Technische Universität München
TUM School of Computation, Information and Technology

Stochastic representations of Marshall–Olkin distributions and upper semilinear copulas

Henrik Sloot

Vollständiger Abdruck der von der TUM School of Computation, Information and Technology der Technischen Universität München zur Erlangung eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Felix Krahmer

Prüfer*innen der Dissertation:

1. Prof. Dr. Matthias Scherer
2. Prof. Dr. Fabrizio Durante
3. Prof. Dr. Wolfgang Trutschnig

Die Dissertation wurde am 26.06.2023 bei der Technischen Universität München eingereicht und durch die TUM School of Computation, Information and Technology am 15.11.2023 angenommen.

Zusammenfassung

Diese Dissertation entwickelt stoch. Darstellungen für upper semilinear Copulas und extendible Marshall-Olkin-Verteilungen, was probabilistische Ansätze und Simulationsalgorithmen ermöglicht. Sie leitet die stoch. Darstellungen von drei USL Copula–Unterklassen her, entwickelt einen effizienten, numerisch stabilen Simulationsalgorithmus für hochdimensionale extendible MO Verteilungen und charakterisiert Survival-Funktionen und de Finetti Darstellungen für generalized MO Verteilungen.

Abstract

This dissertation develops stochastic representations for upper semilinear copulas and extendible Marshall-Olkin distributions, enabling probabilistic approaches and simulation algorithms. It derives three upper semilinear copula subclasses' stochastic representations, proposes an efficient, numerically stable simulation algorithm for high-dimensional extendible Marshall–Olkin, and characterizes survival functions and de Finetti representations for generalized Marshall-Olkin distributions.

Acknowledgements

I would like to express my sincere gratitude to all those who have supported and encouraged me throughout my academic journey. First and foremost, I would like to thank my wife Alexandra for her unwavering support and encouragement that have been essential in helping me persevere through the challenges of completing my dissertation. I am also grateful to my parents, Siegbert and Cilli, and my sister Lena for their love and encouragement.

I am also deeply grateful to my supervisor Matthias Scherer, whose guidance and mentorship have been invaluable. His expertise and encouragement have been instrumental in shaping my research and personal growth.

Finally, I want to extend my thanks to my colleagues, in particular Gabi Zeller, Lexuri Fernández, and Florian Brück, for their valuable insights and collaboration. It has been a pleasure working alongside such talented and dedicated individuals, and I am grateful for the friendships and connections that have been forged along the way.

List of contributed articles

Core articles as principal author

[1] H. Sloot and M. Scherer. “A probabilistic view on semilinear copulas”. In: *Information Sciences* 512 (2020), pp. 258–276. DOI: 10.1016/j.ins.2019.09.069.

[2] H. Sloot. “Implementing Markovian models for extendible Marshall–Olkin distributions”. In: *Dependence Modeling* 10.1 (2022), pp. 308–343. DOI: 10.1515/demo-2022-0151.

Further articles as principal author

[3] M. Scherer and H. Sloot. “Exogenous shock models: analytical characterization and probabilistic construction”. In: *Metrika* 82.8 (2019), pp. 931–959. DOI: 10.1007/s00184-019-00715-8.

[4] H. Sloot. “The deFinetti representation of generalised Marshall–Olkin sequences”. In: *Dependence Modeling* 8.1 (2020), pp. 107–118. DOI: 10.1515/demo-2020-0006.

Further articles as co-author

[5] D. Brigo et al. “Consistent iterated simulation of multivariate defaults: Markov indicators, lack of memory, extreme-value copulas, and the Marshall–Olkin distribution”. In: *Innovations in Insurance, Risk- and Asset Management*. Ed. by K. Glau et al. Conference at the Technical University of Munich, 5–7 April 2017. Singapore: World Scientific, 2018, pp. 47–93. DOI: 10.1142/9789813272569_0003.

I, Henrik Sloot, am the principal author of the articles [1–4] and a co-author of the article [5]. Detailed descriptions of the individual contributions can be found following the respective article summaries in Appendices A to C.

Contents

1 Introduction

This thesis investigates various aspects of modeling multivariate random vectors with parametric distributions. High-dimensional parametric models are at the core of many real-world applications, for example, portfolio credit risk management. Practitioners and researchers have the challenging task of finding suitable models that are well understood while meeting the requirements of their applications. A good understanding of the used distributions is paramount, as even those with similar margins and pairwise correlation coefficients can be very different in nature. Additionally, distributions are not equally tractable. For example, Monte-Carlo simulations require an efficient simulation strategy or at least a known stochastic representation. This thesis provides novel stochastic representations, efficient simulation strategies, and theoretical advances of three classes of high-dimensional families of distributions: Marshall–Olkin distributions, generalized Marshall–Olkin distributions, and upper semilinear copulas.

Multivariate stochastic modeling often relies on the copula approach motivated by Sklar's copula separation theorem suggested in [6]. This approach separates a multivariate distribution function into its marginal distribution functions and a copula; the latter is a distribution function with standard uniform margins comprising dependence information, see [7–11]. Analytically, a d-variate copula is a function with the uniform margins, groundedness, and d-increasingness properties. Furthermore, there exists an analog separation for multivariate survival functions into marginal survival functions and a survival copula, which is itself a copula and is uniquely linked to a copula of the distribution; see [11, Chp. 1]. The copula approach is popular, as it separates margins and dependency for methodological aspects, estimation, and simulation. Hence, one can recombine margins and copulas and employ principles and theory derived from studying univariate distributions. Moreover, the approach is closely linked to rank-based dependence measures such as Spearman's Rho, Kendall's Tau, or Blomqvist's Beta. However, the approach has limits, e.g., when using discontinuous margins or copulas that suggest specific margins, see [12, 13].

Parametric modeling often requires balancing flexibility and structure by choosing a model with appropriate degrees of freedom. For this, classes of distributions are often segmented into nested structures of subclasses characterized by increasingly specialized assumptions. A typical structure encloses extendible subclasses by exchangeable subclasses and these exchangeable subclasses by hierarchical factor classes. Exchangeability requires distributional stability under arbitrary permutations of components, and extendibility requires a stochastic representation as a finite-dimensional margin of an exchangeable sequence. A useful caveat is that exchangeable sequences of random variables are conditionally iid and define an almost surely unique random distribution function via a so-called de Finetti representation; see [14]. For nonnegative sequences, this random distribution function is associated with a subordinator, a nondecreasing stochastic process starting in zero and tending to infinity while having a.s. càdlàg path', which has the interpretation of a prior cumulative hazard rate of an iid sequence. Hierarchical factor classes require the exchangeability of margins with identical characteristics of the factors. This structure is helpful to multivariate normal and exponential distributions, see [15], as it allows leveraging many results about the extendible and exchangeable subclasses for more general hierarchical subclasses.

The first research focus of this dissertation concerns Marshall–Olkin distributions, as introduced in [16]. These distributions of nonnegative random vectors are associated with so-called exogenous shock models with exponential arrival times. Exogenous shock models represent each component of a random vector as the minimum of independent exponentially distributed shock-arrival times associated with subsets of elements, destroying all components in that subset not already destroyed by that time. A series of related works on Marshall–Olkin distributions [15, 17–19], build on each other, cumulate in a framework to build flexible high-dimensional hierarchical Marshall–Olkin distributions based on low-parametric extendible submodels; see also [11, Chp. 3]. However, existing simulation strategies for extendible Marshall–Olkin distributions, see [16, 17, 19–21], suffer from the curse of dimensionality, have numerical issues, or require

substantial restrictions such that a general, numerically stable and efficient end-to-end simulation algorithm for extendible Marshall–Olkin distributions was missing in the literature. Such an algorithm is provided in the contributed core article [2].

The second research focus of this dissertation concerns generalized Marshall–Olkin distributions introduced in [22, 23]. They are associated with a generalized exogenous shock model allowing arbitrary shock-arrival-time distributions, sometimes restricted to those with continuous survival functions. As a proper superclass of Marshall–Olkin distributions, a question standing to reason is which results about Marshall–Olkin distributions can be generalized to this superclass. This question was explored in multiple articles [12, 22–26]. However, the following two questions remained open: Can an analytical characterization for the entire class's survival functions, similar to that of the exchangeable subclass in [25], be established, and are the de Finetti representations of the extendible subclass always associated with additive subordinators? The contributed articles [3, 4] show that the answer to both questions is yes and provide the analytical characterization of the entire class's survival functions and the de Finetti representation of the extendible subclass.

The third research focus of this dissertation concerns upper semilinear copulas introduced in [27, 28]. These are exchangeable copulas whose margins are linear on segments with a constant lowest component connecting the diagonal and a marginal-boundary hyperplane. In the bivariate case, they constitute the copulas of exchangeable generalized Marshall–Olkin distributions. However, in higher dimensions, the following two questions remained open: What is their relationship to generalized Marshall–Olkin distributions, and how can they be represented stochastically? The contributed core article [1] derives the intersection of both classes and provides stochastic models for other subclasses of semilinear copulas.

The thesis is structured as follows: Chapter 2 discusses multivariate lack of memory properties, Sklar's copula separation, multivariate exponential distributions, and their relevance to portfolio credit risk management. Chapter 3 surveys existing research on Marshall–Olkin distributions and summarizes the findings from the contributed core article [2]. Chapter 4 explores existing research on generalized Marshall–Olkin distributions and summarizes the findings from the contributed articles [3, 4]. Chapter 5 examines existing research on upper semilinear copulas and summarizes the findings from the contributed core article [1]. Finally, Chapter 6 provides an outlook and suggests further research opportunities in the presented fields. The contributed articles are attached in Appendices A to C.

2 Multivariate lack of memory properties

This chapter provides a general overview of univariate and multivariate representations of random variables and vectors, focusing on modeling nonnegative, unbounded, and continuous random vectors, subsequently called *multivariate lifetimes*. The term multivariate lifetimes may represent the lifetime of mortal beings, defaultable companies, components in a technical system, or other entities to which a death analogy applies. The primary application of multivariate lifetimes considered in this thesis is modeling defaultable companies, although many of the concepts and techniques discussed also apply to other entities. This section discusses univariate lack of memory properties, Sklar's copula separation, multivariate exponential distributions derived from multivariate lack of memory properties, and the latter's relevance to portfolio credit risk management. This chapter presumes a knowledgeable reader to retain readability and will introduce concepts and terms formally only if it serves a purpose for this chapter. For concepts and terms not formally introduced here, the reader is referred to the provided references, respectively, and for Marshall–Olkin distributions to Chapter 3.

2.1 The univariate lack of memory property

Exponential distributions are central to univariate distributions for nonnegative, unbounded, and continuous lifetimes as an objective base model. For this, recall that the class of exponential distributions has one scale parameter to reproduce all nonnegative expectations and satisfies the characterizing *lack of memory (LOM)* property; see [29, Theorem 1]:

$$\mathbb{P}(\tau > s + t \mid \tau > t) = \mathbb{P}(\tau > s), \quad s, t \geq 0. \tag{2.1}$$

Hence, without prior knowledge about changing death probabilities upon survival over time, it is a canonical choice. Furthermore, nonnegative, unbounded, and continuous random variables τ with survival function $\bar{F}$ have the following representation in terms of their *(cumulative) hazard function* H:

$$\tau = \inf\{t \geq 0 \,:\, H(t) \geq E\} \text{ a.s.}, \tag{2.2a}$$

$$\bar{F}(t) = \exp\{-H(t)\} = \exp\left\{-\int_0^t \lambda(s)\,\mathrm{d}s\right\}, \quad t \geq 0, \tag{2.2b}$$

where E is a unit exponential random variable.[1] Thus, every nonnegative, unbounded, and continuous lifetime model has the interpretation of a time-changed unit exponential model with accumulated time H in the latter model.

This representation can be rewritten such that nonnegative, unbounded, and continuous lifetimes are the *first jump times* of Poisson processes. For this, write

$$\tau = \inf\{t \geq 0 \,:\, Z_t > 0\}\,, \tag{2.3a}$$

$$Z_t = 1 \wedge N_{H(t)} \text{ a.s.}, \quad t \geq 0, \tag{2.3b}$$

where N is a (homogeneous) Poisson process, with the Lebesgue intensity measure having E from Eq. (2.2a) as its first waiting time, and Z is the death-indicator process of τ; see Fig. 2.1.[2] This representation

[1] The representation in Eq. (2.2a) is due to the general distributional transform. Furthermore, the integral representation in Eq. (2.2b) requires the existence of a density f such that the *(infinitesimal) hazard rate* is $\lambda = f/\bar{F}$.

[2] The representation in Eq. (2.3) is due to the canonical representation of Poisson point processes via exponentially distributed waiting times and transformed points; see [30, Ex. 3.3.7, Prop. 3.7]. It potentially requires extending the probability space to support additional independent unit exponential random variables.

is beneficial for applications that require the death indicator's simulation alongside a discretized grid: if H and its generalized inverse $H^{\leftarrow}$ can be evaluated, the memorylessness of the exponentially distributed first jump time of N allows simulating τ given its survival until $t \geq 0$ by $H^{\leftarrow}(H(t) + H(\tilde{\tau}))$ with an independent copy $\tilde{\tau}$ of τ. Using Cauchy's functional equation, this implies that the exponential distributions' LOM property in Eq. (2.1) is equivalent to the death-indicator process being Markov.

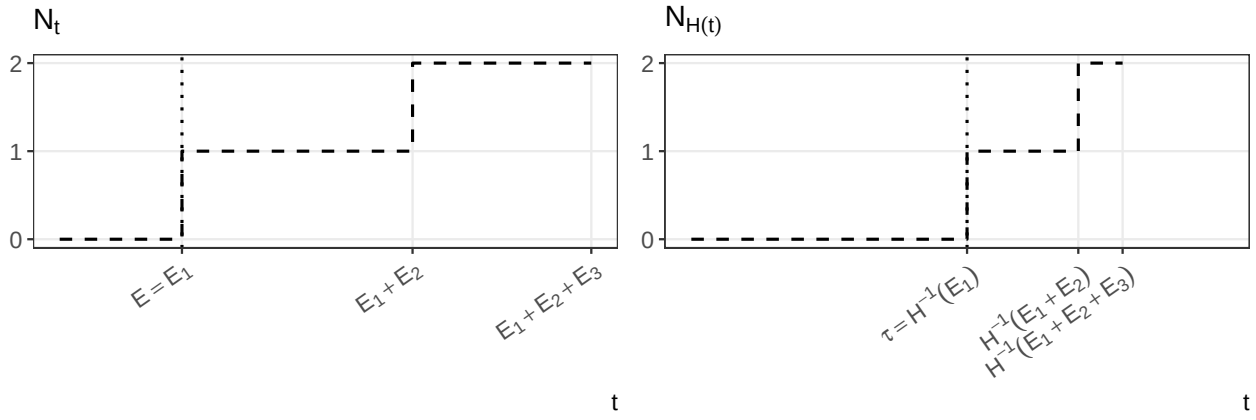


Figure 2.1 A visualization of the representations in Eq. (2.2) and Eq. (2.3) with τ having a Gompertz distribution with scale parameter 1 and shape parameter $1/10$ and a (homogeneous) Poisson process $N_t = \sum_{j \in \mathbb{N}} 1_{\{E_1 + \cdots + E_j \leq t\}}$.

An alternative interpretation of the *hazard time-change*, proposed in [24, Prop. 1], accumulates time with an associative binary operator linked to the hazard function, such that the distribution fulfills a generalized LOM property. For this, consider a strictly increasing, continuous hazard function H and its inverse H^{-1} and define the binary operation $* : \mathbb{R}_+ \times \mathbb{R}_+,\ (x, y) \mapsto H^{-1}(H(x) + H(y))$ with neutral element $H^{-1}(0)$. Then, distributions with hazard functions $\lambda \cdot H$, $\lambda \geq 0$, are uniquely characterized among continuous distributions by the following modified LOM property:

$$\mathbb{P}(\tau > s * t \mid \tau > t) = \mathbb{P}(\tau > s), \quad s, t \geq 0. \tag{2.4}$$

To make this concept more tangible, consider the following example: Suppose you are describing the *waiting time* to be called into a doctor's office. In a queue system that calls patients in the order of their arrival, it is intuitive that the likelihood for waiting at least another five minutes is different whether you just arrived or waited already for an hour. In this concept, the effect of *time-passed* on the waiting-time probabilities can be expressed as follows:

$$\mathbb{P}(\tau > t + s \mid \tau > t) = \mathbb{P}\Big(\tau > H^{-1}\big(H(t+s) - H(t)\big)\Big), \quad s, t \geq 0.$$

Finally, the classes with proportional hazard functions are closed under taking minima of independent random variables: Consider the hazard rate function H and independent $\tau_i \sim \lambda_i \cdot H$, $\lambda_i > 0$, $i \in [d]$.[3] Then,

$$\min_{i \in [d]} \tau_i \sim [\lambda_1 + \cdots + \lambda_d] \cdot H. \tag{2.5}$$

In conclusion, the hazard function representation of Eq. (2.2) and the first jump time representation of Eq. (2.3) allow for building intuitive and vividly interpretable *subjective* univariate lifetimes models, incorporating available data, from a unit exponential *objective base model*, considering only application constraints. Furthermore, the hazard representation is natural for many applications. For example, actuarial life tables are based on estimated yearly mortality rates among a surviving population; see [31]. In practice, this representation motivates the following approach: estimate discretized hazard rates from data, identify parametric classes of distributions with suitable hazard rate shapes, e.g., exponentially increasing rates, estimate the distribution parameters, and perform a model selection. Consequently, it is natural to lift this concept into higher dimensions and seek multivariate objective base models to build subjective models from them in intuitive and vividly interpretable ways.

[3] Consider the notation $[n] \coloneqq \{1, \ldots, n\}$ and $[n]_0 \coloneqq [n] \cup \{0\}$ for $n \in \mathbb{N}$.

2.2 The copula separation

The most comprehensive concept to represent multivariate random vectors is *Sklar's copula separation*, which originated from [6]; see also [7–11] for the following details. A *copula* is the distribution function of a random vector with continuous uniform margins. *Sklar's theorem* allows separating a distribution function F (survival function $\bar{F}$) into its marginal distribution functions F_i (*survival functions* $\bar{F}_i$) and a copula C (survival copula $\hat{C}$) as follows:

$$F(\boldsymbol{x}) = C\Big(F_1(x_1), \dots, F_d(x_d)\Big), \quad \boldsymbol{x} \in \mathbb{R}^d, \text{ and} \tag{2.6a}$$

$$\bar{F}(\boldsymbol{x}) = \hat{C}\Big(\bar{F}_1(x_1), \dots, \bar{F}_d(x_d)\Big), \quad \boldsymbol{x} \in \mathbb{R}^d, \tag{2.6b}$$

where C and $\hat{C}$ are both copulas. They are determined uniquely for a distribution if and only if the marginal distribution functions are continuous. In particular, since a copula's margins are continuous, each copula has a uniquely linked survival copula.

Sklar's copula separation allows extending common distribution classes to include their *meta-distributions* associated with the copulas from this class, see [32, Sec. 7.1.3]. For example, a *meta-Gaussian distribution* has the copula of a multivariate normal distribution in combination with arbitrary marginal distributions. Meta-distribution classes allow the coupling of known and well-tested dependence structures with arbitrary margins. Furthermore, limiting the model scope from arbitrary distributions to *meta-distribution classes* may improve comparability between distributions, as it is often considerably easier to compare distributions in such a meta-distribution class than across different meta-classes; see Section 2.3 and Fig. 2.2.

A d-variate random vector $\boldsymbol{\tau} = (\tau_1, \dots, \tau_d)$ can be standardized using the general distributional transform of [33, Prop. 2.1] to obtain

$$U_i \coloneqq F_i(\tau_i-) + V_i \cdot \big[F_i(\tau_i) - F_i(\tau_i-)\big], \quad i \in [d], \tag{2.7a}$$

$$E_i \coloneqq -\log(1 - U_i), \quad i \in [d], \tag{2.7b}$$

where V_i, $i \in [d]$, are additional iid uniform random variables used to interpolate jumps of the marginal distribution functions uniformly at random.[4] The marginal distributions of the vectors $\boldsymbol{U} = (U_1, \dots, U_d)$ and $\mathbf{1} - \boldsymbol{U}$ are unit uniform, and those of the vector $\boldsymbol{E} = (E_1, \dots, E_d)$ are unit exponential. Furthermore, the (multivariate) distribution function of $\boldsymbol{U}$ is a copula of $\boldsymbol{\tau}$, and that of $\mathbf{1} - \boldsymbol{U}$ is the corresponding survival copula of $\boldsymbol{\tau}$.[5] Moreover, this transformation is almost surely reversible:

$$\tau_i = F_i^{\leftarrow}(U_i) \text{ a.s.}, \quad i \in [d], \tag{2.8a}$$

$$\tau_i = \bar{F}_i^{\leftarrow}(1 - U_i) \text{ a.s.}, \quad i \in [d], \tag{2.8b}$$

$$\tau_i = H_i^{\leftarrow}(E_i) \text{ a.s.}, \quad i \in [d] \tag{2.8c}$$

Conversely, it is also possible to define a random vector $\boldsymbol{\tau}$ with distribution function F and copula C directly via Eq. (2.8) given a random vector $\boldsymbol{U}$ with unit uniform margins and distribution function C.

The copula separation allows separating marginal and dependence aspects for modeling, inference, and simulation. In particular, it permits using existing knowledge and techniques for the margins and dealing with the dependence separately. Overall, as all multivariate models have a copula separation as in Eqs. (2.6) to (2.8), the copula separation helps to understand multivariate distributions in general.

2.3 The copula model building approach

The copula separation motivates a popular methodological approach to build multivariate models by coupling marginal distributions with a copula selected from a diverse pool of candidates from different

[4] For a function f with left limits, denote the left limit of f at x by $f(x-)$.

[5] The standardization in Eq. (2.7) allows almost surely representing $\boldsymbol{\tau}$ via Eq. (2.8). However, if margins are discontinuous there are infinitely many alternative copulas C and associated survival copulas allowing stochastic representation of $\boldsymbol{\tau}$'s distribution via Eq. (2.8) from samples $\boldsymbol{U} \sim C$.

families. To differentiate this concept from the copula separation, applicable to all multivariate models independent of their primary specification, this is called the *copula approach*. The characteristic of this approach is its agnosticism towards specific meta–distributions, as the model's scope is essentially limited only by awareness of copulas and their applicability to the problem at hand, e.g., using maximum likelihood estimators requires a known density. A review of popular textbooks [7–11] yields the common families of *elliptical copulas*, including the popular *Gaussian copulas*, *Archimedean copulas*, *extreme-value (EV) copulas*, including *Marshall–Olkin (MO) copulas*, *Vine copulas*, and copulas obtained via recombination techniques of those.[6]

The most popular Gaussian copulas allow vivid interpretations of their dependence structures by stochastically representing them as linear transformations of iid standard normal random variables. In addition, unlike most other copula models, they are fully characterized by pair-correlation matrices. Similarly, model building often relies on stochastic representations to interpret the nature of the built distribution. However, these stochastic models and their parametrization are often considerably more complex than those of Gaussian copulas. Examples 2.1 to 2.5 present selected popular copula families and their stochastic representations.

Example 2.1 ([11, Exp. 1.17])*. One-factor extendible Gaussian copulas* arise as the copulas of the exchangeable *multivariate normal distributions* with *Pearson correlation coefficients* $\rho \in [0, 1]$. They have the unique stochastic representation

$$U_i \coloneqq F\Big(\sqrt{1-\rho}Z_i + \sqrt{\rho}Z_G\Big), \quad i \in [d],$$

where $Z_G, Z_1, \ldots, Z_d$ are iid standard normally distributed, and F is the standard normal distribution function.[7] There are simple one-to-one mappings from associated normal distributions' Pearson correlations to the popular dependence coefficients *Kendall's* τ and *Spearman's* ρ; see [32, Thm. 7.42].

