}}(k), \ldots, \pi_{\mathbf{t}+s\mathbf{j}}(d)\}\big).
\end{aligned}
$$

Lastly, observe that for each fixed $i \in \{0, 1, \ldots, m-1\}$ and each associated ordered list of binary indices $0 \le j_{[l_i+1]} \le j_{[l_i+2]} \le \ldots \le j_{[l_{i+1}]} \le 1$,

$$
\begin{aligned}
&\mu^{s j_{[l_i+1]}}\big(\{\pi_{\mathbf{t}}(l_i+1), \ldots, \pi_{\mathbf{t}}(d)\}\big) \mu^{s\left(1-j_{[l_{i+1}]}\right)}\big(\{\pi_{\mathbf{t}}(l_{i+1}+1), \ldots, \pi_{\mathbf{t}}(d)\}\big) \\
&\qquad \times \prod_{k=l_i+2}^{l_{i+1}} \mu^{s\left(j_{[k]}-j_{[k-1]}\right)}\big(\{\pi_{\mathbf{t}+s\mathbf{j}}(k), \ldots, \pi_{\mathbf{t}+s\mathbf{j}}(d)\}\big) \\
&= \mu^{s}\big(\{\pi_{\mathbf{t}}(l_{i+1}+1), \ldots, \pi_{\mathbf{t}}(d)\} \cup J_i\big),
\end{aligned}
$$

where $J_i \subseteq \{\pi_{\mathbf{t}}(l_i+1), \pi_{\mathbf{t}}(l_i+2), \ldots, \pi_{\mathbf{t}}(l_{i+1})\}$ is such that for all $k \in \{\pi_{\mathbf{t}}(l_i+1), \pi_{\mathbf{t}}(l_i+2), \ldots, \pi_{\mathbf{t}}(l_{i+1})\}$, we have that $k \in J_i$ if and only if $\mathbf{j}_k = 1$. Therefore, we get the desired factorization (which is, not without reminding of the factorization obtained in Shenkman (2022)):

$$
\begin{aligned}
D(\mathbf{t}, s) &= \kappa \prod_{i=0}^{m-1} \sum_{J_i \subseteq \{\pi_{\mathbf{t}}(l_i+1), \ldots, \pi_{\mathbf{t}}(l_{i+1})\}} (-1)^{|J_i|} \mu^{s}\big(\{\pi_{\mathbf{t}}(l_{i+1}+1), \ldots, \pi_{\mathbf{t}}(d)\} \cup J_i\big) \\
&= \kappa \prod_{i=0}^{m-1} \nabla_{\{\pi_{\mathbf{t}}(l_i+1), \ldots, \pi_{\mathbf{t}}(l_{i+1})\}} \mu^{s}\big(\{\pi_{\mathbf{t}}(l_{i+1}+1), \ldots, \pi_{\mathbf{t}}(d)\}\big) \ \ge\ 0,
\end{aligned} \tag{3.4}
$$

where the inequality is due to Proposition 2.2. This gives us (iii) and also that S is the survival function of a probability measure on $\mathbb{R}^d_{>0}$. To conclude, we remark that S satisfies the LM property (1.2), i.e., a simple computation shows that $S(t_1+s, \ldots, t_d+s) = S(t_1, \ldots, t_d) S(s, \ldots, s)$ for all $\mathbf{t} \in \mathbb{R}^d_+$ and all $s > 0$. □

Expression (3.4) for the S-volume of the hypercube $\prod_{i=1}^{d}(t_i, t_i+s]$ shows how the sets $\mathcal{M}_d$ and $\mathcal{LM}_d$ come into play to guarantee the d-decreasingness of the survival functions. That the S-volume of an arbitrary hypercube can be written as a product of a finite number of S-volumes of "basis" hypercubes is a very special feature, which, to the best of our knowledge, was for the first time observed for the exchangeable $\mathcal{G}_d^{\mathcal{W}}$ distribution in Mai et al. (2013a). For instance, in dimension 2, the "basis" consists of three S-volumes given by $\nabla_I \mu(\emptyset)$, $\emptyset \neq I \subseteq \mathscr{S}_2$: trivially, we get, e.g.,

$$\begin{aligned} V_s\big((0,1] \times (0,1]\big) &= S(0,0) - S(1,0) - S(0,1) + S(1,1) \\ &= \mu(\emptyset) - \mu(\{1\}) - \mu(\{2\}) + \mu(\{1,2\}) \\ &= \nabla_{\{1,2\}} \mu(\emptyset) \geq 0. \end{aligned}$$

In an exemplary manner, this easy computation brings us to discuss an aspect so far neglected, namely the importance of special and low dimensional cases. Since we have not considered a single one of them in this work and repeatedly praised the tractability of the distributions, it may appear that we claim that, at least in theory, to produce such a proof by himself, a skilled mathematician could forgo their investigation and improvise from the very beginning in high dimension without ever needing to look down or sideways. This view does not reflect at all our experience and, therefore, we encourage any interested reader to attempt to engage with the "simpler" and yet distinctive case of the exchangeable $\mathcal{G}_d^{\mathcal{W}}$ distribution in order to get, in a sense, an unbiased feel for the complexity of such an enterprise: All the ingredients required to start the exercise and its complete solution are in Mai et al. (2013a), although it must be said that the general proof given there can be drastically shortened with the help of the factor κ to account in one go for all possible constellations. Alternatively, one can simply assess, by consulting Remark 2.4 in Mai et al. (2013a), which presents the case of the exchangeable $\mathcal{G}_2^{\mathcal{W}}$ distribution in great detail, whether the indispensable intuition and know-how needed to envision and shape a way to solve the multidimensional problem could realistically originate from other sources of inspiration than its thorough study in low dimensions.

3.3. *Proof of sufficiency in Theorem 2.4* (*ii*). In view of Section 3.2, the proof of sufficiency in Theorem 2.4 (*ii*) is now almost just a formality. We make use of Lemma 2.1 in Mai et al. (2013a), which is a discrete equivalent of Lemma 3.1. Observe that $S(\mathbf{0}) = 1$. Further, by d-monotonicity of μ we have that $\mu(I) \leq \mu\big(\{i\}\big)$ for all $\emptyset \neq I \subseteq \mathscr{S}_d$ and any $i \in I$. In view of (2.2b), this gives $\mu(I) < 1$ for all $\emptyset \neq I \subseteq \mathscr{S}_d$. As a result, $\lim_{t_i \to \infty} S(\mathbf{t}) = 0$ for all $i \in \mathscr{S}_d$. Hence, it remains to show that S assigns non-negative mass to each point $\mathbf{t} \in \mathbb{N}^d$, or, equivalently, that

$$\sum_{I \subseteq \mathscr{S}_d} (-1)^{|I|} S\Big(t_1 + \mathbf{1}_I(1), \ldots, t_d + \mathbf{1}_I(d)\Big) \;\geq\; 0$$

for all $\mathbf{t} \in \mathbb{N}_0^d$. However, we have already shown this in the proof of Section 3.2. Therefore, S is the survival function of a probability measure on $\mathbb{N}^d$ which satisfies the discrete LM property (1.2). $\square$

A. Appendix: Proof of Proposition 2.2

Let μ be a positive function on the subsets of $\mathscr{S}_d$ satisfying $\mu^s \in \mathcal{M}_d$ for all $s > 0$. Then, repeating the arguments of Shenkman (2017), for all $n \in \mathbb{N}$, all $T \subseteq \mathscr{S}_d$, and all $\emptyset \neq I \subseteq \overline{T}$, one has

$$\nabla_I \big(\mu^{1/n}(T) - 1 \big) = \nabla_I \, \mu^{1/n}(T) \geq 0,$$

where the equality is due to $I \neq \emptyset$. Using Euler's limit representation of the logarithm, it follows that

$$\nabla_I \, \ln \mu(T) = \lim_{n \to \infty} n \, \nabla_I \big\{ \mu^{1/n}(T) - 1 \big\} \geq 0,$$

which, in turn, gives $\mu \in \mathcal{LM}_d$.

To see that the converse is true, notice first that if $\mu \in \mathcal{LM}_d$, then $\mu^s \in \mathcal{LM}_d$ for all $s > 0$ since for all $T \subseteq \mathscr{S}_d$ and all $\emptyset \neq I \subseteq \overline{T}$, we have $\nabla_I \ln \mu^s(T) = s \nabla_I \ln \mu(T)$. Hence, we only need to show that $\mu \in \mathcal{LM}_d$ implies $\mu \in \mathcal{M}_d$. Before getting started, we need to extract a result from the proof of Proposition 2.6 in Shenkman (2017). For completeness we recall it here: For any two set functions μ and ν on the subsets of $\mathscr{S}_d$, we have