Example 2.2 ([32, Chp. 6 and 7])*. Elliptical copulas* arise as the copulas of *multivariate elliptical distributions* with *characteristic generator* ψ and *dispersion matrix* Σ. They have the canonical stochastic representation

$$U_i \coloneqq F_i\Big(a_i^\top R \cdot \boldsymbol{Z}/\|\boldsymbol{Z}\|_2\Big), \quad i \in [d],$$

where $A = (a_1, \ldots, a_d)^\top \in \mathbb{R}^{d\times k}$ is a matrix with $AA^\top = \Sigma$, $\boldsymbol{Z} = (Z_1, \ldots, Z_k)$ is a vector of iid standard normally distributed random variables, R is a random radial variable whose distribution is associated with ψ, independent of $\boldsymbol{Z}/\|\boldsymbol{Z}\|_2$, and $F_i \equiv F_{\psi,\|a_i\|_2^2}$, $i \in [d]$, are the associated marginal distribution functions. The stochastic representation is unique up to the factorization of Σ and a deterministic scaling of $\boldsymbol{a}_i$ and R.

Example 2.3 ([34])*. Archimedean copulas* arise as the survival copulas of so-called ℓ_1*-norm symmetric distributions* without an atom in $\mathbf{0}$. They have the stochastic representation

$$U_i \coloneqq \bar{F}(R \cdot E_i/\|\boldsymbol{E}\|_1), \quad i \in [d],$$

where $\boldsymbol{E} = (E_1, \ldots, E_d)$ is a vector of iid unit exponential random variables, and R is a radial random variable independent of $\boldsymbol{E}/\|\boldsymbol{E}\|_1$. The stochastic representation is unique up to a deterministic scaling of R. The associated *Archimedean generator* $\varphi_{R,d} = \bar{F}$ is the so-called *Williamson* d*-transform* of the radial variable R. There exists an integral formula to calculate *Kendall's rank correlation coefficient* τ from given generators; see [8, Cor. 5.1.4].

Example 2.4 ([30, Prop. 5.11])*. Extreme-value copulas* arise as the copulas of *extreme-value distributions with standard Fréchet margins*. They have the unique stochastic representation

$$U_i \coloneqq F\left(\eta \cdot \bigvee_{j\in\mathbb{N}} \frac{(\boldsymbol{S}_j)_i}{E_1 + \cdots + E_j}\right), \quad i \in [d],$$

[6] The copula approach is not to be confused with the approach of choosing a distribution from a distinct meta–distribution class, which would require a deliberate restriction of the model's scope to this class.

[7] In this context, *unique* means that the representation is unambiguous in this particular model. However, it does not mean no other stochastic representations from different models exist.

for $\eta > 0$, independent sequences $E_1, E_2, \ldots$ iid unit exponential and $\boldsymbol{S}_1, \boldsymbol{S}_2, \ldots$ iid on the unit simplex with a suitable distribution fulfilling $\mathbb{E}[\boldsymbol{S}_1] = \mathbf{1}/\eta$, and the standard Fréchet distribution function F. EV copulas are also known as the possible limit copulas of individually linearly transformed component-wise maxima of iid sequences of random vectors, and if an EV copula C^{EV} is the limit copula of such a sequence originating from the copula C, we say C is in the *maximum domain of attraction* of C^{EV} and write $C \in \mathrm{MDA}(C^{\mathrm{EV}})$; see [11, Sec. 1.2.5]. In particular, C is an extreme-value copula if and only if it fulfills the property

$$C(\boldsymbol{u}^t) = C(\boldsymbol{u})^t, \quad \forall t \geq 0,\ \boldsymbol{u} \in [\mathbf{0}, \mathbf{1}].$$

Example 2.5 ([14] and [4, Apndx. A]). Extendible copulas (from arbitrary families) arise as the copulas of *conditionally iid distributions*. They have the stochastic *de Finetti* representation

$$U_i \coloneqq \bar{F}\Big(\Psi^{\leftarrow}(V_i)\Big), \quad i \in [d],$$

where Ψ is a random distribution function with $F(x) \coloneqq \mathbb{E}[\Psi(x)]$, $x \in \mathbb{R}$, and $V_1, \ldots, V_d$ is a sequence of iid uniformly distributed random variables. If Ψ has almost surely nonnegative real support, this representation can be rewritten to standardized first-jump times of a random subordinator over independent iid unit exponential barriers:

$$U_i \coloneqq \bar{F}\Big(\inf\{t \geq 0\ :\ \Lambda_t \geq E_i\}\Big), \quad i \in [d],$$

where $\Lambda \coloneqq -\log(1 - \Psi)$ is a subordinator and $\boldsymbol{E} \coloneqq -\log(1 - \boldsymbol{V})$ is a vector of iid unit exponential random variables. Selected examples are: Extendible Archimedean copulas are associated with random conditionally linear maps; see [35]. Extendible Marshall–Olkin survival copulas are associated with Lévy subordinators; see [17]. Extendible generalized Marshall–Olkin survival copulas are associated with additive subordinators, explored in the contributed article [4]; see also Chapter 4 and [25]. Extendible extreme-value copulas and survival copulas of *extendible exponential minima* distributions are associated with so-called *strong and weak infinitely divisible w.r.t. time subordinators*; see [36]. Finally, extendible *min–infinitely divisible* survival copulas are associated with so-called *infinitely divisible subordinators*; see [37].

The copula approach is appropriate for many applications but also promotes an inherently atomistic view of the model, which leads to considerations and limitations discussed in the following. The first three are relevant for all margin-transforming approaches.

- *Discontinuous margins (see [12, 13])*: Coupling copulas with discontinuous marginal distributions implies that other copulas are also associated with the obtained full distribution, resulting in ambiguity about a distribution's dependence without considering its margins. Furthermore, probabilistic definitions of Spearman's ρ and Kendall's τ become margin-dependent for discontinuous margins. [12, Exp. 2.8] couples a specific bivariate copula with two appropriately selected sets of marginal distributions such that those full distributions may also be associated with the comonotonicity and independence copula, respectively.

- *Margin-dependent properties (see [12])*: Some copulas suggest specific margins, and coupling them with other margins results in losing specific properties. A slightly modified and condensed version of an example given in [38, Exp. 2.3] is this:[8] Consider a bivariate MO-distributed random vector $\boldsymbol{\tau}$, scaling factors $\alpha_1 > \alpha_2 > 0$, and define $\boldsymbol{\eta} = \boldsymbol{\alpha} \cdot \boldsymbol{\tau}$. Then, $\boldsymbol{\eta}$ still has exponential margins and a survival copula of MO-kind, but is not MO-distributed itself. In particular, the singular component will not be on the diagonal section anymore, implying some conditional previsibility of its death-indicator process, as $\mathbb{P}(\eta_1 = \eta_2 \cdot (\alpha_1/\alpha_2)) > 0$.

- *Simulation incompatibilities*: The natural approach to simulate a distribution specified with the copula approach is to first simulate $\boldsymbol{U} \sim C$ and then to apply the distribution transform in Eq. (2.8a). However, this requires a simulation algorithm for C and the distribution transform simulation method being

[8] This example anticipates basic knowledge about the bivariate MO distribution and their *singular components*, which will both be discussed in greater detail in Chapter 3.

appropriate for all marginal distributions. For example, this method produces bounded samples for Poisson distributions with *linear congruential random number generators (RNGs)*, see [39, Chp. 2], due to the RNG's inability to produce arbitrarily small numbers. However, there are also less pathological examples for which other methods are preferred over the distribution transform simulation because calculating (or approximating) the generalized inverse of the distribution function is difficult, for example, for Beta distributions; see [39, Exp. 2.2.6].

- *Limits of interpretability and comparability*: Many applications require an intuitive interpretation of the composed model, proven robustness of the fitting procedure, and comprehensible comparisons to alternative models. This is a requirement, in particular, of regulated financial companies' production models, which must be justified to internal validation units, management, and regulators. However, many multivariate lifetime applications lack data for sole statistical model selection and parameter estimation. They require expert judgements and manual modeling decisions about the copulas and margins considered, carefully balancing model structure and flexibility, and suitability assessments. This requires a reasonable understanding of the composed distributions, with their particular specialties and limitations. For example, Gaussian copulas lack tail dependence, and Archimedean copulas are always exchangeable; see [11, Chp. 2 and 4]. Stochastic representations could help make knowledgeable modeling decisions. However, as outlined in the following and highlighted in Fig. 2.2, the stochastic representations of different multivariate distributions are rarely compatible.
  - In *simple* meta-distribution classes such as the meta–extendible Gaussian class, distributions can be interpreted and compared fully and vividly via the stochastic representation in Exp. 2.1.
  - In *complex* meta–distribution classes, such as the meta-Archimedean class, some distributions are difficult to interpret or compare via the stochastic representation in Exp. 2.3. In the case of meta-Archimedean distributions, there exists some intuitions about how properties of the radial distribution relate to specific properties of the copula, e.g., (possible) singular components, see [34, Sec. 4.1], or extremal behavior, see [40]. However, these relationships are highly complex and require specialized knowledge about the radial distribution. Note that alternative stochastic representations not involving the radial variable explicitly require thinking differently about distribution properties; see [41] for an approach to investigate bivariate Archimedean copulas via their *Markov kernels*, which are regular conditional distributions. Moreover, multiple Archimedean copulas may have equal dependence coefficients, and choosing between those copulas is rarely straightforward.
  - *Different* meta–distribution classes, such as meta–extreme-value and meta-elliptical distributions, often have entirely different stochastic representations, making it infeasible to vividly compare two copulas from these classes via their stochastic representations in Exps. 2.2 and 2.4. In the particular example, both meta-distribution classes have infinitely many candidates sharing a particular dependence coefficient value, e.g., a specific value for Kendall's τ. It would be very challenging to manually choose the most suitable candidate if this decision cannot be solely data-based.

2.4 Multivariate lack of memory properties

The copula approach naturally suggests an atomistic view of the model by separating it into its margins and copula. However, a holistic approach with strong model assumptions is needed when margins and copula cannot be selected and estimated individually with sufficient confidence due to little available data. The lack of available data is a problem of many multivariate lifetime applications due to the nonrecurrent and sometimes unprecedented nature of (individual) death events.

An alternative to the copula approach for multivariate lifetimes is using an *objective* multivariate base model that serves as a root for more subjective, interpretable models obtained by intuitive, vivid transformations, similar to unit exponential distributions in the univariate case. Labeling a model *objective* means that

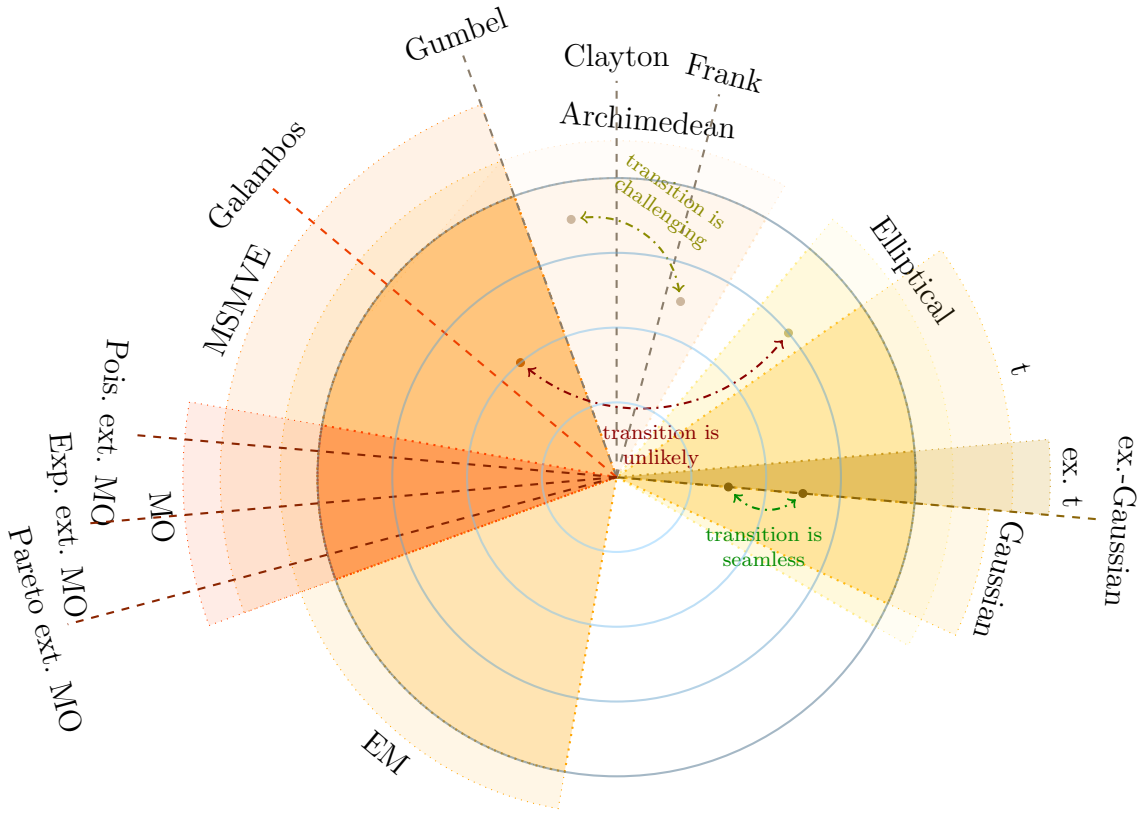


Figure 2.2 A hypothetical, exemplary projection of copulas into the unit disk, mapping the radius to a dependence coefficient and the angular coordinate suitably to preserve copula-class cohesion. The selection is not exhaustive, and the areas do not represent the classes' sizes. The figure indicates the seamlessness to transition in a stochastic representation by deterministic, meaningful parameter transformations to change the copula from one to another, thereby obtaining a vivid comparison of both. The dashed lines represent low-parametric families with stochastic representations allowing seamless transitions. The opaque areas represent larger families with stochastic representations allowing transitions, sometimes only with increased complexity. Transitions between opaque areas are often difficult or impossible in the currently known representation without explicit model switching.

it reflects (stylized) facts, e.g., nonnegative, unbounded, and continuous marginal lifetimes, and that it is a sensible base model easily and intuitively configurable with little to no margin or dependency information.[9] This approach addresses the copula approaches interpretability and comparability issues, mentioned in the last section, by giving up some of its flexibility by restricting the model's scope to a meaningful class of distributions.

It seems natural to search objective base models for multivariate lifetime modeling in multivariate generalizations of univariate exponential distributions' characterizing properties. Recall that exponential distributions are characterized by the LOM property in Eq. (2.1) or, equivalently, by their death-indicator processes being Markov. The following is a brief excerpt from the survey in the contributed article [5] about existing multivariate generalizations of LOM properties. Let $\boldsymbol{\tau}$ be a d-variate nonnegative, unbounded, and continuous random vector, $\boldsymbol{s}, \boldsymbol{t} \geq 0$ be d-variate nonnegative vectors, and $s, t, c \geq 0$ nonnegative values; and consider the following properties:

- $\boldsymbol{\tau}$ has the *multivariate independent exponential lack of memory (MIELOM)* property if

$$\mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I + \boldsymbol{t}_I \mid \boldsymbol{\tau}_I > \boldsymbol{t}_I) = \mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I) \quad \forall \emptyset \neq I \subseteq [d],\ \boldsymbol{s}, \boldsymbol{t} \geq 0. \tag{2.9}$$

- $\boldsymbol{\tau}$ has the *Marshall–Olkin lack of memory (MOLOM)* property if

$$\mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I + t \mid \boldsymbol{\tau}_I > t) = \mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I) \quad \forall \emptyset \neq I \subseteq [d],\ \boldsymbol{s} \geq 0,\ t \geq 0. \tag{2.10}$$

- $\boldsymbol{\tau}$ has the *min-stable multivariate exponential (MSMVE)* LOM property if

$$\mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{c}_I(s + t) \mid \boldsymbol{\tau}_I > \boldsymbol{c}_I t) = \mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{c}_I s) \quad \forall \emptyset \neq I \subseteq [d],\ \boldsymbol{c} \geq 0,\ s, t \geq 0. \tag{2.11}$$

- $\boldsymbol{\tau}$ has the *exponential minima (EM)* LOM property if

$$\mathbb{P}(\boldsymbol{\tau}_I > s + t \mid \boldsymbol{\tau}_I > t) = \mathbb{P}(\boldsymbol{\tau}_I > s) \quad \forall \emptyset \neq I \subseteq [d],\ s, t \geq 0. \tag{2.12}$$

The MIELOM and MOLOM properties in Eqs. (2.9) and (2.10) were first discussed in [16]. This article proved that MIELOM implies independent and exponential margins, and it introduced the class of Marshall–Olkin distributions, which will be discussed in Chapter 3, that fulfill the MOLOM property. As in the univariate case, the MOLOM property can be rewritten as a Markov property, as it is equivalent to all marginal death-indicator processes being Markov, see [21]. The MSMVE LOM property characterizes MSMVE distributions, which are equivalent to compositions of EV–survival copulas and exponential margins, see [42, Lem. 2.44].

These multivariate LOM properties in Eqs. (2.9) to (2.12) and the related distributions were investigated in [38] with the following findings:

- MIELOM implies MOLOM, MOLOM implies MSMVE, and MSMVE implies EM; in particular, MOLOM is the strongest generalization that does not imply independence. Conversely, the article also proves by providing counterexamples that the converse does not always hold. In a similar spirit, we provide two examples in the contributed article [5] showing that the weak version of EM, fulfilling Eq. (2.12) for $I = [d]$, does not always imply EM.
- Random vectors $\boldsymbol{\tau}$ having one of these multivariate LOM properties have a positive dependence in the following sense; see [38, Sec. 5]:
  - If $\boldsymbol{\tau}$ has the MOLOM property, it is also *associated*, i.e., for all nondecreasing real-valued functions f, g for which the following expression is well-defined, it holds that

$$\operatorname{Cov}[f(\boldsymbol{\tau}), g(\boldsymbol{\tau})] \geq 0.$$

[9] The objective base model's choice is subjective in the literal sense and should be understood axiomatically.

– If $\boldsymbol{\tau}$ has the MSMVE property, it is also *positively left and right quadrant dependent*, i.e.,

$$F(\boldsymbol{t}) \geq \prod_{i=1}^{d} F_i(t_i), \quad \forall \boldsymbol{t} \geq 0, \quad \text{and} \quad \bar{F}(\boldsymbol{t}) \geq \prod_{i=1}^{d} \bar{F}_i(t_i), \quad \forall \boldsymbol{t} \geq 0.$$

– If $\boldsymbol{\tau}$ has the EM property, it also holds that

$$F(t\mathbf{1}) \geq \prod_{i=1}^{d} F_i(t), \quad \forall t \geq 0, \quad \text{and} \quad \bar{F}(t\mathbf{1}) \geq \prod_{i=1}^{d} \bar{F}_i(t), \quad \forall t \geq 0.$$

- A *coherent life function* α maps a random vector of lifetimes $\boldsymbol{\tau}$ to an overall system lifetime $\alpha(\boldsymbol{\tau})$ that takes the form

$$\alpha(\boldsymbol{\tau}) = \max_{j \in [n]} \min_{i \in I_j} \tau_i,$$

for sets $I_1, \ldots, I_n \subseteq [d]$. Consider $\boldsymbol{\tau}$ having an EM distribution, then there exists an MO-distributed random vector $\tilde{\boldsymbol{\tau}}$ that is *marginally equivalent in coherent life functions*, i.e., $\alpha(\boldsymbol{\tau}) \overset{d}{=} \alpha(\tilde{\boldsymbol{\tau}})$ for all coherent life functions α.

The suitability of MO distributions for real-world applications that do not involve concurrent deaths, such as credit modeling, may not be initially apparent. Whether this limitation excludes MO distributions from possible choices depends on the modeling decision. However, the finding about marginal equivalence in coherent life functions suggests that MO distributions might still be appropriate for many applications where EM distributions, which do not always have singular components, are suitable. In particular, this is the case when the quantity of interest can be simplified to the overall system lifetime. Additionally, another argument against excluding MO distributions arises when time is discretized.

From a modeling perspective, MO distributions are particularly interesting candidates for objective base models since they allow generalizations of the exponential distributions' hazard and Markov representations.

- Any MO distribution has the representation

$$\tau_i = \inf\{t \geq 0 \ : \ Z_t \ni i\}, \quad i \in [d],$$

for a Markovian death-indicator process Z on the power-set of $[d]$; see [21].

- Any exchangeable Marshall–Olkin (exMO) distribution has the stochastic representation

$$\tau_i \coloneqq \inf\{t \geq 0 \ : \ Z^*_t \geq \Pi(i)\}, \quad i \in [d],$$

for a Markovian death-counting process Z^* on $[d]_0$ and a uniformly random permutation Π, independent of Z^*; this is proven in the contributed core article [2].

- Any extendible Marshall–Olkin (extMO) distribution has the stochastic representation

$$\tau_i \coloneqq \inf\{t \geq 0 \ : \ \Lambda_i \geq E_i\}, \quad i \in [d],$$

for a Lévy subordinator Λ and iid unit exponential random variables $E_1, \ldots, E_d$, independent of Λ; see [17].

These stochastic representations are easily identified as natural multivariate generalizations of the *first-jump-time representation* in Eq. (2.3) or the *hazard representation* in Eq. (2.2) for the extendible case. Furthermore, the *hazard-time-change*, associated with replacing the standard addition with a binary operator linked to an invertible hazard function, may also be used in this multivariate generalization to obtain flexible meta–distribution classes that remain interpretable for multivariate lifetime applications. However, note that, albeit being interpretable, the class of MO distributions also has issues with over-parametrization and parameter-robustness in higher dimensions, since it has $2^d - 1$ parameters. Consequently, practical applications benefit from using structured low-parametric subclasses that are meaningful in the used context. Examples of such approaches can be found in [43–47] and the contributed article [5] as summarized in the following subsection.

2.5 Applications to portfolio credit risk

The following section summarizes key findings from the contributed further article [5], which critically investigates the common industry practice of iteratively simulating default indicators over discretized time grids. It highlights the relevance of LOM properties to portfolio credit risk management. First, it outlines that investigating the terminal dependence for *survival-of-all* events is similar to finding the maximum domain of attraction to which the inter-period distribution's survival copula belongs. In particular, it demonstrates that popular choices for survival copulas for the inter-period distribution may result in unexpected and unintended vanishing terminal dependence. Second, it suggests using hierarchical Marshall–Olkin (hMO) distributions to address a bouquet of methodological challenges of this industry practice.

Practitioners often model default by iteratively simulating *inter-period default-indicators* $Y_k^{(\Delta)}$ over a discretized time grid $\mathcal{T} := (0, \Delta, \dots, n\Delta)$ for some fixed *time increment* $\Delta > 0$ and *terminal time horizon* $T = n\Delta$ alongside other risk factors, i.e., a default-indicator process $Z^{(\Delta)}$ on $\mathcal{T}$ is defined by

$$Z^{(\Delta)} : \mathcal{T} \to \mathbb{P}([d]),\ t \mapsto \bigcup_{k=1}^{t/\Delta} Y_k^{(\Delta)},$$

where the inter-period default-indicators are usually implicitly defined by an inter-period law $\boldsymbol{\zeta}^{(\Delta)}$ with

$$Y_k^{(\Delta)} \overset{d}{=} \Big\{ i \in [d] \ :\ \zeta_{i,k}^{(\Delta)} \le \Delta \Big\}, \quad k \in [n].$$

The inter-period law, if defined, is either modeled explicitly or is defined implicitly by a desired terminal law $\boldsymbol{\tau}^{(\Delta)}$ fulfilling

$$\Big\{ Z_t^{(\Delta)} \ :\ t \in \mathcal{T} \Big\} \overset{d}{=} \Big\{ \Big\{ i \in [d] \ :\ \tau_i^{(\Delta)} \le t \Big\} \ :\ t \in \mathcal{T} \Big\}.$$

In both cases, the inter-period and terminal distributions and their relationship must be feasible and well-understood. In particular, sampling the inter-period indicators and assessing the terminal distribution's suitability must be feasible. Hence, while the inter-period default-indicators may depend on previous timesteps and external risk factors, practical constraints often exclude complex inter-period dependence. Moreover, to retain tractability, they suggest Markovian default-indicator chains, conditioned on external risk factors, or, stronger, embeddability in a continuous model with a Markovian default indicator process.

First, the contributed article [5] outlines this approach's fallacy that terminal dependence can be inadvertently significantly less pronounced than inter-period dependence. For demonstration, it assumes a random walk–type default indicator chain with iid inter-period indicators, independent of other risk-factors, based on an exchangeable distribution with exponential marginal survival function $\bar{F}^{\boldsymbol{\zeta}}$ and exchangeable survival copula $\hat{C}^{\boldsymbol{\zeta}}$, not dependent on the chosen time increment:

$$\bar{F}^{(\boldsymbol{\zeta})} = \hat{C}^{(\boldsymbol{\zeta})} \oplus (\bar{F}^{(\boldsymbol{\zeta})}, \dots, \bar{F}^{(\boldsymbol{\zeta})}).$$

Now, the article observes that if $\hat{C}^{(\boldsymbol{\zeta})} \in \mathrm{MDA}(\hat{C})$ for an extreme-value (survival) copula $\hat{C}$, then

$$\begin{aligned}
\mathbb{P}\Big(Z_t^{(T/n)} = \emptyset\Big) &= \mathbb{P}\Big(\boldsymbol{\tau}^{(T/n)} > t\Big) = \mathbb{P}(\boldsymbol{\zeta} > T/n)^{\lfloor n\cdot(t/T)\rfloor} \\
&= \hat{C}^{(\boldsymbol{\zeta})}\Big(\bar{F}^{(\boldsymbol{\zeta})}(t)^{1/(t/T)/n}, \dots, \bar{F}^{(\boldsymbol{\zeta})}(t)^{1/(t/T)/n}\Big)^{\lfloor n\cdot(t/T)\rfloor} \\
&\overset{n\to\infty}{\longrightarrow} \hat{C}\Big(\bar{F}^{(\boldsymbol{\zeta})}(t), \dots, \bar{F}^{(\boldsymbol{\zeta})}(t)\Big), \quad \forall 0 \le t \le T.
\end{aligned}$$

Hence, for many survival copulas in the independence copula's *maximum domain of attraction*, e.g., Gaussian copulas, Clayton copulas, and Frank copulas, inter-period dependence is not carried over to terminal dependence for *survival-of-all* events.

Second, the contributed article [5] suggests that low-parametric hMO distributions, as described in Rmk. 2.6, have valuable properties as inter-period distributions for iteratively simulating default-indicators over discretized time grids:

- The factor model structure facilitates building low-parametric, high-dimensional factor models that allow adding or removing entities without changing parameters associated with other entities.
- The discrete-time default-indicator chains may be embedded in a universal continuous-time default indicator process, allowing arbitrary time discretizations and horizons.
- The marginal default-indicator chains for arbitrary time discretizations are homogenous Markov chains, which allows simulating subportfolios on arbitrary time discretizations.
- Matching inter-period and terminal distributions simplify reasoning about terminal probabilities.

It is noteworthy that compared to the proposed models in [43, 44], which assume factors with individual and global shocks resulting in individual-, global-, and sector-killing events, this hierarchical model permits a broader range of shocks, as [45], while retaining the low-parametric factor model structure.

Remark 2.6 (Low-parametric hMO iterating default model, see [2, 5])*.* Suppose $d \geq 2$ and a sector-associated partition of defaultable entities $J_1 \dot\cup \cdots \dot\cup J_n = [d]$. Consider a global factor and sector factors associated with extMO distributions via Bernstein functions $\psi^{(0)}, \psi^{(1)}, \ldots, \psi^{(n)}$, which may be chosen from suitable low-parametric families of Bernstein functions.[10] The corresponding hMO distribution has the survival function

$$\bar{F}(\boldsymbol{t}) = \bar{F}^{\psi^{(0)}}(\boldsymbol{t}) \cdot \prod_{k=1}^{n} \bar{F}^{\psi^{(k)}}(\boldsymbol{t}_{J_k}), \quad \boldsymbol{t} \geq 0,$$

and can be sampled as the component-wise minimum of two independent random vectors: a d-variate extMO distributed random vector associated with $\psi^{(0)}$, and a vector of concatenated independent $|J_k|$-variate extMO distributed random vectors associated with $\psi^{(k)}$, respectively, appropriately re-ordered; see [48, Sec. 4.2] and the contributed core article [2] about how to simulate extMO distributions. As a consequence of the *Arnold model (AM)* representation, see [20] and [2, Sec. 3], the iterating death-indicator model associated with this inter-period distribution fulfills the property that a terminal distribution exists and matches the inter-period distribution for arbitrary time discretizations; see the addendum to the contributed article [5].

This model implies exponential margins and exchangeable sector components. However, when non-exponential margins are required, many of these model's features can be lifted into a meta–distribution class that is obtained by a simple time and time-step-size transformation as described in Section 2.1.

[10] Bernstein functions and their relationship to extMO distributions is discussed in Chapter 3. For now, consider them as parameters for extMO distributions.

3 Marshall–Olkin distributions

This chapter primarily discusses the contribution of the core article [2] about implementing simulation algorithms for extendible Marshall–Olkin distributions. The core article also contains detailed introductions to the necessary background theory. The interested reader is referred to the article for all topics only broached in this chapter.

3.1 Marshall–Olkin distributions

Marshall–Olkin (MO) distributions were introduced in [16] as the continuous distributions characterized by the generalized multivariate lack of memory (LOM) property in Eq. (2.10), and they are shown to have the survival function

$$\bar{F}(\boldsymbol{t}) = \exp\left\{ - \sum_{\emptyset \neq I \subseteq [d]} \lambda_I \max_{i \in I} t_i \right\}, \quad \boldsymbol{t} \geq 0,$$

for so-called *shock-arrival intensities* λ_I, $\emptyset \neq I \subseteq [d]$, fulfilling the *marginal finiteness condition*

$$\sum_{I \ni i} \lambda_I > 0, \quad \forall i \in [d].$$

These distributions do not have a multivariate *Lebesgue density* because of the maximum operation in the survival function. Consequently, they have so-called *singular components*, i.e., lower-dimensional subsets with positive probability, on (marginal) diagonal sections.

MO distributions have been researched extensively over the past decades, also w.r.t. aspects not primarily discussed in this thesis. The interested reader is referred to an extensive survey in [11, Chp. 3]; for the exchangeable subclass, see also [49]. In addition, the following list contains notable works about MO distributions that are not discussed elsewhere in the thesis: *Maximum Likelihood estimation* w.r.t. a dominating measure was established in [50, 51]. A moment-based estimation approach for the extendible subclass was researched in [52]. The survival copula of MO distributions was derived in [53], and its *Pickands representation* was derived in [54]. Finally, [55] calculated the distributions of MO-distributed random vectors' means for lower dimensions explicitly and for extendible random vectors in higher dimensions asymptotically.

MO distributions have several *general*, i.e., non-specialized, stochastic representations (see contributed core article [2, Sec. 2.2] for more details):

- The *exogenous shock model (ESM)*; see [16]: The original stochastic representation of MO distributions assumes independent exponential *shock-arrival times* E_I with rates equal to the distributions' shock-arrival intensities λ_I. These are associated with nonempty subsets of components $\emptyset \neq I \subseteq [d]$. A component i's death time τ_i is defined as the first shock-arrival time associated with a subset containing that component:
$$\tau_i \coloneqq \min\{E_I \ : \ I \ni i\}, \quad i \in [d].$$
- The *Arnold model (AM)*; see [20]: Assume iid set-valued *shocks* $Y_1, Y_2, \ldots \subseteq [d]$ with probabilities p_I, $\emptyset \neq I \subseteq [d]$, proportional to the shock-arrival intensities. They arrive after iid exponential waiting times $W_1, W_2, \ldots$, independent of the shocks, with an intensity equal to the sum of all shock-arrival intensities λ. A component i's death time τ_i is triggered by the first shock-set containing it:
$$\tau_i \coloneqq \min\{W_1 + \cdots + W_j \ : \ Y_j \ni i\}, \quad i \in [d].$$

- The Markov representation, subsequently called the *Markov death-set model (MDSM)*: see [21]: The *death-indicator processes* of MO distributions are Markov processes. Furthermore, given a sample path of the death-indicator process Z, a component i's death time is defined as the first time its death is indicated:

$$\tau_i \coloneqq \inf\{t \geq 0 \ : \ Z_t \ni i\}, \quad i \in [d].$$

- Another stochastic representation was proposed in [56, Thm. 3] for a subclass fulfilling the hierarchy condition $\lambda_I = 0 \Rightarrow \lambda_J = 0 \, \forall I \subseteq J$. Those MO distributions arise as the first surpassing-times of d-dimensional Lévy subordinators components' Λ^i over iid unit exponential barrier values $E_1, \ldots, E_d$, independent of the Lévy subordinator, i.e.,

$$\tau_i \coloneqq \inf\left\{t \geq 0 \ : \ \Lambda^i_t \geq E_i\right\}, \quad i \in [d].$$

  Furthermore, [56, Lem. 3] shows that such a d-dimensional Lévy subordinator can be constructed from independent compound Poisson processes with exponential jumps with common rate: There exists $\alpha_I \geq 0$, $\emptyset \neq I \subseteq [d]$, and $\eta > 0$, fulfilling a known recursion, such that we can define independent compound Poisson processes Υ^I with intensity α_I and exponential jumps with rate η:

$$\Lambda^i \coloneqq \sum_{I \ni i} \Upsilon^I, \quad i \in [d].$$

The core article [2] contributes the following novel results about general stochastic representations of MO distributions to the scientific literature:

- [2, Sec. 3] rewrites the death-indicator process Z of an MO-distributed random vector $\boldsymbol{\tau}$ represented by the AM as a random walk $\tilde{Z}$ on the semigroup of $[d]$'s power-set with the union as a conjunction subordinated by a Poisson process N with intensity λ, associated with the AM's waiting times, i.e.,

$$Z_t = \bigcup_{j=1}^{N_t} Y_j = \tilde{Z}_{N(t)}, \quad t \geq 0.$$

  The separation $Z = \tilde{Z} \circ N$ extends the result from [21, p. 62], which previously derived the first part of the equation above, by emphasizing this Markov process being a Poisson process–subordinated Markov chain.

- [2, Thm. 3.1] uses the AM's new death-indicator processes representation as a Poisson process–subordinated Markov chain to calculate the first explicit representation of its infinitesimal Markov generator $Q = (q_{I,J} : I, J \subseteq [d])$:

$$q_{I,J} = \begin{cases} -\sum_{I \subsetneq K \subseteq [d]} \sum_{L \subseteq I} \lambda_{L \cup (K \setminus I)} & , I = J, \\ \sum_{L \subseteq I} \lambda_{L \cup (J \setminus I)} & , I \subsetneq J, \\ 0 & , \text{else}. \end{cases}$$

  Specialized representations of Q for MO subclasses are also provided in [2, Rmk. A.1–A.5].

3.2 Exchangeable Marshall–Olkin distributions

Exchangeable Marshall–Olkin (exMO) distributions are characterized by shock-arrival intensities being equal for shock-sets having the same cardinalities, see [49, Lem. 3.1.1], i.e.,

$$|I| = |J| \ \Leftrightarrow \ \lambda_I = \lambda_J \quad \forall \emptyset \neq I, J \subseteq [d].$$

Consequently, exMO distributions allow the following three reparametrizations, which all require only d instead of $2^d - 1$ parameters:

- As shock-arrival intensities are equal for matching cardinalities, the *exchangeable shock-arrival intensities* λ_i, $i \in [d]$ are indexed by the cardinality of the associated sets.
- Later, the *exchangeable shock-size-arrival intensities* $\eta_i = \binom{d}{i}\lambda_i$, $i \in [d]$, equal to the accumulated shock-arrival intensities of sets with specific cardinalities, become a helpful reparametrization.
- [18] showed that each sequence of exchangeable shock-arrival intensities $\lambda_1, \dots, \lambda_d$ has a unique link to a so-called d*-monotone sequence* $a_0, a_1, \dots, a_{d-1}$, i.e., having nonnegative alternating discrete forward differences $(-1)^j a_k \geq 0$, $0 \leq j + k < d$; see also [11, 49, 57]:

$$a_{i-1} = \sum_{j=0}^{d-i} \binom{d-i}{j} \lambda_{j+1}, \quad i \in [d], \text{ and}$$
$$\lambda_i = (-1)^{i-1} \Delta^{i-1} a_{d-i}, \quad i \in [d].$$

  This reparametrization allows characterizing exMO distributions' survival functions as the reciprocal exponential transformations of a linear combination of the nonincreasingly ordered arguments:

$$\bar{F}(\boldsymbol{t}) = \exp\left\{-\sum_{i=1}^{d} a_{i-1} t_{[i]}\right\}, \quad \boldsymbol{t} \geq 0 \text{ with } t_{[1]} \geq \cdots \geq t_{[d]}.$$

  In addition, it allows seamlessly obtaining margin-parametrizations, since the k-marginal distribution has the reparametrization $a_0, a_1, \dots, a_{k-1}$.

A specialized stochastic representation for exMO distributions was introduced in [19], which models the order statistic of an exMO-distributed random vector $\boldsymbol{\tau}$ recursively and performs a uniformly-at-random shuffling at the end: Given the first k elements of the order statistic, simulate the first shock size and arrival time with the AM of the MO distribution's $d-k$ marginal distribution. For a shock size j, set the next j values of the order statistic to the sum of the obtained shock-arrival time and the largest of the order statistic's first k elements. Repeat this procedure recursively until having the entire order statistic, and shuffle the order statistic uniformly at random to obtain a random vector having the desired exMO distribution.

The core article [2] contributed the following results about stochastic representations of exMO distributions to the scientific literature:

- [2, Thm. 4.1 (a)] shows that the *death-counting process* Z^* of an exMO-distributed random vector $\boldsymbol{\tau}$ is Markov with infinitesimal generator matrix $Q^* = (q^*_{i,j} : 0 \leq i, j \leq d)$:

$$q^*_{i,j} = \begin{cases} -\sum_{l=1}^{d-i} \binom{d-i}{l} \sum_{k=0}^{i} \binom{i}{k} \lambda_{k+l} & , i = j, \\ \binom{d-i}{j-i} \sum_{k=0}^{i} \binom{i}{k} \lambda_{k+(j-i)} & , i < j, \\ 0 & , \text{else.} \end{cases}$$

  The first row's off-diagonal values are the exchangeable shock-size-arrival intensities introduced above, constituting an equivalent reparametrization of exMO distributions. Specialized representations of Q^* for exMO subclasses or reparametrizations are also provided in [2, Rmk. B.1–B.3].

- [2, Thm. 4.1 (b)] proposed the Markov death-counting model (MDCM) by showing that, given a sample of the *death-counting process* Z^*, a component i's death time may be defined as the earliest time by which the death count exceeds $\Pi(i)$, for a uniformly-at-random permutation Π independent of Z^*:

$$\tau_i \coloneqq \inf\{t \geq 0 \ : \ Z^*_t \geq \Pi(i)\}, \quad i \in [d].$$

  As pointed out in the contributed core article, if standard simulation techniques for Markov processes are used to represent Z^*, the resulting stochastic representation is virtually identical to the one proposed in [19]. However, for sampling, using the Markov representation suggests calculating all transition probabilities upfront for all samples, while using the recursive representation from [19] suggests calculating them on-the-fly for each sample. Moreover, having a Markov representation allows using the rich and established theory of Markov processes for further investigations, such as, e.g., in [2, Apndx. C].

3.3 Extendible Marshall–Olkin distributions

Extendible Marshall–Olkin (extMO) distributions are characterized by their stochastic representation as finite margins of an exchangeable sequence with exMO-margins.[1] Consequently, they have two reparametrizations, which, as outlined later, correspond essentially to two sides of the same coin:

- As a consequence of the exchangeable reparametrization $a_0, \dots, a_{d-1}$, there exists a sequence $a_0, a_1, \dots$ that is *completely monotone*, i.e., the first k elements are k-monotone for each $k \in \mathbb{N}$, which parametrizes a particular extension of the extMO distribution.
- De Finetti's theorem implies that every exchangeable sequence of nonnegative random variables $\tau_1, \tau_2, \dots$ defines an a.s. unique subordinator Λ; see [14, Chp. 3] and the contributed article [4, Apndx. A]. Hence, every extMO distribution may be parametrized by a particular subordinator distribution.

In this case, [17] showed that both approaches lead to the same reparametrization: First, a completely monotone sequence uniquely defines a so-called *Bernstein function*; see [58, Cor. 4.2]. A Bernstein function is a nonnegative, nondecreasing, infinitely differentiable function on the positive halfline whose derivatives have alternating signs, extended to the nonnegative halfline by introducing a fixpoint in zero. Second, the subordinators defined by exMO sequences are precisely *Lévy subordinators*, which are characterized by Bernstein functions through their Laplace functionals; see also [17, 49, 57, 59–61]. Specifically, an extMO distribution may be parametrized by a Bernstein function ψ such that

$$\begin{aligned} a_{i-1} &= \psi(i) - \psi(i-1), \quad i \in [d], \text{ and} \\ \lambda_i &= (-1)^{i-1} \Delta^i \psi(d-i), \quad i \in [d], \end{aligned}$$

and they have the so-called *Lévy frailty model (LFM)* stochastic representation

$$\tau_i \coloneqq \inf\{t \geq 0 \,:\, \Lambda_t \geq E_i\}, \quad i \in [d],$$

for a Lévy subordinator Λ associated with ψ and iid unit exponential random variables $E_1, \dots, E_d$ independent of Λ. Note that the LFM is a special case of the de Finetti representation for nonnegative exchangeable random sequences from Exp. 2.5.

The LFM allows thinking about extMO distributions in terms of an associated Lévy subordinator that consists of a killing, a drift, and a pure-jump component; see also the contributed core article [2, Sec. 2.3]: The death times are the subordinator's surpassing times over iid unit exponential barrier values independent of the subordinator. The first extreme example, corresponding to comonotonicity, is the subordinator remaining at zero until jumping to the absorbing state infinity, after an exponential waiting time, such that all barriers are surpassed simultaneously. The second extreme example, corresponding to independence, is the subordinator drifting deterministically at a constant rate without any random jump, such that all barriers are surpassed individually. Otherwise, the likelihood of a joint barrier surpassing depends on the pure-jump component, which is a compound Poisson subordinator or has a stochastic representation as the limit of compound Poisson subordinators. It is noteworthy that unbiased simulation via the LFM requires the pure-jump component to be a compound Poisson process with a suitable jump distribution that can be simulated efficiently; see [11, Sec. 3.3.3]. Finally, associating extMO distributions with Lévy subordinators allows the definition of low-parametric families spanning the dependence range between independence and comonotonicity. See the contributed core article [2, Sec. 2.4] for a curated list of popular examples.

The core article [2] contributes the following to the scientific literature:

[1] Note that extendible random vectors might be stochastically represented as margins of multiple exchangeable sequences with different distributions.

- [2, Sec. 5] outlines that the Bernstein function parametrization implies the following representation of the MDCM's Markov generator matrix Q^* of the default-counting process Z^*:

$$q^*_{i,j} = \binom{d-i}{j-i} \cdot \begin{cases} -\psi(d-i) & , i = j, \\ (-1)^{j-i-1}\Delta^{j-i}\psi(d-j) & , i < j, \\ 0 & , \text{else.} \end{cases}$$

- Furthermore, [2, Sec. 5] contributes an approximation of Q^* using integral representations of the Bernstein function ψ. This approach exploits that applying alternating difference operators to an integrand of *Lévy-Khintchine* or *Stieltjes representations* of Bernstein functions yields a convenient closed-form, nonnegative integrand (in the respective integral).[2] Consequently, Q^* can be approximated with numerical integration techniques after switching the order of integrating and applying the alternating difference operator and multiplying the binomial coefficient. In addition, given the inner values of the Q^*'s first row, the article provides a recursive scheme in [2, Thm. 5.10] and boundary conditions to approximate the remaining entries; see [2, Rmk. 5.11].

- For this, [2, Sec. 5] demonstrates in a numerical study that the naïve alternative of calculating Q^* by recursively applying the alternating difference operator is not numerically stable in higher dimensions due to loss of significant digits. In contrast, a repetition of this study demonstrates for the suggested approach above a numerically stable approximation of the Markov generator matrix for a wide range of test cases. These results are supported by the theoretical results [2, Lem. 5.8 and 5.9] showing that integrands corresponding to the inner values of Q^*'s first row, which are relevant for the numerical integration, are continuous and bounded under mild assumptions. In conclusion, this demonstrates that the novel approach to approximate the Markov generator Q^*, and implicitly the exchangeable shock-size-arrival intensities, provides the first broadly applicable option to use any of the presented models for MO or exMO distributions for simulating extMO distributions specified via low-parametric Bernstein functions in high dimensions.

- Finally, [2, Sec. 6] provides a benchmark study that compares the existing simulation algorithms ESM, AM, LFM, and MDCM, primarily for their speed and setup activities' proportions of runtime. All algorithms were implemented in C++, wrapped in R. They were compared for several selected families using comparable configurations; the article shows only results about the *exponential family* explicitly.[3] The high-level summary of the study's findings is as follows: First, the specialized algorithms based on the LFM and MDCM are significantly faster than the general algorithms based on the ESM and AM, and they do not suffer from parameter storage issues. Second, the MDCM was slower than the LFM for most tested configurations. However, the gap in the runtime of both algorithms is expected to shrink for an increasing number of samples due to the MDCM's comparatively high setup costs. Third, the MDCM is more broadly applicable than the LFM, which requires a suitable compound Poisson representation of the subordinator's pure-jump component. Finally, the MDCM's runtime is bounded, while the LFM's runtime may explode towards specific configurations; see [2, Apndx. C].

Overall, [2, Sec. 5] derives a numerically stable approximation of the MDCM infinitesimal Markov generator matrix for extMO distributions and proposes a specialized simulation algorithm for extMO distributions. This algorithm is shown to be competitive compared to alternative algorithms through a benchmark study in [2, Sec. 6].

[2] Every Bernstein function has a Lévy–Khintchine representation, and every *complete Bernstein functions* has a Stieltjes representation; see [60].

[3] The *exponential family* corresponds to a Lévy subordinator with possible killing and drift components and pure-jump component that is a compound Poisson subordinator with exponential jumps.

3.4 Hierarchical Marshall–Olkin distributions

In light of Chapter 2, the novel simulation algorithm for extMO distributions opens possibilities to implement large-scale factor models for portfolio credit risk modeling with hierarchical Marshall–Olkin (hMO)-distributed default times. There are multiple approaches to defining hMO-distributed random vectors based on extMO-associated factors. For example, [49, Sec. 5.2] defines an hMO distribution by replacing the (global) subordinator in the LFM stochastic representation with individual subordinators that sum independent time-changed global and sector-dependent subordinators. In another example, [48, Sec. 4.2] defines an hMO distribution using a stochastic representation via a minima construction with independent extMO distributed random vectors for global and sector-based factors. HMO distributions are not a primary topic of this thesis, but it is noteworthy that many hierarchical constructions have stochastic representations as refactorings of extMO distributed random vectors. Consequently, simulation algorithms for hMO distributions may be derived from simulation algorithms for extMO distributions; see also the contributed core article [2, Sec. 2.5].

4 Generalized Marshall–Olkin distributions

This chapter primarily discusses the contribution of the further articles [3, 4] that characterize the survival functions of generalized Marshall–Olkin (gMO) distributions and derive the implied de Finetti representation of extendible generalized Marshall–Olkin (extgMO) sequences. The chapter briefly introduces the topic and necessary background theory to discuss the contributed articles' contribution. For detailed introductions, the interested reader is referred to the articles, the provided references, and comprehensive monographs [11, 62, 63].