$$\nabla_I\Big(\mu(T)\,\nu(T)\Big) = \sum_{J\subseteq I}\Big(\nabla_{I\cap\overline{J}}\,\mu\big(T\cup J\big)\Big)\Big(\nabla_J\,\nu(T)\Big) \tag{A.1}$$

for all $T\subseteq\mathscr{S}_d$ and $I\subseteq\overline{T}$. (A.1) is shown by induction on the cardinality of I. Note that for $I=\emptyset$, (A.1) holds trivially. Fix an arbitrary $T\subsetneq\mathscr{S}_d$. Observe first that for all $i\in\overline{T}$,

$$\nabla_i\Big(\mu(T)\,\nu(T)\Big) = \Big(\nabla_i\,\mu(T)\Big)\,\nu(T)\;+\;\mu\big(T\cup\{i\}\big)\nabla_i\,\nu(T). \tag{A.2}$$

For the induction step, assume that (A.1) is true for all $I\subseteq\overline{T}$ satisfying $|I|\le n$ for some $0\le n<\big|\overline{T}\big|$. Take an arbitrary $I\subseteq\overline{T}$ with $|I|=n+1$ and fix an $i\in I$. Then,

$$\begin{aligned}
\nabla_I\Big(\mu(T)\nu(T)\Big) &= \nabla_i\nabla_{I\setminus\{i\}}\Big(\mu(T)\nu(T)\Big) = \nabla_i\sum_{J\subseteq I\setminus\{i\}}\Big(\nabla_{I\cap\overline{\{i\}}\cap\overline{J}}\,\mu\big(T\cup J\big)\Big)\Big(\nabla_J\,\nu(T)\Big)\\
&= \sum_{J\subseteq I\setminus\{i\}}\Big(\nabla_{I\cap\overline{J}}\,\mu\big(T\cup J\big)\,\nabla_J\,\nu(T)\;+\;\nabla_{I\cap\overline{\{i\}}\cap\overline{J}}\,\mu\big(T\cup J\cup\{i\}\big)\,\nabla_{J\cup\{i\}}\,\nu(T)\Big)\\
&= \sum_{\substack{J\subseteq I\\ i\notin J}}\nabla_{I\cap\overline{J}}\,\mu\big(T\cup J\big)\,\nabla_J\,\nu(T)\;+\;\sum_{\substack{J\subseteq I\\ i\in J}}\nabla_{I\cap\overline{J}}\,\mu\big(T\cup J\big)\,\nabla_J\,\nu(T)\\
&= \sum_{J\subseteq I}\nabla_{I\cap\overline{J}}\,\mu\big(T\cup J\big)\,\nabla_J\,\nu(T),
\end{aligned}$$

where the second equality is due to the induction hypothesis, and the third equality follows by applying (A.2) to the product of $\widetilde{\mu}(T)=\nabla_{I\cap\overline{\{i\}}\cap\overline{J}}\,\mu\big(T\cup J\big)$ and $\widetilde{\nu}(T)=\nabla_J\,\nu(T)$, $T\subseteq\mathscr{S}_d$, for each fixed $J\subseteq I\setminus\{i\}$.

Now let $\mu\in\mathcal{LM}_d$. Then,

$$\nabla_i\,\mu(T) = \left(\frac{\mu(T)}{\mu\big(T\cup\{i\}\big)}-1\right)\mu\big(T\cup\{i\}\big)\;\ge\;0$$

for all $T\subsetneq\mathscr{S}_d$ and $i\in\overline{T}$, where the inequality follows by making use of $\nabla_i\ln\mu(T)\ge 0$ and of the positivity of μ. Striving for a contradiction, assume (2.1) is not true. Then there exists an $n\in\mathbb{N}$ satisfying $1\le n<d$ such that

$$\nabla_I\,\mu(T)\ge 0\quad\forall\,T\subseteq\mathscr{S}_d\text{ and }\forall\,I\subseteq\overline{T},\;I\neq\emptyset,\;|I|\le n, \tag{A.3}$$

but where (2.1) does not hold for some $T\subseteq\mathscr{S}_d$ and some $\emptyset\neq J\subseteq\overline{T}$ with $|J|=n+1$. Take any $i\in J$. Then, by application of (A.1),