4.1 Generalized Marshall–Olkin distributions

Marshall–Olkin (MO) distributions, having stochastic representations via the exogenous shock model (ESM) with exponentially distributed shock-arrival times, can be generalized to gMO distributions by allowing nonnegative, unbounded, and continuous shock-arrival times in a generalized exogenous shock model (gESM). Formally, the class of admissible hazard functions $\mathcal{H}$ is

$$\mathcal{H} \coloneqq \Big\{H : [0,\infty) \to [0,\infty) \ : \ H \in \mathcal{C}^{(0)},\ \Delta H \geq 0,\ H(0) = 0\Big\}.$$

Additionally, for a.s. finite margins, the class of admissible marginal hazard functions $\mathcal{H}_0$ is

$$\mathcal{H}_0 \coloneqq \{H \in \mathcal{H} \ : \ H(\infty-) = \infty\}.$$

Finally, the survival functions of gMO distributions are

$$\bar{F}(\boldsymbol{t}) = \exp\left\{-\sum_{\emptyset \neq I \subseteq [d]} H_I\Big(\max_{i\in I} t_i\Big)\right\}, \quad \boldsymbol{t} \geq 0,$$

for $H_I \in \mathcal{H}$ with $\sum_{I \ni i} H_I \in \mathcal{H}_0$, $\forall i \in [d]$. It is readily observable that random vectors with this survival function $\bar{F}$ have a stochastic representation as a *gESM* such that for independent random variables $Z_I \sim H_I$, $\emptyset \neq I \subseteq [d]$, the following random vector $\boldsymbol{\tau}$ has survival function $\bar{F}$:

$$\tau_i \coloneqq \min\{Z_I \ : \ I \ni i\}, \quad i \in [d].$$

Generalizing MO distributions by allowing nonnegative, unbounded, and continuous shock-arrival times in the associated gESM has been suggested in multiple articles. For example, the special case with proportional hazard functions was suggested in [24] to solve a modified Marshall–Olkin lack of memory (MOLOM) property with other binary operators for time accumulation; see Section 2.1. For the bivariate case, the gESM was suggested and analyzed in [12, 22]. The multivariate case is discussed in [23]. Furthermore, the associated survival copulas were investigated for the bivariate case in [27], for the exchangeable Armageddon case, which only allows individual and global shocks, in [64], and for the exchangeable case in [25]; an exhaustive treatment of the exchangeable case can be found in [62]. Finally, the contributed articles [3, 4] consider a slightly restricted class with shock-arrival hazard functions in $\mathcal{H}$ and marginal hazard functions in $\mathcal{H}_0$. The shock-arrival hazard functions' finiteness, continuity, fixpoints in zero, and the overarching marginal finiteness condition restrict the class to nonnegative, unbounded, and continuous random vectors.

There are multiple noteworthy alternatives or broader approaches to generalize classical MO distributions, which are not discussed in detail in this thesis or the contributed articles but are noteworthy; see also [65].

For example, [66, 67] discuss *Archimax copulas* that enclose extreme-value (EV) copulas and Archimedean copulas, which implies they also enclose MO survival copulas. A particular subclass of Archimax copulas arises as the survival copulas of (classical) ESM's whose (exponential) shock-arrival times are scaled by an independent positive random variable; see [67, Prop. 3.1 and Alg. 4.1]. This subclass can be embedded in a class proposed in [68] called *mixed (multivariate) gMO distributions* whose members have a stochastic representation as a gESM-represented, gMO-distributed random vector scaled by an independent positive random variable. Moreover, a broad approach allowing arbitrary dependency between the random shock-arrival times of the gESM and alternative aggregations is proposed in [69]. Alternative generalizations of extendible Marshall–Olkin (extMO) distributions that replace the Lévy subordinator in the Lévy frailty model (LFM) by subordinators that are *strongly* or *weakly infinitely divisible w.r.t. time* or, more general, *infinitely divisible (ID)* were discussed in [36, 37], respectively, and correspond to min-stable multivariate exponential (MSMVE), exponential minima (EM), and min–infinitely divisible (min-ID) distributions, respectively; see also [63]. Finally, other alternative approaches that generalize the ESM by associating each shock-arrival time with a realized shock triggered by an event-related Poisson process and replacing these Poisson processes with *Cox processes* are proposed in [45–47].

4.2 Exchangeable generalized Marshall–Olkin distributions

Analogous to classical exchangeable Marshall–Olkin (exMO) distributions, exchangeable generalized Marshall–Olkin (exgMO) distributions are characterized by the property that their associated shock-arrival distributions are equal for shock-sets with matching cardinalities, see [62, Prop. 3.1.2]., i.e.,

$$|I| = |J| \;\Leftrightarrow\; H_I = H_J \quad \forall \emptyset \neq I, J \subseteq [d].$$

Consequently, exgMO distributions can also be reparametrized to d *exchangeable shock-arrival hazard functions* H_i, $i \in [d]$. The reparametrization was first investigated on the survival copula level in [25, Thm. 1.1] and translated to the level of survival functions in the contributed article [4, Lem. 1]. The exchangeable hazard functions are uniquely linked to a sequence of hazard functions $A_0, \dots A_{d-1} \in \mathcal{H}$ with $A_0 \in \mathcal{H}_0$ that is pointwise d-monotone:

$$\begin{aligned} A_{i-1}(t) &= \sum_{j=0}^{d-i} \binom{d-i}{j} H_{j+1}(t), \quad t \geq 0,\ i \in [d], \text{ and,} \\ H_i(t) &= (-1)^{i-1} \Delta^{i-1} A_{d-i}(t), \quad t \geq 0,\ i \in [d]. \end{aligned}$$

The reparametrization characterizes exgMO distributions' survival functions as the reciprocal exponentials of sums of nonincreasingly ordered, individually transformed arguments, with order-independent hazard-function-transformation fulfilling specific and monotonicity requirements:

$$\bar{F}(\boldsymbol{t}) = \exp\left\{ -\sum_{i=1}^{d} A_{i-1}(t_{[i]}) \right\}, \quad \boldsymbol{t} \geq 0 \text{ with } t_{[1]} \geq \cdots \geq t_{[d]}.$$

This reparametrization of exgMO distributions is significant for several reasons:

- It allows the identification of distributions by their survival functions as exgMO distributions. For example, [25, Prop. 3.1] shows that the additive frailty model (AFM), a generalization of the LFM, see Chapter 3, with additive subordinators instead of Lévy subordinators, implies exgMO distributed random vectors.
- It allows the investigation of properties of exgMO distributions. For example, [25, Prop. 2.2] determines that exgMO survival copulas have the extreme-value property if and only if they are exMO survival copulas, and [62, Thm. 3.4.2] identifies the *radial symmetric* subclass of exgMO survival copulas.

The contributed article [3] generalized this characterization for non-exchangeable gMO distributions. [3, Thm. 1] shows that a continuous survival function of a nonnegative random vector is that of a gMO distribution if and only if it is the reciprocal exponential function of sums of nonincreasingly ordered, individually transformed arguments, with order-dependent hazard-function-transformations fulfilling certain monotonicity and continuity requirements:[1]

$$\bar{F}(\boldsymbol{t}) = \exp\left\{-\sum_{i=1}^{d} A_{i-1}^{\pi}(t_{\pi(i)})\right\}, \quad \boldsymbol{t} \geq 0,\ \pi \in \mathcal{S}_d,\ t_{\pi(1)} \geq \cdots \geq t_{\pi(d)},$$

with

$$A_{i-1}^{\pi}(t) = \sum_{I:\pi(i)\in I\subseteq\pi(i,\ldots,d)} H_I(t), \quad t \geq 0,\ i \in [d], \pi \in \mathcal{S}_d,$$

and

$$H_I(t) = \sum_{J:m\in J\subseteq I} (-1)^{|J|-1} A_{d-|I|+|J|-1}^{\pi_J}(t), \quad t \geq 0,\ \emptyset \neq I \subseteq [d],\ m \in I,$$

where $\{\pi_J : m \in J \subseteq I\}$ is an arbitrary family of permutations with $\pi_J(1, \ldots, |J \cup ([d] \setminus I)|) = J \cup ([d] \setminus I)$ and $\pi_J(|J \cup ([d] \setminus I)|) = m$.

The article provides applications of this characterization result in [3, Sec. 3]:

- The AFM can be generalized to a hierarchical framework in the spirit of the hierarchical models suggested in [48, Sec. 4.2] by considering component-specific additive trigger subordinators $\Lambda_1, \ldots, \Lambda_d$ that are linear recombinations of a shared pool of independent additive subordinators $\boldsymbol{\Upsilon} = (\Upsilon_1, \ldots, \Upsilon_d)$ with parametrizing Bernstein function families $\psi_1, \ldots, \psi_n$, and coefficient matrix $\Theta = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)^\top$:

  $$\Lambda_{i,t} = \boldsymbol{\theta}_i^\top \boldsymbol{\Upsilon}_t, \quad i \in [d].$$

  Using the characterization theorem, one can easily determine that the following obtained survival function is gMO:

  $$\bar{F}(\boldsymbol{t}) = \exp\left\{-\sum_{i=1}^{d}\sum_{k=1}^{n}\left[\psi_{k,t_{\pi(i)}}\left(\sum_{j=1}^{i}\Theta_{\pi(i),k}\right) - \psi_{k,t_{\pi(i)}}\left(\sum_{j=1}^{i-1}\Theta_{\pi(i),k}\right)\right]\right\},$$

  $$\boldsymbol{t} \geq 0,\ \pi \in \mathcal{S}_d\ :\ t_{\pi(1)} \geq \cdots \geq t_{\pi(d)}.$$

  Following, [3, Sec. 3] demonstrates how to use the characterization theorem to calculate the global shock event's probability from the additive subordinators' characterizing Bernstein function families.

- [70] discusses so-called *regenerative composition structures*. These are Markovian sequences of compositions of positive integers fulfilling specific consistency and stationarity properties. [70, Thm. 5.2] proves that they are uniquely linked to extMO sequences generated via the LFM as follows: Traverse the increasingly ordered unit exponential barrier values and recursively count until the range between two adjacent (ordered) barrier values contains an element of the subordinator path's closure. Regenerative composition structures can be described by so-called *decrement matrices* representing the Arnold model (AM)'s first-shock probabilities for all margins. As the characterization theorem's specialization for MO distributions implicitly links Lévy subordinators and decrement matrices, it can be used to generalize the theory of regenerative composition structures.

[1] The stated result in [3, Thm. 1] is in terms of survival functions, but a translation in terms of hazard function follows as in [4, Lem. 1].

4.3 Extendible generalized Marshall–Olkin distributions

ExtgMO distributions have a stochastic representation as the margin of an exchangeable sequence with exgMO distributed margins. Recall that, as outlined in Chapter 3, exMO sequences' distributions are uniquely linked to Lévy subordinators' distributions and their characterizing Bernstein functions via the LFM. Furthermore, [25, Prop. 3.1] proved that extending the LFM to the AFM by using additive subordinators produces exgMO sequences.

The contributed article [4] contributed the following to the scientific literature: [4, Thm. 1] shows that every exgMO sequence $\tau_1, \tau_2, \ldots$ uniquely defines an additive subordinator Λ on the same probability space characterized by a family of Bernstein functions $\{\psi_t : t \geq 0\}$ with specific consistency properties:[2]

$$\tau_i = \inf\{t \geq 0 \ : \ \Lambda_t \geq E_i\} \text{ a.s.}, \quad i \in \mathbb{N},$$

for an implicitly defined sequence of iid unit exponential random variables $E_1, E_2, \ldots$ independent of Λ.[3] Furthermore, using [25, Prop. 3.1], the associated survival function takes the form

$$\bar{F}(\boldsymbol{t}) = \exp\left\{-\sum_{i=1}^{d} \psi_{t_{[i]}}(i) - \psi_{t_{[i]}}(i-1)\right\}, \quad \boldsymbol{t} \geq 0, \ t_{[1]} \geq \cdots \geq t_{[d]}.$$

This contribution's significance is twofold: First, it answers an open research question from [62, p. 147]. Second, it was a stepping stone toward subsequent generalizations using nonnegative, nondecreasing, and ID càdlàg subordinators to obtain min-ID sequences in [37].

In addition, [4, Sec. 3] demonstrates for an exemplary extgMO distribution how to determine an exgMO sequence's subordinator distribution and how to recover the subordinator explicitly.

[2] A family of Bernstein functions $\{\psi_t : t \geq 0\}$ parametrizes an additive subordinator's distribution if and only if $\psi_0 \equiv 0$, $\psi_s - \psi_t$ is a Bernstein function for all $s > t \geq 0$, and the mapping $t \mapsto \psi_t(x)$ is continuous for all $x \geq 0$; see [4, Eq. 12].

[3] Technically, obtaining the iid sequence of unit exponential random variables requires introducing an independent sequences of unit uniform random variables to uniformly-at-random interpolate jumps of the subordinator's reciprocal exponential when transforming the sequence to iid unit uniform random variables; see [4, Cor. 2].

5 Upper semilinear copulas

This chapter primarily discusses the contribution of the core article [1] that derives stochastic representations for multiple subclasses of upper semilinear (USL) copulas. Necessary background information is briefly discussed. For a detailed introduction, the reader is referred to [7–11], [28], and the article itself.

5.1 Copula compatibility problems

Recall that copulas are standardized distribution functions with continuous uniform margins. However, they are also functions and have a purely analytical characterization with applications outside classical probability theory: A copula is a d-variate function that is *grounded*, has the *uniform margin property*, and is d*-increasing*.[1] Moreover, bivariate copulas as functions are also *conjunctors* or *t-seminorms*, and extendible copulas give rise to *conjunctive aggregation operations*.[2] For this reason, many copulas were first derived and studied as analytical functions without knowing any stochastic representation. A well-known example is the class of Archimedean copulas proposed in the field of probabilistic metric spaces without a stochastic representation, see [72, 73], and for which a comprehensive stochastic representation was derived decades later in [34]. However, knowing a stochastic representation is useful even without interest in sampling, since it can be used to analyze the function and derive properties.

There exists a rich body of literature that, among interesting candidate functions, derives copulas exhibiting specific geometrical features or characterizes copulas as solutions to *compatibility problems*. The classical compatibility problem asks under which circumstances $\binom{d}{k}$ k-copulas can be embedded as k-margins into a d-copula, see [8, p. 107]. However, more generally, a compatibility problem asks under which circumstances certain marginal probabilities can be embedded into a d-copula with specific features; see [74, Sec. 1.8]. An extensive compilation of works about copulas solving geometric compatibility problems, e.g., having a given support or given marginal horizontal, vertical, diagonal, or affine sections, is [74, Sec. 1.7.3]; see also [8, Chp. 3]. [75] outlines that marginal diagonal sections comprise many interesting quantities, e.g., *tail-dependence coefficients*. Consequently, compatibility problems for given marginal diagonal sections are frequently studied; see [75–83]. For example, [82] derives conditions on a diagonal such that the pointwise infimum over all copulas with this diagonal, the *Bertino quasi-copula*, is itself a copula.

5.2 Upper semilinear copulas

Bivariate *lower semilinear (LSL)* and *USL* copulas were first discussed in [27]. The underlying question was: given a bivariate copula's diagonal, can an exchangeable copula be constructed that is linear on all sections parallel to one axis that connect a point on a lower (LSL) or upper (USL) marginal-boundary hyperplane,

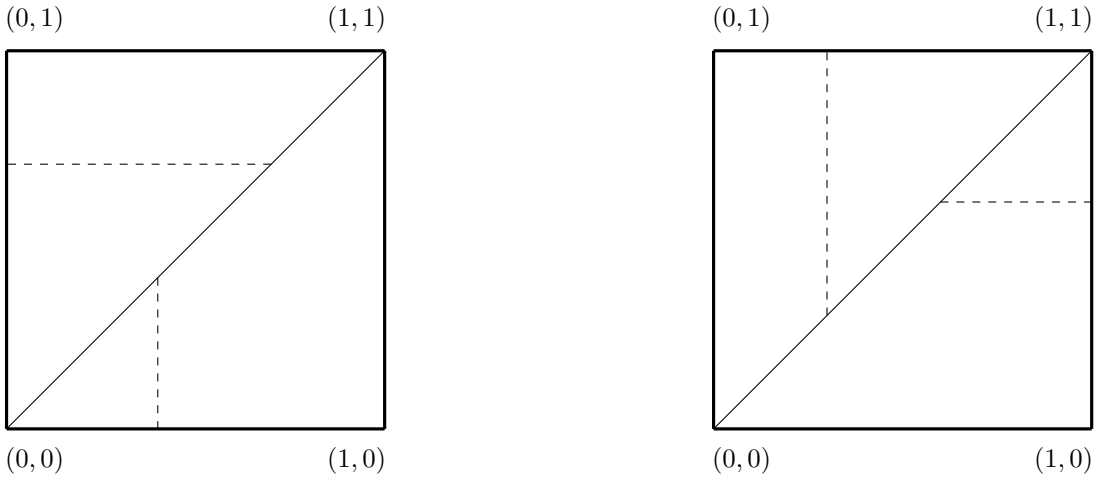
[1] A function $C : [0,1]^d \to [0,1]$ is called *grounded* if $u_i = 0$ implies $C(\boldsymbol{u}) = 0$, has the *uniform-margin* property if $u_j = 1 \forall j \neq i$ implies $C(\boldsymbol{u}) = u_i$, and is d*-increasing* if the associated volume of arbitrary d*-boxes*, calculated with the inclusion-exclusion formula pretending C would be a distribution function, is nonnegative; see [8, Def. 2.10.5]

[2] A *conjunctor* is a monotone binary operator on the the unit interval whose binary restriction is the boolean conjuction, a *t-seminorm* is a conjunctor with neutral element 1, a *t-norm* is an associative and commutative t-seminorm, and a *conjunctive aggregation operation* is a variadic operator on unit intervals that maps zero-vectors to zero and one-vectors to one, has the uniform margin property, is component-wise nondecreasing, and bounded by the *Fréchet–Hoeffding upper bound*; see [64, 71].

respectively, to a point on the diagonal section; see Fig. 5.1. [27, Lems. 2 and 3] showed that semilinearity allows a recursive diagonal-based deconstruction leading to copulas of the form:

$$C^{\mathsf{LSL}}(\boldsymbol{u}) = u_{(1)} \cdot \frac{\delta_{C^{\mathsf{LSL}}}(u_{(2)})}{u_{(2)}}, \quad \boldsymbol{u} \in [0,1]^2,\ u_{(1)} \le u_{(2)}, \text{ and}$$
$$C^{\mathsf{USL}}(\boldsymbol{u}) = u_{(1)} - (1-u_{(2)})\frac{u_{(1)} - \delta_{C^{\mathsf{USL}}}(u_{(1)})}{1-u_{(1)}}, \quad \boldsymbol{u} \in [0,1]^2,\ u_{(1)} \le u_{(2)}.$$

In addition, bivariate USL copulas are the survival copulas of LSL copulas. Moreover, [27, Thm. 4] derived conditions for bivariate copula diagonals to admit a lower semilinear copula. Finally, the authors show in [27, Sec. 4] that lower semilinear copulas are symmetric *Marshall copulas* defined in [12]. Subsequently, they are the survival copulas of bivariate exchangeable generalized Marshall–Olkins (exgMOs) distributions and have a stochastic representation via bivariate generalized exogenous shock models (gESMs).



(a) lower semilinear **(b)** upper semilinear

Figure 5.1 Illustration of "linear segments" (dashed lines) of bivariate LSL and USL copulas; from [1, p. 266], cf. [27, p. 65].

A multivariate extension of USL copulas was defined and characterized in [28]. A USL copula must be exchangeable and linear on all segments connecting the diagonal to an upper marginal-boundary hyperplane with a constant lower component. Formally, these segments are for the permutations of segments of the form

$$\mathcal{S}_{d,\boldsymbol{v}} = \left\{ \begin{pmatrix} \boldsymbol{u} \\ \mathbf{1} \end{pmatrix} \in [0,1]^{j+(d-j)} \ :\ u_i = \lambda v_1 + (1-\lambda)v_i,\ i \in \{1,\dots,j\},\ \lambda \in [0,1] \right\},$$

for an ordered vector $\boldsymbol{v} \in [0,1]^j$ with $v_1 \le \cdots \le v_j = 1$. [28] showed that semilinearity allows a recursive diagonal-based deconstruction of USL copulas C with marginal diagonals $\delta_2, \dots, \delta_d$:[3]

$$\begin{aligned} C(\boldsymbol{u}) &= \frac{1-u_{(d)}}{1-u_{(1)}} \cdot \delta_{d-1} + \frac{u_{(d)} - u_{(1)}}{1-u_{(1)}} C_{d-1}\left(u_{(1)} + \frac{1-u_{(1)}}{u_{(d)} - u_{(1)}} \left(\boldsymbol{u} - u_{(1)}\mathbf{1} \right) \right) \\ &= \frac{\sum_{i=1}^{d} (u_{(i+1)} - u_{(i)}) \delta_i(u_{(1)})}{1-u_{(1)}}, \quad \boldsymbol{u} \in [0,1], u_{(1)} \le \cdots \le u_{(d)}. \end{aligned}$$

This representation of USL copulas highlights two significant, interesting properties: First, they are fully specified by their diagonals. Second, they have a simple closed-form expression, given that the diagonal

[3] The following formula uses the convention $u_{(d+1)} \coloneqq 1$.

functions are not complicated. [28, Thm. 1] provides a characterization of USL copulas in the form of specific necessary and sufficient coefficients to marginal diagonal functions $\delta_2, \dots, \delta_d$ to define a USL copula, which is equivalent to the d-increasingness property under the exchangeability and semilinearity assumption. Finally, the corresponding proof showed that realizations of USL copulas must be concentrated on at most two distinct random values, e.g., if C is a USL copula and $U \sim C$, then

$$\mathbb{P}(U_1 \neq U_2, U_2 \neq U_3, U_1 \neq U_3) = 0.$$

On a macro level, the core article [1] contributed the following findings to the scientific literature:

- Recall that bivariate USL copulas are the survival copulas of bivariate LSL copulas. As bivariate LSL copulas are survival copulas of bivariate exgMO distributions, bivariate USL copulas can also be generalized into higher dimensions as copulas of exgMO distributions. Consequently, it was an open question if those two generalizations intersect and if USL copulas have a stochastic gESM representation. [1, Sec. 5] investigates this question and determines the intersection between both generalizations; see Section 5.3.
- Except for trivial cases or the bivariate subclass, no explicit stochastic representations for USL copulas have been proposed. As stochastic representations are frequently used to derive properties of copulas, they are helpful even if one is not primarily interested in using the copula in a stochastic context. [1] provides stochastic representations for the intersecting subclass with exgMO copulas in [1, Thm. 3], see Section 5.3, the subclass with identical marginal diagonal functions suggested in [28, Cor. 1], see Section 5.4, and an extendible subclass in [1, Cor. 6], see Section 5.5, which highlights the usefulness of stochastic representations by using it to show this subclass is *radially symmetric*.

5.3 Upper semilinear copulas and exchangeable generalized MO distributions

Recall that bivariate USL copulas are copulas of bivariate exgMO distributions. Furthermore, each exgMO distribution has a stochastic representation via an exchangeable gESM. Hence, the corresponding survival copula has a representation via a *dual gESM* as featured in [25]: Let $\hat{C}$ be an exgMO survival copula, then there exists a stochastic realization $U \sim \hat{C}$ with

$$U_i \coloneqq \max\{Z_I : I \ni i\}, \quad i \in [d],$$

for independent, $[0,1]$-valued $Z_I \sim F_I$, $\emptyset \neq I \subseteq [d]$ with specific distribution functions; shocks that are almost surely equal to zero are *ineffective* and may be omitted. A natural question is which multivariate USL copulas are the copulas of exgMO distributions and, by extension, have a stochastic exgMO representation.

The core article [1, Sec. 5] contributed the following to the scientific literature:

- [1, Thm. 3] shows that a USL copula is an exgMO copula if and only if the associated dual gESM has no effective shocks with associated sets missing more than one component. This result implies that at most one component may deviate from the joint minimum.
- [1, Cors. 3 and 4] derive simplified characterizations for this subclass, equivalent to recovering valid distribution functions for an associated gESM.

5.4 Upper semilinear copulas with identical marginal diagonal functions

[28, Sec. 3] investigated a subclass of USL copulas for which all marginal diagonals are identical, i.e., $\delta_2 = \cdots = \delta_d \equiv \delta$. In particular, they provide a simplified characterization theorem [28, Cor. 3] and several examples in [28, Sec. 4].

The core article [1, Sec. 6] contributed the following to the scientific literature:

- [1, Lem. 8] uses that this particular subclass permits only realizations for which at most one component deviates from the joint maximum to derive a conditional sampling approach based on conditional sampling of bivariate copulas; see [11, Alg. 1.2]. Moreover, it also derives the required probabilities and (conditional) distribution functions. The sampling algorithm is summarized in Alg. 1.
- [1, Cor. 5] uses the novel stochastic representation to derive another simplified characterization theorem for this subclass, equivalent to recovering proper probabilities and distribution functions required for Alg. 1. This results is also an example of how to use stochastic representations to derive properties of copulas.

Algorithm 1 Sampling algorithm for an USL copula C with equal multivariate diagonals $\delta_j \equiv \delta$, $j \in \{2, \dots, d\}$, see [1, Alg. 1].

input An admissible d-diagonal δ.
output A sample from the USL copula with identical marginal diagonal δ.
function SAMPLEUSLC(δ)
  Draw $I \sim \mathrm{Bernoulli}(p)$ with $p = d \int_0^1 \frac{x-\delta(x)}{1-x} \, \mathrm{dx}$.
  if $I = 0$ **then**
    Draw $U_\wedge \sim \Big(\delta(u) - d \int_0^u \frac{x-\delta(x)}{1-x} \, \mathrm{dx}\Big)/(1-p)$.
    Set $U_1 = \cdots = U_d = U_\wedge$.
  else
    Draw $U_\wedge \sim d\Big((u - \delta(u)) + \int_0^u \frac{x-\delta(x)}{1-x} \, \mathrm{dx}\Big)/p$.
    Draw $U_\vee \sim \mathcal{U}_{[U_\wedge, 1]}$.
    Draw K uniform from the set $\{1, \dots, d\}$.
    Set $U_K \coloneqq U_\wedge$ and $U_j \coloneqq U_\vee$, $j \neq K$.
  end if
  return $\boldsymbol{U} = (U_1, \dots, U_d)'$.
end function

5.5 An extendible family of upper semilinear copulas

Extendible subfamilies are an essential cornerstone of every copula family. They are interesting for many reasons, for example, they imply a stochastic de Finetti representation, a *conditionally iid* representation of the random vector, which is both simple and natural to many applications, see [63, Sec. 1.2]. In addition, specific features of copula classes sometimes allow a *guess-and-verify* approach to finding the de Finetti representation. Hence, for instances without known stochastic representations or few known examples, finding the extendible subclass can be a promising approach to finding new stochastic representations and interesting exemplary members of this class.

The core article [1, Sec. 7] contributes the following to the scientific literature:

- [1, Cor. 6] derives an extendible subclass corresponding to the following stochastic representation: Draw an random probability in $[1/2, 1]$, draw a bivariate sample from the independence copula, and independently conduct for each component a Bernoulli experiment conditioned on the sampled success probability to decide whether to set the component to the first or second bivariate sample. For a random probability Q, the corresponding copula C is

$$C_d(\boldsymbol{u}) = u_{(1)} \left(\mathbb{E}[Q^d] + \mathbb{E}\Big[(1-Q)^d\Big] + \sum_{j=2}^{d} u_{(j)} \Big(\mathbb{E}\Big[Q(1-Q)^{j-1}\Big] + \mathbb{E}\Big[(1-Q)Q^{j-1}\Big] \Big) \right).$$

  This subclass extends the *bivariate Dirichlet copulas* from [62, Theorem 3.5.3].

- [1, Cor. 7] uses this stochastic representation to show that the corresponding extendible subclass is radially symmetric. First, this is an excellent example of how stochastic representations may be used to derive properties of copulas as analytical functions. Second, this highlights that this extendible subclass is a meaningful multivariate extension of the bivariate Dirichlet copula, since it also extends the radial symmetric property.

6 Outlook

This section aims to identify open questions and potential further research topics arising from the contributed articles [1–5].

Simulation of high-dimensional, non-trivial Marshall–Olkin (MO) distributions is furthered by the simulation algorithm for extendible Marshall–Olkin (extMO) distributions in the contributed core article [2]. However, there remain plenty of research questions and open problems involving the simulation of MO distributions:

- The stochastic representations for the Markovian default-indicator processes of exchangeable Marshall–Olkin (exMO) distributions proposed in [2] is not unique. In particular, *possibly looping* stochastic representations can be derived from the Arnold model (AM) by splitting the shock-set-sampling into two parts: First, sample the shock size with appropriate probabilities. Second, sample the concrete shock set uniformly at random. For simulating the default-counting process, we do not need to know the concrete shock sets immediately. However, we need to be able to sample the actual size of newly dead components given the original shock size and the number of already dead components, which can be done using the *hypergeometric distribution*.[1] While a simulation algorithm based on this stochastic representation certainly requires on average more iterations to reach the absorbing state, it also requires only calculating the first row of the exMO Markov generator matrix and allows for simpler parameter representations. However, implementing this algorithm has challenges: Sampling algorithms for hypergeometric distributions are not part of the standard repertoire of all programming languages. In addition, there are multiple options to sample from hypergeometric distributions, including elaborate *acceptance/rejection* schemes and *alias sampling*; see [84]. In this particular instance, it would be interesting to explore this simulation algorithm with all its various implementation options and compare it to the suggested extMO simulation algorithm from [2].

- [85] develops a simulation algorithm for *continuous max-id processes*, and *exchangeable max-id sequences* in particular, which implies a simulation algorithm for MO distributions as a special case. However, the details for this particular example need to be worked out. In particular, how to map the distribution parameter to the algorithm parameter, and how it compares to the other existing simulation algorithms.

- A promising idea for future research is using the novel Markov representation or the simulation algorithm of extMO distributions for applied problems. The following straightforward corollary shows that the Markov representation of extMO distributions' death-counting processes implies that non-path-dependent derivatives on portfolio loss processes of extMO-distributed default times can be priced almost in closed form, requiring only approximations to calculate the Markov generators and matrix exponentials.

  Corollary 6.1. *Consider d-variate extMO-distributed multivariate default times with infinitesimal Markov generator Q^* for the death-counting process Z^*. Then, for a deterministic* recovery fraction *$\delta \in [0,1]$, the associated portfolio loss process $L = (1-\delta)/d \cdot Z^*$ is Markov with*

  $$\mathbb{P}(L_t = i/d \cdot (1-\delta)) = \boldsymbol{\delta_0}^\top \mathrm{e}^{tQ^*} \boldsymbol{\delta_i}, \quad t \geq 0, i \in [d]_0,$$

  where $(\boldsymbol{\delta_i})_j = 1_{\{i=j\}}$, $\forall i, j \in [d]_0$.

[1] The *hypergeometric distribution* describes the number of red balls drawn among a number of draws without replacement from a bowl with specific amounts red and black balls.

Furthermore, a significant achievement of [2] is making the simulation of high-dimensional extMO distribution or derived hierarchical Marshall–Olkin (hMO) distributions feasible. Hence, an application that stands to reason is the calibration and pricing of portfolio credit derivatives with Cor. 6.1 or Monte-Carlo simulations.

- The practical application of the simulation algorithm involves writing statistical software, which is ideally thoroughly tested. Besides possible implementation errors, implementing simulation algorithms requires avoiding non-obvious pitfalls which the implementers may not be aware of. An example that stands to reason is that some implementations require approximations whose stability needs to be assessed. Less intuitive examples are related to the nature of random number generators, e.g., linear congruential random number generators (RNGs) can theoretically iterate through their periods and produce ties — which is considerably more likely for multivariate random vector simulation algorithms that increment the underlying bit-sequence multiple times per sample. Bottom line, implementations of simulation algorithms should be tested statistically for the samples' independence and their distribution assumptions, see [2, Apndx. D]. The tight constant for the multivariate *Dvoretzky–Kiefer–Wolfowitz (DKW) inequality* derived in [86] is an interesting option to test distributional assumptions. On the face of it, using the DKW inequality does not require additional theory. However, if used in many tests concurrently, how to choose the appropriate thresholds should be explored. Additionally, as the optimal constant is valid asymptotically, optimal sample sizes for testing should be explored.

- It is an open question if and how the AM and the Markov death-set model (MDSM) for MO distributions can be generalized for generalized Marshall–Olkin (gMO) distributions, and if and how the Markov death-counting model (MDCM) for exMO distributions can be generalized to exchangeable generalized Marshall–Olkin (exgMO) distributions. For this, consider the following conjecture:

  Conjecture 6.2. *Consider a gMO distribution with infinitesimal hazard shock-arrival rates $\lambda_I(t)$, $t \geq 0$, $\emptyset \neq I \subseteq [d]$ and $\lambda = \sum_{\emptyset \neq I \subseteq [d]} \lambda_I$. Then, the following random vector $\boldsymbol{\tau}$ has the desired distribution:*

  $$\tau_i = \min\{W_1 + \cdots + W_j \ : \ Y_j \ni i\}, \quad i \in [d],$$

  where $\{(W_1 + \cdots + W_j, Y_j) : j \in \mathbb{N}\}$ are points of a Poisson random measure (PRM) *with intensity measure*

  $$\mu(\mathrm{d}s \times \mathrm{d}\boldsymbol{y}) = \lambda(s)\mathrm{d}s \times \left(\sum_{\emptyset \neq I \subseteq [d] : I \in \mathrm{d}\boldsymbol{y}} \frac{\lambda_I(s)}{\lambda(s)} \right).$$

  If this conjecture is true, this would also open the door to generalizing the MDSM for gMO and the MDCM for exgMO distributions. However, it also would raise the question for which gMO distributions those representations could feasibly be used for sampling. From the second part of the PRM's intensity measure, it is apparent that proportional hazard rates would be more straightforward.

The contributed article [3] generalized a characterization for exgMO survival copula from [25] to gMO distributions. [25, Prop. 2.2] used the characterization for exgMO survival copulas to show that their extreme-value (EV) subclass is precisely the class of exMO survival copulas. Consequently, it would be interesting to generalize this result and determine whether the EV subclass of gMO survival copulas is also the class of MO survival copulas.

The contributed core article [1] provided several stochastic representations of multivariate upper semilinear (USL) copulas. These representations allow using probabilistic means to reason about these copulas and to use stochastic representations to derive analytical properties. Moreover, the extendible class also provided a large class of nontrivial USL copulas. The article leaves the following possible avenues to continue research:

- The extendible subclass described in [1] describes a model which draws with replacement from a bivariate independent sample with a possibly random probability. However, it remains an open question of how the entire extendible subclass looks and whether the semilinearity can be linked explicitly to a property of the de Finetti representation. To this end, consider the following conjecture:

Conjecture 6.3. *Let F be a distribution function on $[1/2, 1]$, let $\tilde{C}_2$ be a bivariate USL copula, and consider the following stochastic model. Let $Q \sim F$, $V \sim \tilde{C}_2$, and $\{J_i\}_{i\in\mathbb{N}}$ a family of conditionally iid Bernoulli distributed random variables with random success parameter Q. Define*

$$U_i \coloneqq J_i V_1 + (1 - J_i)V_2, \quad i \in \mathbb{N}.$$

Then, for each $d \geq 2$, the random vector $\boldsymbol{U} = (U_1, \ldots, U_d)$ has for $\boldsymbol{u} \in [0, 1]^d$ with $u_{(1)} \leq \cdots \leq u_{(d)}$ the following distribution function that is a USL copula:

$$C_d(\boldsymbol{u}) = u_{(1)}\mathbb{E}[Q^d] + u_{(1)}\mathbb{E}[(1 - Q)^d] + \sum_{j=2}^{d} \tilde{C}_2(u_{(1)}, u_{(j)})\Big(\mathbb{E}\Big[Q(1 - Q)^{j-1}\Big] + \mathbb{E}\Big[(1 - Q)Q^{j-1}\Big]\Big).$$

- The stochastic conditional sampling representation of the subclass with equal multivariate diagonals uses this subclass's distinct distribution of values onto the random vector's components. An interesting question is if this conditional sampling approach is feasible for other subclasses of semilinear copulas with similarly limited options to distribute minimum and maximum onto the vector's components.

Bibliography

[1] H. Sloot and M. Scherer. “A probabilistic view on semilinear copulas”. In: *Information Sciences* 512 (2020), pp. 258–276. DOI: 10.1016/j.ins.2019.09.069.

[2] H. Sloot. “Implementing Markovian models for extendible Marshall–Olkin distributions”. In: *Dependence Modeling* 10.1 (2022), pp. 308–343. DOI: 10.1515/demo-2022-0151.

[3] M. Scherer and H. Sloot. “Exogenous shock models: analytical characterization and probabilistic construction”. In: *Metrika* 82.8 (2019), pp. 931–959. DOI: 10.1007/s00184-019-00715-8.

[4] H. Sloot. “The deFinetti representation of generalised Marshall–Olkin sequences”. In: *Dependence Modeling* 8.1 (2020), pp. 107–118. DOI: 10.1515/demo-2020-0006.

[5] D. Brigo et al. “Consistent iterated simulation of multivariate defaults: Markov indicators, lack of memory, extreme-value copulas, and the Marshall–Olkin distribution”. In: *Innovations in Insurance, Risk- and Asset Management*. Ed. by K. Glau et al. Conference at the Technical University of Munich, 5–7 April 2017. Singapore: World Scientific, 2018, pp. 47–93. DOI: 10.1142/9789813272569_0003.

[6] A. Sklar. “Fonctions de répartition a n dimensions et leurs marges”. In: *Publications de l’Institut de statistique de l’Université de Paris* 8 (1959), pp. 229–231.

[7] H. Joe. *Multivariate models and multivariate dependence concepts*. New York, NY: Chapman and Hall/CRC, 1997. DOI: 10.1201/b13150.

[8] R. B. Nelsen. *An introduction to copulas*. 2nd ed. New York, NY: Springer, 2006. DOI: 10.1007/0-387-28678-0.

[9] H. Joe. *Dependence modeling with copulas*. New York, NY: Chapman and Hall/CRC, 2014. DOI: 10.1201/b17116.

[10] F. Durante and C. Sempi. *Principles of copula theory*. New York, NY: Chapman and Hall/CRC, 2015. DOI: 10.1201/b18674.

[11] J.-F. Mai and M. Scherer. *Simulating copulas: stochastic models, sampling algorithms and applications*. 2nd ed. Singapore: World Scientific, 2017. DOI: 10.1142/10265.

[12] A. W. Marshall. “Copulas, marginals, and joint distributions”. In: *Distributions with fixed marginals and related topics*. Ed. by L. Rüschendorf, B. Schweizer, and M. D. Taylor. Vol. 28. Lecture Notes–Monograph Series. Hayward, CA: Institute of Mathematical Statistics, 1996, pp. 213–222. DOI: 10.1214/lnms/1215452620.

[13] C. Genest and J. Nešlehová. “A primer on copulas for count data”. In: *Astin Bulletin* 37.2 (2007), pp. 475–515. DOI: 10.2143/AST.37.2.2024077.

[14] D. J. Aldous. “Exchangeability and related topics”. In: *École d’Été de probabilités de Saint-Flour XIII — 1983*. Ed. by P. L. Hennequin. Vol. 1117. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer, 1985, pp. 1–198. DOI: 10.1007/BFb0099421.

[15] J.-F. Mai and M. Scherer. “H-extendible copulas”. In: *Journal of Multivariate Analysis* 110 (2012), pp. 151–160. DOI: 10.1016/j.jmva.2012.03.011.

[16] A. W. Marshall and I. Olkin. “A multivariate exponential distribution”. In: *Journal of the American Statistical Association* 62.317 (1967), pp. 30–44. DOI: 10.2307/2282907.

[17] J.-F. Mai and M. Scherer. “Lévy-frailty copulas”. In: *Journal of Multivariate Analysis* 100.7 (2009), pp. 1567–1585. DOI: 10.1016/j.jmva.2009.01.010.

[18] J.-F. Mai and M. Scherer. “Reparameterizing Marshall–Olkin copulas with applications to sampling”. In: *Journal of Statistical Computation and Simulation* 81.1 (2011), pp. 59–78. DOI: 10.1080/00949650903185961.

[19] J.-F. Mai and M. Scherer. “Sampling exchangeable and hierarchical Marshall-Olkin distributions”. In: *Communications in Statistics. Theory and Methods* 42.4 (2013), pp. 619–632. DOI: 10.1080/03610926.2011.615437.

[20] B. C. Arnold. “A characterization of the exponential distribution by multivariate geometric compounding”. In: *Sankhyā: The Indian Journal of Statistics, Series A* 37.1 (1975), pp. 164–173. ISSN: 0581-572X.

[21] D. Brigo, J.-F. Mai, and M. Scherer. “Markov multi-variate survival indicators for default simulation as a new characterization of the Marshall–Olkin law”. In: *Statistics & Probability Letters* 114 (2016), pp. 60–66. DOI: 10.1016/j.spl.2016.03.013.

[22] X. Li and F. Pellerey. “Generalized Marshall–Olkin distributions and related bivariate aging properties”. In: *Journal of Multivariate Analysis* 102.10 (2011), pp. 1399–1409. DOI: 10.1016/j.jmva.2011.05.006.

[23] J. Lin and X. Li. “Multivariate generalized Marshall–Olkin distributions and copulas”. In: *Methodology and Computing in Applied Probability* 16.1 (2014), pp. 53–78. DOI: 10.1007/s11009-012-9297-4.

[24] P. Muliere and M. Scarsini. “Characterization of a Marshall–Olkin type class of distributions”. In: *Annals of the Institute of Statistical Mathematics* 39.2 (1987), pp. 429–441. DOI: 10.1007/BF02491480.

[25] J.-F. Mai, S. Schenk, and M. Scherer. “Exchangeable exogenous shock models”. In: *Bernoulli* 22.2 (2016), pp. 1278–1299. DOI: 10.3150/14-BEJ693.

[26] J.-F. Mai, S. Schenk, and M. Scherer. “Two novel characterizations of self-decomposability on the half-line”. In: *Journal of Theoretical Probability* 30.1 (2017), pp. 365–383. DOI: 10.1007/s10959-015-0644-6.

[27] F. Durante et al. “Semilinear copulas”. In: *Fuzzy Sets and Systems* 159.1 (2008), pp. 63–76. DOI: 10.1016/j.fss.2007.09.001.

[28] J. J. Arias-García, H. De Meyer, and B. De Baets. “Multivariate upper semilinear copulas”. In: *Information Sciences* 360 (2016), pp. 289–300. DOI: 10.1016/j.ins.2016.04.028.

[29] J. Aczél. *Lectures on functional equations and their applications*. New York, NY: Academic Press, 1966. ISBN: 978-0-12-043750-4.

[30] S. I. Resnick. *Extreme values, regular variation, and point processes*. New York, NY: Springer, 1987. DOI: 10.1007/978-0-387-75953-1.

[31] J. R. Wilmoth et al. *Methods protocol for the Human Mortality Database*. Tech. rep. Available at www.mortality.org [24.01.2023]. Human Mortality Database (HMD). Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France), 2021.

[32] P. Embrechts, R. Frey, and A. McNeil. *Quantitative risk management*. 2nd ed. Princeton, NJ: Princeton University Press, 2015.

[33] L. Rüschendorf. “On the distributional transform, Sklar’s theorem, and the empirical copula process”. In: *Journal of Statistical Planning and Inference* 139.11 (2009), pp. 3921–3927. DOI: 10.1016/j.jspi.2009.05.030.

[34] A. J. McNeil and J. Nešlehová. “Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions”. In: *Annals of Statistics* 37.5B (2009), pp. 3059–3097. DOI: 10.1214/07-AOS556.

[35] A. W. Marshall and I. Olkin. “Families of multivariate distributions”. In: *Journal of the American Statistical Association* 83.403 (1988), pp. 834–841. DOI: 10.1080/01621459.1988.10478671.

[36] J.-F. Mai and M. Scherer. “Characterization of extendible distributions with exponential minima via processes that are infinitely divisible with respect to time”. In: *Extremes* 17.1 (2014), pp. 77–95. DOI: 10.1007/s10687-013-0175-4.

[37] F. Brück, J.-F. Mai, and M. Scherer. “Exchangeable min-id sequences: characterization, exponent measures and non-decreasing id-processes”. In: *Extremes* 26.1 (2023), pp. 183–239. DOI: 10.1007/s10687-022-00450-w.

[38] J. D. Esary and A. W. Marshall. “Multivariate distributions with exponential minimums”. In: *Annals of Statistics* 2.1 (1974), pp. 84–98. DOI: 10.1214/aos/1176342615.

[39] P. Glasserman. *Monte Carlo methods in financial engineering*. New York, NY: Springer, 2003. DOI: 10.1007/978-0-387-21617-1.

[40] M. Larsson and J. Nešlehová. “Extremal behavior of Archimedean copulas”. In: *Advances in Applied Probability* 43.1 (2011), pp. 195–216. DOI: 10.1239/aap/1300198519.

[41] J. F. Sánchez and W. Trutschnig. “Singularity aspects of Archimedean copulas”. In: *Journal of Mathematical Analysis and Applications* 432.1 (2015), pp. 103–113. DOI: 10.1016/j.jmaa.2015.06.036.

[42] G. Bernhart. “Advances in financial engineering: Bondesson densities, the construction of MSMVE distributions, and the modeling of discrete cash dividends”. Dissertation. Munich: Technical University of Munich, 2015. URL: https://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20151019-1253702-1-8.

[43] F. Lindskog and A. J. McNeil. “Common Poisson shock models: applications to insurance and credit risk modelling”. In: *Astin Bulletin* 33.2 (2003), pp. 209–238. DOI: 10.1017/S0515036100013441.

[44] K. Giesecke. “A simple exponential model for dependent defaults”. In: *Journal of Fixed Income* 13.3 (2003), pp. 74–83. DOI: 10.3905/jfi.2003.319362.

[45] Y. Elouerkhaoui. “Pricing and hedging in a dynamic credit model”. In: *International Journal of Theoretical and Applied Finance* 10.4 (2007), pp. 703–731. DOI: 10.1142/S0219024907004408.

[46] T. R. Bielecki et al. “A bottom-up dynamic model of portfolio credit risk; Part I: Markov copula perspective”. Mar. 2013. DOI: 10.2139/ssrn.1844574.

[47] T. R. Bielecki et al. “A bottom-up dynamic model of portfolio credit risk; Part II: common-shock interpretation, calibration and hedging issues”. Mar. 2013. DOI: 10.2139/ssrn.2245130.

[48] J.-F. Mai. “Multivariate exponential distributions with latent factor structure and related topics”. Habilitation thesis. Munich: Technical University of Munich, 2014. URL: https://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20140220-1236170-0-7.

[49] J.-F. Mai. “Extendibility of Marshall–Olkin distributions via Lévy subordinators and an application to portfolio credit risk”. Dissertation. Munich: Technical University of Munich, 2010. URL: https://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20100628-969547-1-8.

[50] P. Sullo. “Estimating the parameters of a multivariate exponential distribution”. PhD thesis. Tallahassee, FL: The Florida State University, 1973.

[51] F. Proschan and P. Sullo. “Estimating the parameters of a multivariate exponential distribution”. In: *Journal of the American Statistical Association* 71.354 (1976), pp. 465–472. DOI: 10.1080/01621459.1976.10480370.

[52] C. Hering and J.-F. Mai. “Moment-based estimation of extendible Marshall–Olkin copulas”. In: *Metrika* 75.5 (2012), pp. 601–620. DOI: 10.1007/s00184-011-0344-x.

[53] H. Li. “Tail dependence comparison of survival Marshall–Olkin copulas”. In: *Methodology and Computing in Applied Probability* 10.1 (2008), pp. 39–54. DOI: 10.1007/s11009-007-9037-3.

[54] J.-F. Mai and M. Scherer. “The Pickands representation of survival Marshall–Olkin copulas”. In: *Statistics & Probability Letters* 80.5-6 (2010), pp. 357–360. DOI: 10.1016/j.spl.2009.11.010.

[55] L. Fernández, J.-F. Mai, and M. Scherer. “The mean of Marshall–Olkin-dependent exponential random variables”. In: *Marshall–Olkin distributions — advances in theory and applications*. Ed. by U. Cherubini, F. Durante, and S. Mulinacci. Vol. 141. Springer Proceedings in Mathematics & Statistics. Cham, Switzerland: Springer, 2015, pp. 33–50. DOI: 10.1007/978-3-319-19039-6_3.

[56] Y. Sun, R. Mendoza-Arriaga, and V. Linetsky. “Marshall–Olkin distributions, subordinators, efficient simulation, and applications to credit risk”. In: *Advances in Applied Probability* 49.2 (2017), pp. 481–514. DOI: 10.1017/apr.2017.10.