$$\begin{aligned}
\nabla_J\,\mu(T)\; &= \nabla_{J\setminus\{i\}}\left(\left(\frac{\mu(T)}{\mu\big(T\cup\{i\}\big)}-1\right)\mu\big(T\cup\{i\}\big)\right)\\
&= \sum_{I\subseteq J\setminus\{i\}}\nabla_{(J\setminus\{i\})\cap\overline{I}}\left(\frac{\mu\big(T\cup I\big)}{\mu\big(T\cup\{i\}\cup I\big)}-1\right)\nabla_I\,\mu\big(T\cup\{i\}\big),
\end{aligned}$$

where $\nabla_I\,\mu(T\cup\{i\})\ge 0$ for all $I\subseteq J\setminus\{i\}$ by (A.3). Hence, we will obtain a contradiction if we can show that

$$\nabla_I\left(\frac{\mu(K)}{\mu\big(K\cup\{j\}\big)}-1\right)\;\ge\;0 \tag{A.4}$$

for all $K \subsetneq \mathscr{S}_d$, all $j \in \overline{K}$, and all $I \subseteq \overline{K}\backslash\{j\}$. To this end, observe first that

$$\nabla_I \ln\left(\frac{\mu(K)}{\mu\big(K \cup \{j\}\big)}\right) = \nabla_{I\cup\{j\}} \ln \mu(K) \;\geq\; 0$$

for all $K \subsetneq \mathscr{S}_d$, all $j \in \overline{K}$, and all $I \subseteq \overline{K}\backslash\{j\}$. Applying inductively the product rule (A.1), we deduce that

$$\nabla_I \ln^m\left(\frac{\mu(K)}{\mu\big(K \cup \{j\}\big)}\right) \;\geq\; 0 \tag{A.5}$$

for all $m \in \mathbb{N}$, all $K \subsetneq \mathscr{S}_d$, all $j \in \overline{K}$, and all $I \subseteq \overline{K}\backslash\{j\}$. Moreover, for $m = 0$, the left-hand side of (A.5) equals 1 if $I = \emptyset$, and 0 otherwise. Using the power series representation of the exponential function, we obtain

$$\nabla_I\left(\frac{\mu(K)}{\mu\big(K \cup \{j\}\big)}\right) = \sum_{m=0}^{\infty} \frac{1}{m!} \nabla_I \ln^m\left(\frac{\mu(K)}{\mu\big(K \cup \{j\}\big)}\right) \geq 0$$

for all $K \subsetneq \mathscr{S}_d$, all $j \in \overline{K}$, and all $I \subseteq \overline{K}\backslash\{j\}$. Consequently, we get (A.4) for all $\emptyset \neq I \subseteq \overline{K}\backslash\{j\}$. Having in mind that for all $j \in \overline{K}$, $\big(\mu(K)/\mu(K \cup \{j\})\big) \geq 1$, we infer that (A.4) actually holds for all $K \subsetneq \mathscr{S}_d$, all $j \in \overline{K}$, and all $I \subseteq \overline{K}\backslash\{j\}$. This provides the desired contradiction and concludes the proof. $\square$

Acknowledgements

I sincerely thank an associate editor for the valuable comments on a first version of this manuscript. In particular, addressing the substantiated criticism that the article was geared towards a confidential audience gave me the chance to facilitate the dissemination of its content. Finally, I acknowledge having adopted the wordings "sophisticated technical tools" and "generative constructions" of the referee and of the associate editor to describe the reparametrizations and the probabilistic constructions, respectively.

References

Arnold, B. C. A characterization of the exponential distribution by multivariate geometric compounding. *Sankhyā Ser. A*, **37** (1), 164–173 (1975). MR440792.

Azlarov, T. A. and Volodin, N. A. On the discrete analog of Marshall-Olkin's distribution. In *Stability problems for stochastic models (Moscow, 1982)*, volume 982 of *Lecture Notes in Math.*, pp. 17–23. Springer, Berlin (1983). MR715457.

Balakrishnan, N. and Lai, C.-D. *Continuous bivariate distributions*. Springer, Dordrecht, second edition (2009). ISBN 978-0-387-09613-1. MR2840643.

Bennett, G. Lower bounds for matrices. II. *Canad. J. Math.*, **44** (1), 54–74 (1992). MR1152666.

Esary, J. D. and Marshall, A. W. Multivariate geometric distributions generated by a cumulative damage process. Technical report, Naval Postgraduate School, Monterey, California (1973). Available online at https://calhoun.nps.edu/handle/10945/30003.