[57] C. Berg, J. P. R. Christensen, and P. Ressel. *Harmonic analysis on semigroups*. New York, NY: Springer, 1984. DOI: 10.1007/978-1-4612-1128-0.

[58] A. Gnedin and J. Pitman. “Moments of convex distribution functions and completely alternating sequences”. In: *Institute of Mathematical Statistics collections*. Ed. by D. Nolan and T. Speed. Vol. 2. Institute of Mathematical Statistics, 2008, pp. 30–41. DOI: 10.1214/193940307000000374.

[59] P. Ressel. “Monotonicity properties of multivariate distribution and survival functions — with an application to Lévy-frailty copulas”. In: *Journal of Multivariate Analysis* 102.3 (2011), pp. 393–404. DOI: 10.1016/j.jmva.2010.10.001.

[60] R. L. Schilling, R. Song, and Z. Vondracek. *Bernstein functions*. 2nd ed. Berlin, Boston: De Gruyter, 2012. DOI: 10.1515/9783110269338.

[61] P. Ressel. “Finite exchangeability, Lévy-frailty copulas and higher-order monotonic sequences”. In: *Journal of Theoretical Probability* 26.3 (2013), pp. 666–675. DOI: 10.1007/s10959-011-0389-9.

[62] S. Schenk. “Exchangeable exogenous shock models”. Dissertation. Munich: Technical University of Munich, 2016. URL: https://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20160314-1278138-1-5.

[63] J.-F. Mai. “The infinite extendibility problem for exchangeable real-valued random vectors”. In: *Probability Surveys* 17 (2020), pp. 677–753. DOI: 10.1214/19-PS336.

[64] F. Durante, J. J. Quesada-Molina, and M. Úbeda-Flores. “On a family of multivariate copulas for aggregation processes”. In: *Information Sciences* 177.24 (2007), pp. 5715–5724. DOI: 10.1016/j.ins.2007.07.019.

[65] G. Bernhart et al. “A survey of dynamic representations and generalizations of the Marshall–Olkin distribution”. In: *Marshall–Olkin distributions — advances in theory and applications*. Ed. by U. Cherubini, F. Durante, and S. Mulinacci. Vol. 141. Springer Proceedings in Mathematics & Statistics. Cham, Switzerland: Springer, 2015, pp. 1–13. DOI: 10.1007/978-3-319-19039-6_1.

[66] P. Capéraà, A.-L. Fougères, and C. Genest. “Bivariate distributions with given extreme value attractor”. In: *Journal of Multivariate Analysis* 72.1 (2000), pp. 30–49. DOI: 10.1006/jmva.1999.1845.

[67] A. Charpentier et al. “Multivariate Archimax copulas”. In: *Journal of Multivariate Analysis* 126 (2014), pp. 118–136. DOI: 10.1016/j.jmva.2013.12.013.

[68] S. Mulinacci. “Marshall–Olkin machinery and power mixing: the mixed generalized Marshall–Olkin distribution”. In: *Marshall–Olkin distributions — advances in theory and applications*. Ed. by U. Cherubini, F. Durante, and S. Mulinacci. Vol. 141. Springer Proceedings in Mathematics & Statistics. Cham, Switzerland: Springer, 2015, pp. 65–86. DOI: 10.1007/978-3-319-19039-6_5.

[69] F. Durante, S. Girard, and G. Mazo. “Copulas based on Marshall–Olkin machinery”. In: *Marshall–Olkin distributions — advances in theory and applications*. Ed. by U. Cherubini, F. Durante, and S. Mulinacci. Vol. 141. Springer Proceedings in Mathematics & Statistics. Cham, Switzerland: Springer, 2015, pp. 15–31. DOI: 10.1007/978-3-319-19039-6_2.

[70] A. Gnedin and J. Pitman. “Regenerative composition structures”. In: *Annals of Probability* 33.2 (2005), pp. 445–479. DOI: 10.1214/009117904000000801.

[71] F. Durante et al. “Conjunctors and their residual implicators: characterizations and construction methods”. In: *Mediterranean Journal of Mathematics* 4.3 (2007), pp. 343–356. DOI: 10.1007/s00009-007-0122-1.

[72] B. Schweizer and A. Sklar. “Associative functions and statistical triangle inequalities”. In: *Publicationes Mathematicae Debrecen* 8.1-2 (1961), pp. 169–186. DOI: 10.5486/PMD.1961.8.1-2.16.

[73] B. Schweizer and A. Sklar. “Associative functions and abstract semigroups”. In: *Publicationes Mathematicae Debrecen* 10.1-4 (1963), pp. 69–81. DOI: 10.5486/PMD.1963.10.1-4.09.

[74] F. Durante and C. Sempi. “Copula theory: an introduction”. In: *Copula theory and its applications*. Ed. by P. Jaworski et al. Vol. 198. Lecture Notes in Statistics. Berlin, Heidelberg: Springer, 2010, pp. 3–31. DOI: 10.1007/978-3-642-12465-5_1.

[75] P. Jaworski. “On copulas and their diagonals”. In: *Information Sciences* 179.17 (2009), pp. 2863–2871. DOI: 10.1016/j.ins.2008.09.006.

[76] R. B. Nelsen and G. A. Fredricks. “Diagonal copulas”. In: *Distributions with given marginals and moment problems*. Ed. by V. Beneš and J. Štěpán. Dordrecht, Netherlands: Springer, 1997, pp. 121–128. DOI: 10.1007/978-94-011-5532-8_15.

[77] G. A. Fredricks and R. B. Nelsen. “Copulas constructed from diagonal sections”. In: *Distributions with given marginals and moment problems*. Ed. by V. Beneš and J. Štěpán. Dordrecht, Netherlands: Springer, 1997, pp. 129–136. DOI: 10.1007/978-94-011-5532-8_16.

[78] F. Durante, R. Mesiar, and C. Sempi. “On a family of copulas constructed from the diagonal section”. In: *Soft Computing* 10.6 (2005), pp. 490–494. DOI: 10.1007/s00500-005-0523-7.

[79] F. Durante et al. “Copulas with given diagonal sections: novel constructions and applications”. In: *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 15.4 (2007), pp. 397–410. DOI: 10.1142/s0218488507004753.

[80] F. Durante and P. Jaworski. “Absolutely continuous copulas with given diagonal sections”. In: *Communications in Statistics. Theory and Methods* 37.18 (2008), pp. 2924–2942. DOI: 10.1080/03610920802050927.

[81] R. Mesiar and J. Kalická. “Diagonal copulas”. In: *Aggregation functions in theory and in practise*. Ed. by H. Bustince et al. Vol. 228. Advances in Intelligent Systems and Computing. Berlin, Heidelberg: Springer, 2013, pp. 67–74. DOI: 10.1007/978-3-642-39165-1_10.

[82] J. J. Arias-García, H. De Meyer, and B. De Baets. “Multivariate Bertino copulas”. In: *Journal of Mathematical Analysis and Applications* 434.2 (2016), pp. 1346–1364. DOI: 10.1016/j.jmaa.2015.09.059.

[83] J. Fernández-Sánchez and M. Úbeda-Flores. “Constructions of copulas with given diagonal (and opposite diagonal) sections and some generalizations”. In: *Dependence Modeling* 6.1 (2018), pp. 139–155. DOI: 10.1515/demo-2018-0009.

[84] V. Kachitvichyanukul and B. Schmeiser. “Computer generation of hypergeometric random variates”. In: *Journal of Statistical Computation and Simulation* 22.2 (1985), pp. 127–145. DOI: 10.1080/00949658508810839.

[85] F. Brück. “Exact simulation of continuous max-id processes with applications to exchangeable max-id sequences”. In: *Journal of Multivariate Analysis* 193 (2023), pp. 1–22. DOI: 10.1016/j.jmva.2022.105117.

[86] M. Naaman. “On the tight constant in the multivariate Dvoretzky–Kiefer–Wolfowitz inequality”. In: *Statistics & Probability Letters* 173 (2021), pp. 1–8. DOI: 10.1016/j.spl.2021.109088.

[87] H. Sloot. “Exogenous shock models”. MA thesis. Munich: Technical University of Munich, 2017.

[88] D. Brigo and K. Chourdakis. “Consistent single- and multi-step sampling of multivariate arrival times: a characterization of self-chaining copulas”. Apr. 2012. DOI: 10.2139/ssrn.2047474.

A Core articles as primary author

A.1 A probabilistic view on semilinear copulas

A probabilistic view on semilinear copulas

Matthias Scherer and Henrik Sloot

The article [1] develops novel stochastic representations of upper semilinear copulas. These are exchangeable copulas whose margins are linear on segments with a constant lowest component connecting the diagonal and an upper marginal-boundary hyperplane. Consequently, they are fully characterized by their marginal diagonal functions. This article answers how the correspondence between bivariate upper semilinear copulas and exchangeable exogenous shock models carries over to higher dimensions. Furthermore, it provides stochastic models for several subclasses, which previously had no probabilistic interpretation and were studied as analytical objects only.

The introduction recalls copulas and upper semilinear copulas. We point out that many copulas, including some discussed in this article, were described as analytical functions and have no known stochastic representations. We introduce upper semilinear copulas and preview our new stochastic representations: The first uses reflections from exogenous shock models for sampling the subclass intersecting with survival copulas of exchangeable generalized Marshall–Olkin survival copulas. The second uses conditional sampling for the subclass having equal diagonal functions for all multivariate margins. The third is a de Finetti representation for an extendible subclass.

Sections 2 to 4 recall background on copulas, semilinear copulas, and exchangeable generalized Marshall–Olkin survival copulas, respectively. In particular, we examine a characterization theorem, proving that the d-increasingness property reduces for upper semilinear copula candidate functions to the verification of non-negativity of squares $[a,b]^d$ and rectangles of the form $[a_1,b_1]^{d-m}\times[a_2,b_2]^m$, $b_1\leq a_2$ on the diagonal section. Furthermore, it provides equivalent monotonicity conditions for three univariate functions. Subsequently, we extract from the theorem's proof that upper semilinear copulas' stochastic realizations concentrate almost surely on at most two values. Also, we prove that the independence copula is not upper semilinear from dimension three onwards.

Section 5 investigates upper semilinear copulas' connection to exchangeable generalized Marshall–Olkin copulas and their exogenous shock models. We show that a copula is simultaneously upper semilinear and the copula of an exchangeable generalized Marshall–Olkin distribution if and only if solely shocks corresponding to sets missing at most one element are non-degenerate. This fact implies that at most one component of a realization differs from its joint minimum. Furthermore, we show that this is equivalent to constant incremental differences between marginal diagonal functions.

Section 6 focuses on the case that all diagonal functions of multivariate margins are equal. We prove that this is equivalent to the case where at most one component of a realization differs from its joint maximum. Following, we develop a multi-step simulation algorithm, each step conditioning on the outcomes of previous steps: First, determine in a Bernoulli trial whether the minimum equals the maximum. Second, sample the joint minimum. Third, if the minimum does not equal the maximal, sample the joint maximum and pick the component taking the minimum uniformly at random. We also calculate all involved probabilities and conditional distribution functions.

Section 7 looks at an extendible subclass of upper semilinear copulas. This subclass corresponds to the following stochastic model: First, draw a sample from the bivariate independence copula. Second, conduct independent Bernoulli trials for each component to decide whether it takes the first or second value of the previous step's sample. Optionally, the success probability of these trials can be randomized. We show that the realizations are from an upper semilinear copula and derive its explicit form and that of its marginal diagonal functions. In contrast to the subclasses above, this model allows arbitrary concentrations of a realization's components on its joint minimum and maximum. Finally, we also conclude from the stochastic model that these copulas are radially symmetric.

Statement of individual contribution

I, Henrik Sloot, am the primary author of this article. I am responsible for the conceptualization, the proofs, the software, the analysis, the visualization, and the writing. The role of Matthias Scherer was that of a supervisor, giving feedback on results, the written drafts, related literature, and double-checking results.

Addendum about the notation

To avoid confusion about the notation, I want to highlight that the following article [1] called the survival copulas of exchangeable generalized Marshall–Olkin (exgMO) distributions “exgMO copulas” while this thesis and the preceding summary calls them “survival exgMO copulas”, “exgMO survival copulas”, or “survival copulas of exgMO copulas” for consistency reasons. In particular, note that the copulas in class (A) of the article are called “exgMO copulas” in this thesis.

Erratum

The original publication contained a minor error. The copula in Corollary 6 should read

$$C_d(\boldsymbol{u}) = u_{(1)}\left(\mathbb{E}[Q^d] + \mathbb{E}\left[(1-Q)^d\right] + \sum_{j=2}^{d} u_{(j)}\left(\mathbb{E}\left[Q(1-Q)^{j-1}\right] + \mathbb{E}\left[(1-Q)Q^{j-1}\right]\right)\right).$$

Permission to include

H. Sloot and M. Scherer. “A probabilistic view on semilinear copulas”. In: *Information Sciences* 512 (2020), pp. 258–276. DOI: 10.1016/j.ins.2019.09.069

Reprinted according to the author's rights for publishing with Elsevier;[1] see archived excerpt below.

Copyright

Overview Author rights Institution rights Government rights Find out r

Overview

In order for Elsevier to publish and disseminate research articles, we need certain publishing rights from authors, which are determined by a publishing agreement between the author and Elsevier.

For articles published open access, the authors license exclusive rights in their article to Elsevier where a CC BY-NC-ND end user license is selected, and license non-exclusive rights where a CC BY end user license is selected.

For articles published under the subscription model, the authors typically transfer copyright to Elsevier. In some circumstances, authors may instead grant us (or the learned society for whom we publish) an exclusive license to publish and disseminate their work.

Regardless of whether they choose to publish open access or subscription with Elsevier, authors have many of the same rights under our publishing agreement, which support their need to share, disseminate and maximize the impact of their research.

For open access articles, authors will also have additional rights, depending on the Creative Commons end user license that they select. This Creative Commons license sets out the rights that readers (as well as the authors) have to re-use and

[1] https://www.elsevier.com/about/policies/copyright.

share the article: please see here (https://www.elsevier.com/about/policies/open-access-licenses) for more information on how articles can be re-used and shared under these licenses.

This page aims to summarise authors' rights when publishing with Elsevier; these are explained in more detail in the publishing agreement between the author and Elsevier.

Irrespective of how an article is published, Elsevier is committed to protect and defend authors' works and their reputation. We take allegations of infringement, plagiarism, ethical disputes, and fraud very seriously.

Author rights

The below table explains the rights that authors have when they publish with Elsevier, for authors who choose to publish either open access or subscription. These apply to the corresponding author and all co-authors.

Author rights in Elsevier's proprietary journals	**Published open access**	**Published subscription**
Retain patent and trademark rights	√	√
Retain the rights to use their research data freely without any restriction	√	√
Receive proper attribution and credit for their published work	√	√
Re-use their own material in new works without permission or payment (with full acknowledgement of the original article): 1. Extend an article to book length 2. Include an article in a subsequent compilation of their own work 3. Re-use portions, excerpts, and their own figures or tables in other works.	√	√

Use and share their works for scholarly purposes (with full acknowledgement of the original article): 1. In their own classroom teaching. Electronic and physical distribution of copies is permitted 2. If an author is speaking at a conference, they can present the article and distribute copies to the attendees 3. Distribute the article, including by email, to their students and to research colleagues who they know for their personal use 4. Share and publicize the article via Share Links, which offers 50 days' free access for anyone, without signup or registration 5. Include in a thesis or dissertation (provided this is not published commercially) 6. Share copies of their article privately as part of an invitation-only work group on commercial sites with which the publisher has a hosting agreement	√	√
Publicly share the preprint on any website or repository at any time.	√	√
Publicly share the accepted manuscript on non-commercial sites	√	√ using a CC BY-NC-ND license and usually only after an embargo period (see Sharing Policy (https://www.elsevier.com/about/policies/sharing) for more information)
Publicly share the final published article	√ in line with the author's choice of end user license	×
Retain copyright	√	×

Institution rights

Regardless of how the author chooses to publish with Elsevier, their institution has the right to use articles for classroom teaching and internal training. Articles can be used for these purposes throughout the author's institution, not just by the author:

Institution rights in Elsevier's proprietary journals (providing full acknowledgement of the original article is given)	**All articles**
Copies can be distributed electronically as well as in physical form for classroom teaching and internal training purposes	√
Material can be included in coursework and courseware programs for use within the institution (but not in Massive Open Online Courses)	√
Articles can be included in applications for grant funding	√
Theses and dissertations which contain embedded final published articles as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publication on ScienceDirect	√

Government rights

For US government employees, works created within the scope of their employment are considered to be public domain and Elsevier's publishing agreements do not require a transfer or license of rights for such works.

In the UK and certain commonwealth countries, a work created by a government employee is copyrightable, but the government may own the copyright (Crown copyright). Click here (https://www.elsevier.com/open-access/agreements/uk-national-archives) for information about UK government employees publishing open access.

Information Sciences 512 (2020) 258–276

Contents lists available at ScienceDirect

Information Sciences

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

A probabilistic view on semilinear copulas

Henrik Sloot, Matthias Scherer*

Technische Universität München, Lehrstuhl für Finanzmathematik, Parkring 11, Garching 85748, Germany

ARTICLE INFO

Article history:
Received 15 March 2019
Revised 23 September 2019
Accepted 26 September 2019
Available online 27 September 2019

Keywords:
Copula
Diagonal section
Semilinear copula
Exogenous shock model
Marshall–Olkin copula

ABSTRACT

This article advances the theory on multivariate upper semilinear copulas. Probabilistic features of this class are discussed and three subclasses are investigated in detail. The first subclass consists of upper semilinear copulas whose survival copulas are generalised Marshall–Olkin copulas. The second subclass is defined in that they possess identical multivariate diagonals. The third subclass is a family of extendible upper semilinear copulas. Stochastic models and analytical characterisation theorems are derived for each of these subclasses.

1. Introduction

Sklar's theorem, see [26], allows the decomposition of a multivariate distribution function F into a copula C and marginal distribution functions $F_1, \dots, F_d$ via:

$$F(\boldsymbol{x}) = C\big(F_1(x_1), \dots, F_d(x_d)\big), \quad \boldsymbol{x} = (x_1, \dots, x_d) \in \mathbb{R}^d. \tag{1}$$

The copula C is itself a multivariate distribution function whose marginal distribution functions are standardised to the uniform distribution function on [0, 1].

The analytical decomposition in Eq. (1) is useful for stochastic modelling as well as statistical inference. As a result, copulas have been intensively studied over the recent decades in analysis, probability theory, and statistics, see [6,11,16,19,25]. Furthermore, they are used in practical applications, e.g. in quantitative risk management, credit risk, and insurance mathematics, see [7,9].

Copulas can be equivalently characterised by analytical properties involving the notion of d-monotonicity, groundedness, and the uniform margin property, see Eqs. (3a)–(3c). Consequently, in spite of their probabilistic nature, they are also studied as purely analytical objects in other mathematical fields, e.g. in fuzzy set theory. Two examples thereof are illustrated by the following:

- Bivariate copulas have uniform margins and are 2-increasing. In particular, the latter property implies that they are non-decreasing in each component. Therefore, they belong to the classes of *conjunctors* and *semicopulas* (sometimes also called *t-seminorms*).[1]

* Corresponding author.
E-mail addresses: henrik.sloot@tum.de (H. Sloot), scherer@tum.de (M. Scherer).

[1] A conjunctor is a monotone non-decreasing extension of the boolean conjunction from {0, 1} to the interval [0, 1] and a semicopula is a conjunction with neutral element 1, see [3].

https://doi.org/10.1016/j.ins.2019.09.069
0020-0255/

- Copulas of higher dimensions also possess the same aforementioned properties. Particularly, this implies that the distribution functions of countable infinite exchangeable sequences of random variables with uniform margins form a subclass of *conjuctive aggregation operations*, see [5].

As copulas can be treated and analysed as purely analytical objects, there are various families of copulas that have not historically originated from a stochastic model, but have been identified by their analytical properties from a broader class of functions. It is usually a difficult undertaking to link the analytical properties of a copula to a stochastic model. However, a stochastic representation is required for applications, e.g. for simulation and model construction. Additionally, stochastic representations can yield valuable analytical insights as, for example, analytical characterisations or proof of certain characteristics. A well-known example is the class of Archimedean copulas, see [23,24]. These originally emerged in the field of probabilistic metric spaces as a subclass of so-called *t-norms*. A complete characterisation of this class has been presented only recently in [18] by means of a unified stochastic representation. Another example is presented in Section 7, in which radial symmetry follows in a natural way from the stochastic model whereas an analytical proof would require tedious calculations.

In this article, we discuss d-variate upper semilinear copulas (abbr. as USL_d for $d \geq 2$). These were introduced in [2] as the solution to the following compatibility problem: given copula diagonals δ_j, $j \in \{2, \ldots, d\}$, what are the necessary and sufficient conditions for these to be the j-dimensional marginal diagonals of an exchangeable d-variate copula C that is linear on the segments

$$\mathcal{S}_{\boldsymbol{v},d} := \left\{ \begin{pmatrix} \boldsymbol{u} \\ \mathbf{1} \end{pmatrix} \in [0,1]^{j+(d-j)} \;:\; u_i = \lambda v_1 + (1-\lambda) v_i,\ i \in \{1, \ldots, j\},\ \lambda \in [0,1] \right\}, \tag{2}$$

where $j \in \{1, \ldots, d\}$ and $\boldsymbol{v} \in [0,1]^j$ with $v_1 \leq \ldots \leq v_j = 1$? These copulas are an extension of *bivariate* upper semilinear copulas, which were proposed in [4]. [2] describes the following recursive construction for C. In this case, denote for $k \in \{1, \ldots, d\}$ the k-margins of C by C_k and denote for $\boldsymbol{u} = (u_1, \ldots, u_k) \in [0,1]^k$ its ordered version by $u_{(1)} \leq \ldots \leq u_{(k)}$. Furthermore, represent $\boldsymbol{u}$ as the linear combination[2]

$$\boldsymbol{u} = \lambda (u_{(1)} \cdot \mathbf{1}) + (1-\lambda) \boldsymbol{u}^\star,$$

where $\lambda = (1 - u_{(k)})/(1 - u_{(1)})$ and

$$u^\star_{(j)} := u_{(1)} + \frac{1}{1-\lambda} (u_{(j)} - u_{(1)}), \quad j \in \{1, \ldots, k\}.$$

In particular, $u_{(1)} \cdot \mathbf{1}$ is a value on the (marginal) diagonal and $\boldsymbol{u}^\star$ is a value on the (marginal) boundary. Consequently, the fact that C is linear on the segments in Eq. (2) implies for $k \in \{2, \ldots, d\}$ the following recursion

$$\begin{aligned} C_k(\boldsymbol{u}) &= \lambda \cdot C_k(u_{(1)} \cdot \mathbf{1}) + (1-\lambda) \cdot C_k(\boldsymbol{u}^\star) \\ &= \lambda \cdot \delta_k(u_{(1)}) + (1-\lambda) \cdot C_{k-1}(u^\star_1, \ldots, u^\star_{k-1}). \end{aligned}$$

To the best of our knowledge, the class of multivariate upper semilinear copulas has only been investigated analytically for $d > 2$. Furthermore, apart from the special case of comonotonicity, we are not aware of any stochastic representation for a multivariate upper semilinear copula with $d > 2$. We intend to fill this gap by investigating the subclasses (A), (B), and (C) introduced below.[3] For all families, we provide a stochastic representation and a characterisation theorem.

(A) One subclass of upper semilinear copulas can be linked to so-called *exchangeable exogenous shock models*. Note that the survival copula $\hat{C}_2$ of a bivariate upper semilinear copula C_2 is called a *lower semilinear copula*, see [4]. A bivariate lower semilinear copula has the form

$$\hat{C}_2(\boldsymbol{u}) = u_{(1)} \cdot \frac{\delta_2^L(u_{(2)})}{u_{(2)}}, \quad \boldsymbol{u} \in [0,1]^2.$$

These copulas are also known as *bivariate exchangeable generalised Marshall–Olkin copulas* (exGMO_2), see [12,13]. Furthermore, they have a multivariate extension with a stochastic representation which is called the exchangeable exogenous shock model, see [15]. The question that arises is: *What is the intersection between the classes survival exGMO and USL for $d > 2$?* In Section 5, we provide an answer to this question by deriving the necessary and sufficient conditions for a copula to be in the intersection of both of these classes.

(B) One example, which is presented and discussed in [2], is that of identical multivariate diagonals. We show that a realisation $\boldsymbol{U} \sim C$ of an upper semilinear copula C with identical multivariate diagonals can have at most one component that differs from the joint maximum. This implies that the ordered version of $\boldsymbol{U}$ is determined by the minimum component $U_\wedge$, the maximum component $U_\vee$, and the event $\{U_\wedge \neq U_\vee\}$. In Section 6, we use this observation to derive a stochastic model, which is based on conditional sampling. We also provide a novel characterisation theorem for this subclass.

[2] Note, that we have made a small correction to the representation of $\boldsymbol{u}^\star$ compared to the original reference [2].

[3] For the purpose of readability, we refer to these three subclasses in the remainder of this introduction as (A), (B), and (C).

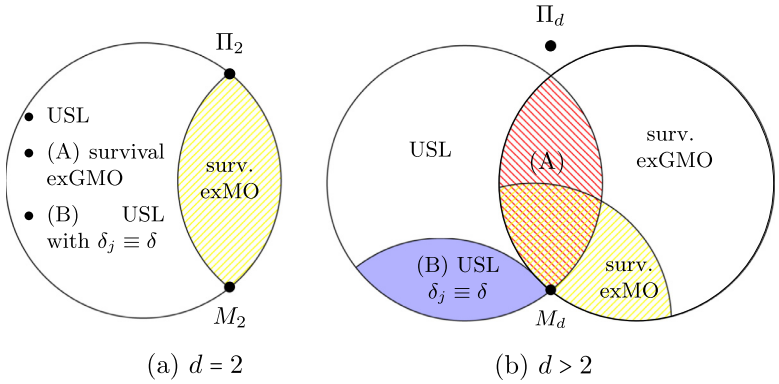


Fig. 1. A Venn diagram of the subclasses of upper semilinear copulas and survival exchangeable generalised Marshall–Olkin copulas. For the sake of completeness, the class of survival exchangeable Marshall–Olkin copulas is included. M_d and Π_d are the Fréchet–Hoeffding upper bound and the independence copula for dimension d, respectively.

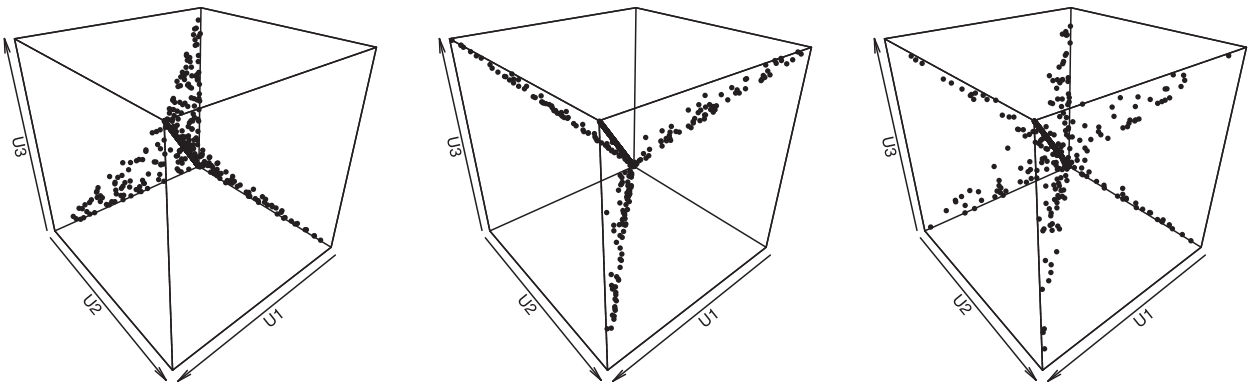


Fig. 2. Three scatterplots for 3-margins of 5-dimensional realisations of members of the subclasses (A), (B), and (C) (left to right). All three copulas were calibrated to the same multivariate lower tail-dependence parameter.

(C) In Section 7, we present a subclass which is extendible in the class of upper semilinear copulas[4] and provide its explicit deFinetti representation. A deFinetti representation is a two-step model where in the first step a random distribution function is sampled, from which — in the second step — an iid sample is drawn, see [1, Chapters 2 and 3]. This particular subclass extends the bivariate Dirichlet copula, see [22, Theorem 3.5.3], and is radially symmetric. Furthermore, members of this subclass are conjunctive aggregation operations.

The extension of the first two subclasses to higher dimensions is illustrated in Fig. 1.[5] We observe the following: in the bivariate case, both subclasses (A) and (B) coincide with the entire class of upper semilinear copulas. In the case $d > 2$, the only copula in the intersection of the subclasses (A) and (B) is the comonotonicity copula.[6] Furthermore, the independence copula is not upper semilinear if $d > 2$. Fig. 2 shows exemplary scatterplots for all three subclasses, (A), (B), and (C).

The remaining paper is organised as follows: we briefly introduce the necessary key concepts from copula theory in Section 2 as well as the classes of upper semilinear copulas and exchangeable generalised Marshall–Olkin copulas in Section 3 and Section 4, respectively. In Sections 5, 6, 7, we discuss the subclasses (A), (B), and (C), while we offer a conclusion in Section 8.

2. Notation and mathematical background

In this section, we introduce copulas from an analytical and probabilistic point of view, summarise relevant results, and establish the notation which is used throughout this article.

[4] We call a copula $C_d \in \mathrm{USL}_d$ extendible in the class of upper semilinear copulas if a sequence of random variables $\{U_i\}_{i\in\mathbb{N}}$ exists such that each finite margin is upper semilinear and the d-margin is C_d.

[5] The illustration in Fig. 1 anticipates a few minor results which will be proven in the later sections.

[6] This statement can be extended as follows: in the case $d > 2$, the only copula in at least two of the subclasses (A), (B), or (C) is the comonotonicity copula.

We use the following conventions for a concise notation. We use bold letters for vectors and capital letters for random variables and vectors. We apply operators component-wise, i.e. $\boldsymbol{U} \leq \boldsymbol{u}$ means $U_i \leq u_i$ for all i. For a vector $\boldsymbol{u} \in [0,1]^d$, we denote its ordered version by $u_{(1)} \leq \ldots \leq u_{(d)}$. For a (multivariate) distribution function F, we use the notation $\boldsymbol{X} \sim F$ if the random vector $\boldsymbol{X}$ has the distribution function F. Furthermore, for two random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$, we denote $\boldsymbol{X} \overset{d}{=} \boldsymbol{Y}$ if $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same distribution function. If not stated otherwise, $d \in \mathbb{N}$ denotes the dimension. Finally, we define $[n] := \{1, \ldots, n\}$ for $n \in \mathbb{N}$.

We use mathematical expressions involving vectors $\boldsymbol{u} \in [0,1]^k$, resp. their ordered versions $0 \leq u_{(1)} \leq \ldots \leq u_{(k)} \leq 1$, and diagonal functions which may contain fractions having a numerator and denominator equal to zero for some specific values on the boundary of $[0,1]^k$. For readability, we omit the treatment of these special cases in the following as the existence of an analytical continuation can, in all involved cases, easily be inferred from the properties of a diagonal function and/or the fact that $0 \leq u_{(1)} \leq \ldots \leq u_{(k)}$.

2.1. Copulas, symmetry, and diagonal functions

Below, we summarise the most important definitions and results from the theory on copulas.

We call a function $C\colon [0,1]^d \to [0,1]$ *copula* if it fulfills the following conditions.

$$C(\boldsymbol{u}) = 0 \quad \text{for all } \boldsymbol{u} \in [0,1]^d \text{ with } u_i = 0 \text{ for some } i \in [d]. \tag{3a}$$

$$C(\boldsymbol{u}) = u_i \quad \text{for all } \boldsymbol{u} \in [0,1]^d \text{ with } u_j = 1\ \forall j \neq i. \tag{3b}$$

$$V_C([\boldsymbol{a}, \boldsymbol{b}]) := \sum_{\boldsymbol{\gamma} \in \times_{i=1}^d \{a_i, b_i\}} (-1)^{|\{i\,:\, a_i = \gamma_i\}|}\, C(\boldsymbol{\gamma}) \geq 0 \quad \text{for all } [\boldsymbol{a}, \boldsymbol{b}] \subseteq [0,1]^d. \tag{3c}$$

Furthermore, we call $V_C([\boldsymbol{a}, \boldsymbol{b}])$ the C-volume of the rectangle $[\boldsymbol{a}, \boldsymbol{b}]$, $\boldsymbol{a} \leq \boldsymbol{b}$, and we call the properties in Eqs. (3a)–(3c) *groundedness, uniform margin property*, and *d-increasingness*.

Given a set of functions from $[0,1]^d$ to $[0,1]$, an interesting problem is to determine the copula subclass, i.e. the functions that fulfill Eqs. (3a)–(3c). To emphasise that we often discuss *potential* copulas, we call arbitrary functions $C\colon [0,1]^d \to [0,1]$ *copula candidate functions*. There is a second equivalent definition for copulas that characterises copulas as probabilistic objects.

Lemma 1 [20, Theorem 8]**.** *A function $C\colon [0,1]^d \to [0,1]$ is a copula if and only if a random vector $\boldsymbol{U}$ on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ exists with*

$$\mathbb{P}(U_i \leq u) = u, \quad \forall u \in [0,1],\ i \in [d],$$
$$\mathbb{P}(\boldsymbol{U} \leq \boldsymbol{u}) = C(\boldsymbol{u}), \quad \forall \boldsymbol{u} \in [0,1]^d.$$

Hence, we can identify a copula with a probability measure on $[0,1]^d$.

We define the *survival copula* $\hat{C}$ of a copula C by

$$\hat{C}(\boldsymbol{u}) := V_C([\mathbf{1} - \boldsymbol{u}, \mathbf{1}]), \quad \boldsymbol{u} \in [0,1]^d.$$

A simple calculation shows that if $\boldsymbol{U} \sim C$, we have $\mathbf{1} - \boldsymbol{U} \sim \hat{C}$. In particular, this implies that $\hat{C}$ is itself a copula. Furthermore, note that $\hat{\hat{C}} = C$. We call a copula *radially symmetric* if $C = \hat{C}$. In the language of probability theory, this is equivalent to $\boldsymbol{U} \overset{d}{=} \mathbf{1} - \boldsymbol{U}$ for $\boldsymbol{U} \sim C$.

We call a copula (candidate function) C *exchangeable* if

$$C(\boldsymbol{u}) = C(u_{\pi(1)}, \ldots, u_{\pi(d)}), \quad \forall \boldsymbol{u} \in [0,1]^d$$

and for all permutations π on the index set $[d]$. For a random vector $\boldsymbol{U} \sim C$, the corresponding probabilistic interpretation is that a permutation of the components of $\boldsymbol{U}$ does not change its distribution function.

We call a copula C *extendible (in a class $\mathcal{C}$)* if an exchangeable sequence $\{U_i\}_{i \in \mathbb{N}}$ exists such that $(U_1, \ldots, U_d)' \sim C$ (and each finite margin is from the class $\mathcal{C}$).[7] Note that this implies that an extendible copula is always exchangeable. The converse, however, is not true as one can construct simple examples of exchangeable copulas that are not extendible (e.g. the bivariate counter-monotonicity copula, see [16, Remark 1.4]).

For an exchangeable copula (candidate function) C and $k \in [d]$, we define the *k-margin* C_k and the *k-diagonal* δ_k of C by

$$C_k : [0,1]^k \to [0,1], \quad \boldsymbol{u} \mapsto C(\boldsymbol{u}, 1, \ldots, 1),$$
$$\delta_k : [0,1] \to [0,1], \quad u \mapsto C_k(u, \ldots, u).$$

[7] Note that our definition of extendibility *in a class $\mathcal{C}$* is stricter than the common definition of extendibility, see, e.g. [16, p. 43], where it is usually not required that each finite margin is in the prespecified class $\mathcal{C}$.

In a probabilistic setting, C_k is the distribution function of a k-dimensional marginal vector, e.g. $(U_1, \ldots, U_k)$, and δ_k is the distribution of the minimum of k components, e.g. $\min_{i\in[k]} U_i$.

Given the k-margin of a copula candidate function C, one can use the copula definition or Lemma 1 to check if C_k is a proper copula. Similarly, there are conditions to verify whether a function $\delta\colon [0,1]\to[0,1]$ is the diagonal of a d-dimensional copula. We call such a function a *d-diagonal*. Note that we can determine if a function δ is a d-diagonal with the following lemma.

Lemma 2 [10] or [21]**.** *A function $\delta\colon [0,1]\to[0,1]$ is a d-diagonal, i.e. the diagonal of a d-dimensional copula, if and only if the following conditions are fulfilled:*

$$\delta(1) = 1. \tag{4a}$$

$$\delta(u) \le u, \quad \forall u \in [0,1]. \tag{4b}$$

$$0 \le \delta(v) - \delta(u) \le d(v-u), \quad \forall u, v \in [0,1] \text{ with } u < v. \tag{4c}$$

3. Semilinear copulas

In this section, we introduce the class of upper semilinear copulas. Traditionally, this class is not defined by a stochastic model, but by an analytic, recursive construction principle. This recursive construction allows the specification of an upper semilinear copula solely by its diagonal functions. A concise overview of various approaches to specifying a copula via given diagonal functions can be found in [2, Section 1].

We placed a special emphasis on the characteristics and peculiarities of this class. One noteworthy property is that a realisation from such a copula has at most two distinct components. With the following sections of this article in mind, this property is crucial for the derivation of stochastic models and characterisation theorems.

Definition 1. We call a d-variate copula candidate function $C\colon [0,1]^d\to[0,1]$ *(upper) semilinear* if C is *exchangeable* and if C is linear on the sections

$$\mathcal{S}_{v,d} := \left\{ \begin{pmatrix} \boldsymbol{u} \\ \mathbf{1} \end{pmatrix} \in [0,1]^{j+(d-j)} \;:\; \boldsymbol{u} = \lambda \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_1 \end{pmatrix} + (1-\lambda)\cdot \begin{pmatrix} v_1 \\ \vdots \\ v_j \end{pmatrix},\ \lambda \in [0,1] \right\},$$

where $j \in \{1, \ldots, d\}$ and $\boldsymbol{v} \in [0,1]^j$ with $v_1 \le \ldots \le v_j = 1$.

This property has three important implications, which can be derived by straightforward calculations, see [2]. For this, let $\boldsymbol{u} = (u_1, \ldots, u_k)' \in [0,1]^k$ with ordered version $u_{(1)} \le \ldots \le u_{(k)}$.

(a) The k-margins C_k of C are upper semilinear copulas (resp. copula candidate functions).

(b) The following recursion holds:

$$C_k(\boldsymbol{u}) = \frac{1-u_{(k)}}{1-u_{(1)}}\cdot\delta_k(u_{(1)}) + \frac{u_{(k)}-u_{(1)}}{1-u_{(1)}}\cdot C_{k-1}(\boldsymbol{u}^\star), \tag{5a}$$

where $\boldsymbol{u}^\star \in [0,1]^{k-1}$ with

$$\boldsymbol{u}^\star_{(i)} := u_{(1)} + \frac{(u_{(i)}-u_{(1)})\cdot(1-u_{(1)})}{u_{(k)}-u_{(1)}},\ i \in \{1,\ldots,k-1\}. \tag{5b}$$

(c) Using the convention that $u_{(k+1)} = 1$, we can write the k-margin C_k as

$$C_k(\boldsymbol{u}) = \frac{\sum_{i=1}^{k}(u_{(i+1)}-u_{(i)})\cdot\delta_i(u_{(1)})}{1-u_{(1)}}. \tag{6}$$

[2] provides the following characterisation theorem for determining if an upper semilinear copula candidate function C is a proper copula.

Theorem 1 (Characterisation, see [2, Theorem 1])**.** *Let C be an upper semilinear copula candidate function such that the corresponding diagonal sections $\delta_2, \ldots, \delta_d$ are proper diagonals. C is a copula if and only if the following three conditions hold:*

(a) *For any $m \in \{1, \ldots, d-1\}$, the function $v_d^{(m)} : [0,1) \to [0,\infty)$, defined by*

$$v_d^{(m)}(u) := \frac{1}{1-u}\cdot\sum_{j=0}^{m}(-1)^j\binom{m}{j}\delta_{d-m+j}(u),$$

is non-decreasing. The C-volume of $[0,u]^{d-m}\times[u,1]^m$ is $(1-u)\cdot v_d^{(m)}(u)$.

(b) *The function* ζ_d: $[0, 1] \to [0, 1]$, *defined by*

$$\zeta_d(u) := 1 + \sum_{j=1}^{d} (-1)^j \binom{d}{j} \delta_j(u),$$

is non-increasing. The C-volume of $[u, 1]^d$ *is* $\zeta_d(u)$.

(c) *The inequality*

$$\left(\frac{\delta_d(u)}{1-u}\right)' \geq \frac{1-\zeta_d(u)}{(1-u)^2} \tag{7}$$

holds almost everywhere with respect to the Lebesgue measure on $(0, 1)$.

The proof of this theorem, see [2, p. 293–296], shows that the three conditions from Theorem 1 are equivalent to C being d-increasing.[8] This proof utilises some properties of semilinear copulas (resp. copula candidate functions) which we will use in subsequent sections. Therefore, we outline the proof and highlight important intermediate results. For the complete proof, we refer the interested reader to the aforementioned reference.

In the proof, it is first observed that each *d-box* $[\boldsymbol{a}, \boldsymbol{b}] \subseteq [0, 1]^d$ can be decomposed into d-boxes, with disjoint interiors, of the form

$$\bigtimes_{i=1}^{d} [u_i, v_i],$$

such that for all $i \neq j$ either $u_j \geq v_i$, $u_i \geq v_j$, or $[u_i, v_i] = [u_j, v_j]$. Thus, due to the exchangeability of C, it suffices to check the d-increasingness property on d-boxes of the form

$$\bigtimes_{j=1}^{r} [u_j, v_j]^{m_j},$$

with $m_1 + \ldots + m_r = d$ and

$$0 \leq u_1 < v_1 \leq u_2 < v_2 \leq \ldots \leq u_r < v_r \leq 1.$$

The remaining proof is split into three parts for the cases $r = 1$, $r = 2$, and $r > 2$ and the following is shown: the d-increasingness property is equivalent to condition a) or conditions b) and c) for $r = 1$ or $r = 2$, respectively. In the case $r > 2$, the C-volume of the d-box is equal to zero. The last statement is summarised in the following lemma.

Lemma 3 [2, Proof of Theorem 1]**.** *Let* C *be an upper semilinear copula candidate function and* $r > 2$, $\sum_{j=1}^{r} m_j = d$, *and* $u_i > v_j \; \forall i > j$. *Then*

$$V_C\left(\bigtimes_{j=1}^{r} [u_j, v_j]^{m_j}\right) = 0. \tag{8}$$

Lemma 3 highlights a significant characteristic of upper semilinear copulas, since Eq. (8) implies that a realisation $\boldsymbol{U} \sim C$ must be concentrated on at most two distinct (random) values. More formally, let C be an upper semilinear copula and $\boldsymbol{U} \sim C$. Then we can conclude with a simple probabilistic argument that

$$\mathbb{P}\big(U_1 \neq U_2, U_1 \neq U_3, U_2 \neq U_3\big) = 0.$$

This implies a very strong, albeit unusual, dependence structure between the components of $\boldsymbol{U}$.

Another important corollary from the proof of Theorem 1 is the following collection of closed-form expressions for C-volumes.

Corollary 1 [2, Proof of Theorem 1]**.** *Let* C *be an upper semilinear copula. Then for* $0 \leq u < v \leq 1$ *and* $0 \leq u_1 < v_1 \leq u_2 < v_2 \leq 1$, *we have*

$$\begin{aligned}
V_C([u, v]^d) &= \delta_d(v) - \left(\frac{1-v}{1-u} \cdot \delta_d(u) + \frac{v-u}{1-u} \cdot (1 - \zeta_d(u))\right), \\
V_C\big([u_1, v_1]^{d-m} \times [u_2, v_2]^m\big) &= (v_2 - u_2) \cdot \big(\nu_d^{(m)}(v_1) - \nu_d^{(m)}(u_1)\big), \\
V_C\big([0, v]^d \setminus ([0, u]^d \cup [u, v]^d)\big) &= \frac{v-u}{1-u} \cdot (1 - \zeta_d(u) - \delta_d(u)) \\
&= \frac{v-u}{1-u} \cdot V_C\big([0, 1]^d \setminus ([0, u]^d \cup [u, 1]^d)\big).
\end{aligned}$$

[8] It follows directly from Eq. (6) that C is also grounded and possesses the uniform margin property.

Proof. For the first identity, see [2, p. 296], and for the second identity, see [2, p. 294]. We obtain the third identity with the following calculation:

$$\begin{aligned} V_C\big([0,v]^d \setminus ([0,u]^d \cup [u,v]^d)\big) &= V_C\big([0,v]^d\big) - V_C([0,u]^d) - V_C\big([u,v]^d\big) \\ &= \delta_d(v) - \delta_d(u) - \delta_d(v) + \left(\frac{1-v}{1-u}\cdot\delta_d(u) + \frac{v-u}{1-u}\cdot(1-\zeta_d(u))\right) \\ &= \frac{v-u}{1-u}\cdot(1-\zeta_d(u)-\delta_d(u)) \\ &= \frac{v-u}{1-u}\cdot V_C\big([0,1]^d \setminus ([0,u]^d) \cup [u,1]^d\big). \end{aligned}$$

□

We conclude this section with an interesting finding: while the independence copula Π_d is upper semilinear for $d = 2$, we show in the following corollary that this is not the case for $d > 2$.

Corollary 2. *Let $d > 2$ and C be an upper semilinear copula candidate function with $\delta_j(u) = u^j$, $j \geq 2$. Then C is not a proper copula.*

Proof. We check this claim by contradiction. Let C be an upper semilinear copula candidate function with $d > 2$ and diagonal functions $\delta_j(u) = u^j$, $j \geq 2$. Assume that C is a proper copula. This implies that the 3-margin C_3 is also a proper copula. A simple calculation shows that $\zeta_3(u) = (1-u)^3$. Hence, we find that the third condition of the characterisation theorem, $(\delta_3(u)/(1-u))' \geq (1-\zeta_3(u))/(1-u)^2$ for almost every $u \in [0,1]$, is equivalent to

$$3u^2(1-u) + u^3 \geq 1 - (1-u)^3.$$

This implies for $u = 1/2$ that $4/8 \geq 1 - 1/8$ or $5/8 \geq 1$, which is a contradiction. □

4. Exchangeable generalised Marshall–Olkin copulas

Upper semilinear copulas are connected to exchangeable Marshall–Olkin distributions. These were introduced in the seminal paper [17], which also showed that they are uniquely linked to so-called exogenous shock models with independent, exponentially distributed *shocks*. The aforementioned connection emerges in the bivariate case, where the survival copula of a bivariate upper semilinear copula is an exchangeable generalised Marshall–Olkin copula and vice versa. One of the initial questions leading to this article was under which circumstances this relationship holds in higher dimensions. To answer this question, we briefly discuss exchangeable generalised Marshall–Olkin distributions and the exogenous shock model.

Definition 2 (Exchangeable generalised Marshall–Olkin copula)**.** We call a d-variate copula candidate function $C\colon [0,1]^d \to [0,1]$ *exchangeable* and of *generalised Marshall–Olkin type* if there are functions $g_2, \ldots, g_d$ such that

$$C(\boldsymbol{u}) = u_{(1)} \cdot g_2(u_{(2)}) \cdot \ldots \cdot g_d(u_{(d)}), \quad \boldsymbol{u} = (u_1, \ldots, u_d)' \in [0,1]^d. \tag{9}$$

Furthermore, we call a proper copula of that form *exchangeable generalised Marshall–Olkin* copula (abbr. as exGMO_d).

An extensive monograph on these copulas is [22]; [15] is a concise article, containing all results presented in this section.