Esary, J. D. and Marshall, A. W. Multivariate distributions with exponential minimums. *Ann. Statist.*, **2**, 84–98 (1974). MR362704.

Galambos, J. and Kotz, S. *Characterizations of probability distributions. A unified approach with an emphasis on exponential and related models*, volume 675 of *Lecture Notes in Mathematics*. Springer, Berlin (1978). ISBN 3-540-08933-0. MR513423.

Gnedin, A. and Pitman, J. Moments of convex distribution functions and completely alternating sequences. In *Probability and statistics: essays in honor of David A. Freedman*, volume 2 of *Inst. Math. Stat. (IMS) Collect.*, pp. 30–41. Inst. Math. Statist., Beachwood, OH (2008). MR2459948.

Joe, H. *Multivariate models and dependence concepts*, volume 73 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London (1997). ISBN 0-412-07331-5. MR1462613.

Langberg, N., Proschan, F., and Quinzi, A. J. Converting dependent models into independent ones, with applications in reliability. In *The theory and applications of reliability, with emphasis on Bayesian and nonparametric methods, Vol. I (Conf., Univ. South Florida, Tampa, Fla., 1975)*, pp. 259–276. Academic Press, New York-London (1977). MR451596.

Mai, J.-F. and Scherer, M. Lévy-frailty copulas. *J. Multivariate Anal.*, **100** (7), 1567–1585 (2009a). MR2514148.

Mai, J.-F. and Scherer, M. A tractable multivariate default model based on a stochastic time-change. *Int. J. Theor. Appl. Finance*, **12** (2), 227–249 (2009b). MR2522454.

Mai, J.-F. and Scherer, M. Reparameterizing Marshall-Olkin copulas with applications to sampling. *J. Stat. Comput. Simul.*, **81** (1), 59–78 (2011). MR2747378.

Mai, J.-F. and Scherer, M. *Simulating copulas. Stochastic models, sampling algorithms, and applications*, volume 4 of *Series in Quantitative Finance*. Imperial College Press, London (2012). ISBN 978-1-84816-874-9; 1-84816-874-8. MR2906392.

Mai, J.-F. and Scherer, M. Extendibility of Marshall-Olkin distributions and inverse Pascal triangles. *Braz. J. Probab. Stat.*, **27** (3), 310–321 (2013). MR3064726.

Mai, J.-F., Scherer, M., and Shenkman, N. An Analytical Characterization of the Exchangeable Wide-Sense Geometric Law. In Kruse, R. et al., editors, *Synergies of Soft Computing and Statistics for Intelligent Data Analysis*, volume 190, pp. 331–340. Springer Berlin Heidelberg (2013a). ISBN 978-3-642-33042-1. DOI: 10.1007/978-3-642-33042-1_36.

Mai, J.-F., Scherer, M., and Shenkman, N. Multivariate geometric distributions, (logarithmically) monotone sequences, and infinitely divisible laws. *J. Multivariate Anal.*, **115**, 457–480 (2013b). MR3004570.

Marshall, A. W. and Olkin, I. A multivariate exponential distribution. *J. Amer. Statist. Assoc.*, **62**, 30–44 (1967). MR215400.

Marshall, A. W. and Olkin, I. Functional equations for multivariate exponential distributions. *J. Multivariate Anal.*, **39** (1), 209–215 (1991). MR1128681.

McNeil, A. J. and Nešlehová, J. Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. *Ann. Statist.*, **37** (5B), 3059–3097 (2009). MR2541455.

Rhoades, B. E. A sufficient condition for total monotonicity. *Trans. Amer. Math. Soc.*, **107**, 309–319 (1963). MR147810.

Shenkman, N. A natural parametrization of multivariate distributions with limited memory. *J. Multivariate Anal.*, **155**, 234–251 (2017). MR3607893.

Shenkman, N. A primer on the characterization of the exchangeable Marshall-Olkin copula via monotone sequences. *Braz. J. Probab. Stat.*, **34** (1), 127–135 (2020). MR4058974.

Shenkman, N. Min-infinite divisibility of the bivariate Marshall-Olkin copulas. *Comm. Statist. Theory Methods*, **51** (1), 226–231 (2022). MR4361632.

Sun, Y., Mendoza-Arriaga, R., and Linetsky, V. Marshall-Olkin distributions, subordinators, efficient simulation, and applications to credit risk. *Adv. in Appl. Probab.*, **49** (2), 481–514 (2017). MR3668386.