Classical exchangeable Marshall–Olkin copulas arise as a special case if the functions g_i, $i \geq 2$, are power functions. They obtain their name from the eponymous multivariate exponential distribution which was proposed in [17] and whose survival copulas are of this form. An extensive monograph on exchangeable Marshall–Olkin copulas is [14].

[22] provides three equivalent characterising conditions for an exchangeable copula candidate function of generalised Marshall–Olkin type to be a proper copula. However, for our purposes, we only require the one presented below. For this, let $\mathcal{D}$ be the set of continuous distribution functions on $[0,1]$ which are positive on $(0,1]$, i.e.

$$\mathcal{D} := \left\{ F \in \mathcal{C}^{(0)}([0,1]) \;:\; \Delta F \geq 0,\ 0 \notin F((0,1]),\ F(1) = 1 \right\}.$$

Theorem 2 (Characterisation, see [22, Theorem 3.3.1])**.** *Let C be an exchangeable copula candidate function of Marshall–Olkin type having a representation as in Eq. (9) for functions $g_2, \ldots, g_d$ with $g_i(1) = 1$, $i \geq 2$. Then C is a proper copula if and only if $H_i \in \mathcal{D}$ for all $i \in \{1, \ldots, d\}$, where*

$$H_i(u) := \begin{cases} \prod_{j=0}^{i-1} \left(g_{d-i+1+j}(u)\right)^{(-1)^j \binom{i-1}{j}} & u \in (0,1] \\ \lim_{v \searrow 0} H_i(v) & u = 0. \end{cases} \tag{10}$$

The functions H_i in Eq. (10) can be used to define a stochastic representation called *exchangeable exogenous shock model*, see [22, p. 61 sqq.]. For this purpose, consider a proper copula $C \in \mathrm{exGMO}_d$ with a representation as in Eq. (9) for functions $g_2, \ldots, g_d$. Let $\{Z_I\}_{\emptyset \neq I \subseteq \{1,\ldots,d\}}$ be a family of independent random variables with

$$Z_I \sim H_{|I|}, \qquad \emptyset \neq I \subseteq \{1, \ldots, d\}.$$

Then it is the case that $\boldsymbol{U} \sim C$, where $\boldsymbol{U}$ is defined by

$$U_i = \max\left\{Z_I \ : \ i \in I\right\}, \qquad i \in \{1, \dots, d\}. \tag{11}$$

Note that the prefix *exchangeable* highlights that the resulting vector $\boldsymbol{U}$ is exchangeable. Furthermore, it can be shown that $\boldsymbol{U}$ is exchangeable if and only if the shock distribution functions H_I depend only on the cardinality of the corresponding set I, see [22, Propostion 3.1.2].

The following lemma shows that every exchangeable exogenous shock model also implies an exGMO copula. For this reason, $C \sim \text{exGMO}_d$ is sometimes called *exchangeable exogenous shock model copula*.

Lemma 4 [22, p. 61 sq.]**.** *Let $H_I \in \mathcal{D}$, $\emptyset \neq I \subseteq [d]$, such that*

$$\prod_{j=1}^{d} H_j^{\binom{d-1}{j-1}}(u) = u, \quad \forall\ u \in [0,1].$$

Furthermore, let $\{Z_I\}_{\emptyset \neq I \subseteq [d]}$ be a family of independent random variables with $Z_I \sim H_I$. Define $\boldsymbol{U}$ by Eq. (11). Then the distribution function of $\boldsymbol{U}$ is an exGMO copula with

$$g_i(u) = \prod_{j=1}^{d+1-i} H_j^{\binom{d-i}{j-1}}(u), \quad \forall\ u \in [0,1],\ i \in \{2, \dots, d\}.$$

5. Upper semilinear and exchangeable GMO copulas

The classes of survival exGMO copulas and upper semilinear copulas are two multivariate generalisations of the same bivariate copula family. The former is attained as survival copulas in the generalisation of the bivariate exogenous shock model. In contrast, the latter is a result of a generalisation of the recursive construction principle for the copula function. Thus, both class are obtained by lifting different features of the same bivariate class to higher dimensions. However, to the best of our knowledge, the similarities and differences between these two generalisations have not been investigated in the scientific literature, yet.

A natural problem is the identification of conditions that allow a copula to be both survival exGMO and upper semilinear. It should be noted that a natural stochastic model for upper semilinear copulas, or at least a subclass thereof, has yet to be proposed. Thus, this problem is related to the problem of finding stochastic representations for upper semilinear copulas. In the remainder of this section, we present such conditions. These correspond to strong restrictions on the shock model representation as well as a strong restriction on the possible choices of diagonal functions in the recursive construction principle for upper semilinear copulas.

5.1. Bivariate semilinear copulas

We start by proving for the bivariate case that the classes of survival exGMO copulas and upper semilinear copulas are identical. In order to observe this, we recall results on bivariate semilinear copulas from [4]. In the bivariate case, an upper semilinear copula candidate function has the form

$$C(\boldsymbol{u}) = \frac{(1 - u_{(2)}) \cdot \delta_2(u_{(1)}) + (u_{(2)} - u_{(1)}) \cdot u_{(1)}}{1 - u_{(1)}}, \quad \boldsymbol{u} \in [0,1]^2.$$

A simple calculation shows that the survival counterpart, $\hat{C}$, of C is defined by the following equations.

$$\hat{C}(\boldsymbol{u}) = V_C([\mathbf{1} - \boldsymbol{u}, \mathbf{1}]) = u_{(1)} \cdot \frac{\delta_2^L(u_{(2)})}{u_{(2)}}, \qquad \boldsymbol{u} \in [0,1]^2. \tag{12a}$$

$$\delta_2^L(u) = 2u - 1 + \delta_2(1-u), \qquad u \in [0,1]. \tag{12b}$$

A copula of this form is called *(bivariate) lower semilinear copula* (abbr. as LSL_2). We define $g_2(u) := \delta_2^L(u)/u,\ u \in [0,1]$, and we prove the claim by comparing Eq. (12a) with Eq. (9). The linear segments of bivariate lower and upper semilinear copulas are illustrated in Fig. 3a.

5.2. Characterisation of $\text{USL}_d \cap \text{surv. exGMO}_d$ for $d > 2$

Note that there are simple examples for copulas which are survival exGMO but not upper semilinear (e.g. the independence copula for $d > 2$) as well as for copulas which are survival exGMO and upper semilinear (e.g. the comonotonicity copula). Another example, which is more complicated, is that of a copula which is upper semilinear but not survival exGMO (non-comonotonic with identical multivariate diagonal functions), see Section 6. As a result, we have already established that neither the discussed intersection of copula classes is empty nor is one class a subclass of the other.

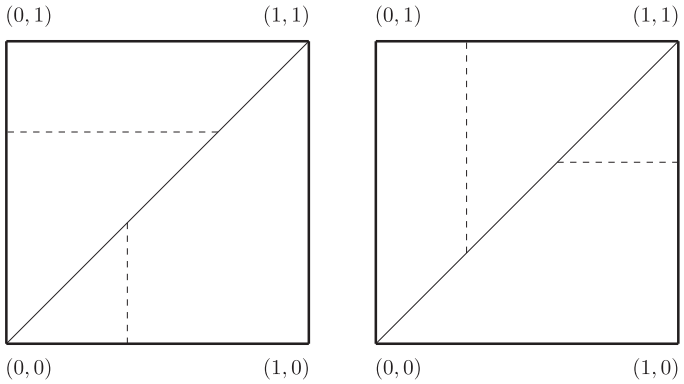


Fig. 3. Illustration of "linear segments" (dashed lines) of lower and upper semilinear copulas for $d = 2$, cf. [4, p. 65].

Below, we present necessary and sufficient conditions for a copula to be both survival exGMO and upper semilinear.

Theorem 3. *Let $d > 2$ and consider the notation of Theorem 2. A d-variate exGMO copula C is the survival copula of an upper semilinear copula $\hat{C}$, if and only if*

$$H_I \equiv 1_{[0,\infty)}, \quad \forall I \ : \ |I| < d-1. \tag{13}$$

Remark. The condition in Eq. (13) has the following intuitive interpretation. For this purpose, consider the stochastic representation in Eq. (11), i.e. $Z_I \sim H_I$, $\emptyset \neq I \subseteq [d]$, are independent random variables and $\boldsymbol{U} \sim C$ is defined by

$$U_i := \max\left\{Z_I \ : \ i \in I\right\}, \qquad i \in \{1,\dots,d\}. \qquad \text{(11 rev.)}$$

Thus, Eq. (13) becomes equivalent to

$$Z_I = 0 \text{ a.s.} \quad \forall I \ : \ |I| < d-1$$

and the shock model simplifies considerably

$$U_i := \max\left\{Z_I \ : \ i \in I, \ |I| \in \{d-1, d\}\right\}, \qquad i \in \{1,\dots,d\}.$$

This imposes a strong constraint on the original exchangeable exogenous shock model definition. In this regard, these copulas constitute a rather small subclass of all survival exGMO copulas.

Proof of Theorem 3. Firstly, we prove by contradiction that the exchangeable exogenous shock model of an exGMO copula whose survival copula is upper semilinear fulfills Eq. (13). Therefore, assume that C is a d-variate exGMO copula and that its survival copula $\hat{C}$ is upper semilinear. Furthermore, assume that there exists $i < d-1$ such that $H_i \not\equiv 1_{[0,\infty)}$. Exchangeability of the shock model implies the existence of sets $\emptyset \neq I_1, I_2, I_3 \subseteq [d]$ with $|I_j| = i$,

$$I_j \cap \{1,2,3\} = \{j\}, \quad j \in \{1,2,3\}$$

and $0 < \epsilon_1 < \epsilon_2 < \epsilon_3 < \epsilon_4 \leq 1$ such that

$$\begin{aligned}
\mathbb{P}(U_1 \neq U_2, U_2 \neq U_3, U_1 \neq U_3) &\geq \mathbb{P}\left(\left(\max_{\substack{\emptyset \neq I \subseteq \{1,\dots,d\} \\ I_j \neq I, \{1,2,3\} \cap I \neq \emptyset}} Z_I\right) < Z_{I_1} < Z_{I_2} < Z_{I_3}\right) \\
&\geq \left[\prod_{\emptyset \neq I \subseteq [d], I_j \neq I, \{1,2,3\} \cap I \neq \emptyset} H_{|I|}(\epsilon_1)\right] \cdot [H_i(\epsilon_2) - H_i(\epsilon_1)] \times [H_i(\epsilon_3) - H_i(\epsilon_2)] \cdot [H_i(\epsilon_4) - H_i(\epsilon_3)] \overset{(*)}{>} 0.
\end{aligned}$$

Here, $(*)$ holds because the assumptions imply the existence of an interval $(a, b] \subseteq [0, 1]$ on which the functions H_{I_j} are strictly monotone. Consequently, $(*)$ holds for arbitrary $\epsilon_j \in (a, b]$ with $\epsilon_1 < \dots < \epsilon_4$, see Fig. 4. This contradicts the result of Lemma 3, i.e. that all components of $\mathbf{1} - \boldsymbol{U}$ are concentrated on at most two distinct points.

Secondly, we prove that Eq. (13) is a sufficient qualification such that the survival copula of an exGMO copula is upper semilinear. Therefore, assume that C is a d-variate exGMO copula such that Eq. (13) holds. Lemma 4 implies that $g_i \equiv 1_{[0,\infty)}$ for $i > 2$ and

$$C(\boldsymbol{u}) = u_{(1)} \cdot g_2(u_{(2)}), \qquad \boldsymbol{u} \in [0,1]^d,$$

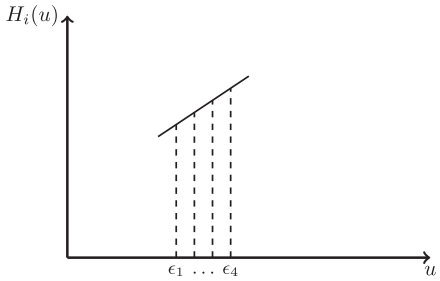


Fig. 4. A stylised strictly monotone section of H_i and a possible choice for $\epsilon_1, \dots, \epsilon_4$.

$$g_2(u) = H_{d-1}(u), \qquad u \in [0,1].$$

Subsequently, we obtain for $\boldsymbol{u} \in [0,1]^d$

$$\begin{aligned}
\hat{C}(\boldsymbol{u}) &= 1 + \sum_{k=1}^{d} (-1)^k \sum_{1 \le i_1 < \dots < i_k \le d} C_k(1-u_{(i_1)}, \dots, 1-u_{(i_k)}) \\
&= 1 - \sum_{i=1}^{d} (1-u_{(i)}) + \sum_{k=2}^{d} (-1)^k \sum_{1 \le i_1 < \dots < i_k \le d} (1-u_{(i_k)}) \cdot g_2(1-u_{(i_{k-1})}) \\
&= 1 - \sum_{i=1}^{d} (1-u_{(i)}) + \sum_{k=1}^{d-1} \sum_{j=k+1}^{d} (1-u_{(j)}) \cdot g_2(1-u_{(k)}) \sum_{i=0}^{k-1} (-1)^{i+2} \binom{k-1}{i} \\
&= 1 - \sum_{i=1}^{d} (1-u_{(i)}) + g_2(1-u_{(1)}) \sum_{i=2}^{d} (1-u_{(i)}).
\end{aligned}$$

Here, we used the equation $\sum_{i=0}^{k-1} (-1)^i \binom{k-1}{i} = 1_{\{k=1\}}$. We conclude that the k-diagonal of $\hat{C}$ equals $\delta_k(u) = u - (k-1)(1-u)(1-g_2(1-u))$ and that

$$g_2(1-u) = 1 + \frac{\delta_k(u) - \delta_{k-1}(u)}{1-u}, \quad k \in \{2, \dots, d\}, \ u \in [0,1].$$

Finally, we plug this identity into the last equation for $\hat{C}$ and can ascertain after a lengthy but straightforward calculation that

$$\hat{C}(\boldsymbol{u}) = \frac{\sum_{i=1}^{d} (u_{(i+1)} - u_{(i)}) \cdot \delta_i(u_{(1)})}{1 - u_{(1)}}, \quad \boldsymbol{u} \in [0,1]^d.$$

Here, we use the convention $u_{(d+1)} \equiv 1$. This shows that $\hat{C}$ is an upper semilinear copula. □

Remark. Recall that one property of upper semilinear copulas is that all components of a realisation are concentrated on at most two distinct (random) values. Theorem 3 implies that, if we additionally assume that its survival copula is exGMO, at most one component of a realisation may differ from the joint minimum. Furthermore, it follows that for $d > 2$, the only extendible copula of this subclass is the comonotonicity copula.

We can combine Theorem 3 with Theorem 2 to obtain the following analytical characterisation for an upper semilinear copula with a survival exGMO copula.

Corollary 3. *Let δ_k, $k \in \{2, \dots, d\}$, be k-diagonals and let C be the corresponding upper semilinear copula candidate function defined by Eq. (6). C is a copula and has an exGMO survival copula if and only if*

$$\delta_k(u) - \delta_{k-1}(u) = \delta_j(u) - \delta_{j-1}(u), \quad \forall k \neq j, \ u \in [0,1],$$

and the functions g_2 and $u/g_2(u)^{d-1}$ are non-decreasing on $[0,1]$ with $g_2(1) = 1$, where g_2 is (for some $k \ge 2$) defined by

$$g_2(u) = 1 + \frac{\delta_k(1-u) - \delta_{k-1}(1-u)}{u}, \quad u \in [0,1].$$

The second condition, namely g_2 and $u/g_2(u)^{d-1}$ being non-decreasing on $[0,1]$, can be replaced such that we attain the following characterisation, cf. [4, Corollary 5].

Corollary 4. *Let δ_k, $k \in \{2, \dots, d\}$, be k-diagonals and let C be the corresponding upper semilinear copula candidate function defined by Eq. (6). Assume that*

$$\delta_k(u) - \delta_{k-1}(u) = \delta_j(u) - \delta_{j-1}(u), \quad \forall k \neq j, \ u \in [0,1],$$

and define for some $k \geq 2$

$$g_2(u) = 1 + \frac{\delta_k(1-u) - \delta_{k-1}(1-u)}{u}, \quad u \in [0,1].$$

Furthermore, assume that g_2 *is absolutely continuous and* $g_2(u) > 0, \forall u > 0$. *Then,* C *is a copula if and only if*

$$0 \leq u \cdot (d-1) \cdot g_2'(u) \leq g_2(u),$$

for all $u \in (0,1)$ *where* $g_2'(u)$ *exists.*

Proof. Corollary 3 implies that C is a copula if and only if $g_2(u)$ and $u/(g_2(u))^{d-1}$ are non-decreasing in u. The claim of this corollary follows, since g_2 is absolutely continuous and it holds that

$$\begin{aligned} & g_2(u) \text{ and } \frac{u}{(g_2(u))^{d-1}} \text{ non-decreasing} \\ \Leftrightarrow \quad & g_2'(u) \geq 0 \text{ and } \left(\log \frac{u}{(g_2(u))^{d-1}}\right)' \geq 0 \text{ a.e.} \\ \Leftrightarrow \quad & g_2(u) \geq u \cdot (d-1) \cdot g_2'(u) \geq 0 \text{ a.e.} \end{aligned}$$

□

Remark. Adapting the notation of Theorem 1, we can show for $C \in \mathrm{USL}_d \cap \mathrm{surv.exGMO}_d$ that

$$\begin{aligned} & \left(\frac{\delta_d(u)}{1-u}\right)' \geq \frac{1-\zeta_d(u)}{(1-u)^2}, \quad \forall u \in [0,1] \\ \Leftrightarrow \quad & u \cdot (d-1) \cdot g_2'(u) \leq g_2(u), \quad \forall u \in [0,1]. \end{aligned}$$

Furthermore, we can show that the non-decreasingness of g_2 is equivalent to the non-decreasingness of $\nu_d^{(m)}$ or the non-increasingness of ζ_d, respectively.

6. Identical multivariate diagonals

In this section, we explore the special case in which all multivariate diagonals are identical. In particular, throughout this section we assume that C is an upper semilinear copula (candidate function) with diagonals $\delta_j \equiv \delta$, $j \geq 2$, for a d-diagonal function δ.

This assumption implies for a realisation $\boldsymbol{U} \sim C$ and distinct i, j, k that

$$\mathbb{P}\big(U_i \leq u, U_j \leq u, U_k > u\big) = 0, \quad \forall u \in [0,1].$$

Thus, at most one component U_i, $i \in [d]$, of $\boldsymbol{U}$ may differ from the joint maximum $\max_{i \in [d]} U_i$. Consequently, and because $\boldsymbol{U}$ is exchangeable, we can reduce sampling $\boldsymbol{U} \sim C$ to sampling the first two components of the ordered version of $\boldsymbol{U}$ and a random shuffling. All bivariate random vectors have a stochastic model, which is based on conditional sampling, see, e.g., [16, Algorithm 1.2]. We will use this to construct a stochastic model for $\boldsymbol{U} \sim C$.

The subclass in question has already been discussed and characterised in [2]. Therein, the authors established that the copula (candidate function) simplifies to

$$C(\boldsymbol{u}) = \frac{(1-u_{(2)}) \cdot \delta(u_{(1)}) + (u_{(2)} - u_{(1)}) \cdot u_{(1)}}{1-u_{(1)}}, \quad \boldsymbol{u} \in [0,1]^d.$$

Furthermore, they present the following theorem, which is a refinement of their general characterisation theorem, see Theorem 1.

Theorem 4 (Characterisation, see [2, Corollary 1])**.** *Let* C *be an upper semilinear copula candidate function such that* $\delta_j \equiv \delta$, $j \in \{2, \ldots, d\}$. C *is a copula if and only if the function* ζ *is non-increasing and the functions* ϕ *and* ν *are non-decreasing, where*

$$\begin{aligned} & \zeta : [0,1] \to [0,1], \ u \mapsto 1 - du + (d-1)\delta(u), \\ & \phi : [0,1) \to \mathbb{R}, \ u \mapsto \frac{1 - du + (d-1)\delta(u)}{(1-u)^d}, \\ & \nu : [0,1) \to \mathbb{R}, \ u \mapsto \frac{u - \delta(u)}{1-u}. \end{aligned}$$

The assumption of identical multivariate diagonals allows us to simplify several expressions. We use these simplifications, which are summarised in the following two lemmas, to derive the probability distributions for the conditional sampling of $(U_{(1)}, U_{(2)})$.

Lemma 5. *Let C be an upper semilinear copula candidate function with equal diagonals $\delta_j \equiv \delta$, $j \in \{2, \dots, d\}$. Then, for $u > 0$, we have*

$$\nu_d^{(m)}(u) = \begin{cases} \frac{u-\delta(u)}{1-u} & m = d-1, \\ 0 & \text{else} \end{cases} \quad \text{and}$$
$$\zeta_d(u) = 1 - du + (d-1)\delta(u).$$

Proof. Firstly, note that the binomial formula implies for $m \geq 1$ the identity $\sum_{j=0}^{m} (-1)^j \binom{m}{j} = 0$. Hence, we have

$$\begin{aligned} \nu_d^{(m)}(u) &= \frac{1}{1-u} \sum_{j=0}^{m} (-1)^j \binom{m}{j} \delta_{d-m+j}(u) \\ &= \begin{cases} (u-\delta(u))/(1-u) & , \quad m = d-1, \\ 0 & , \quad \text{else}, \end{cases} \end{aligned}$$

where we use the assumption that all multivariate diagonals are identical, i.e. $\delta_{d-m+j} = \delta, d-m+j > 1$. Similarly, we can obtain the second identity with

$$\zeta_d(u) = 1 + \sum_{j=1}^{d} (-1)^j \binom{d}{j} \delta_j(u) = 1 - du + d\delta(u) - \delta(u).$$

□

Lemma 6. *Let C be an upper semilinear copula with equal diagonals $\delta_j \equiv \delta$, $j \in \{2, \dots, d\}$. Then, for $0 \leq u < v \leq 1$ and $0 \leq u_1 < v_1 \leq u_2 < v_2 \leq 1$, we have*

$$V_C((u,v]^d) = \delta(v) - \delta(u) - d\,\frac{v-u}{1-u}\,(u - \delta(u)) \text{ and}$$
$$V_C\big((u_1,v_1] \times (u_2,v_2]^{d-1}\big) = (v_2-u_2) \cdot \left[\frac{v_1 - \delta(v_1)}{1-v_1} - \frac{u_1 - \delta(u_1)}{1-u_1} \right].$$

Proof. We use Corollary 1 and Lemma 5 to prove the claim. Particularly, we have

$$\begin{aligned} V_C((u,v]^d) &= \delta(v) - \left(\frac{1-v}{1-u} \cdot \delta(u) + \frac{v-u}{1-u} \cdot (1 - \zeta_d(u)) \right) \\ &= \delta(v) - \left(\frac{1-v}{1-u} \cdot \delta(u) + \frac{v-u}{1-u} \cdot (1 - 1 + du - (d-1) \cdot \delta(u)) \right) \\ &= \delta(v) - \delta(u) - d\,\frac{v-u}{1-u}\,(u - \delta(u)) \end{aligned}$$

and

$$\begin{aligned} V_C\big((u_1,v_1] \times (u_2,v_2]^{d-1}\big) &= (v_2-u_2) \cdot \left(\nu_d^{(d-1)}(v_1) - \nu_d^{(d-1)}(u_1) \right) \\ &= (v_2-u_2) \cdot \left(\frac{v_1 - \delta(v_1)}{1-v_1} - \frac{v_1 - \delta(u_1)}{1-u_1} \right). \end{aligned}$$

□

6.1. Conditional sampling approach

The fundamental idea behind the conditional sampling approach for bivariate random vectors is to perform a separation of the joint probability distribution. In our case, we can write this separation as (for $0 \leq u < v \leq 1$)

$$\begin{aligned} &\mathbb{P}\big(U_{(1)} \leq u, U_{(2)} \leq v\big) \\ &\quad = \mathbb{P}\big(U_{(1)} = U_{(2)}\big) \cdot \mathbb{P}\big(U_{(1)} \leq u \;\big|\; U_{(1)} = U_{(2)}\big) + \mathbb{P}\big(U_{(1)} \neq U_{(2)}\big) \cdot \\ &\qquad \int \mathbb{P}\big(U_{(2)} \in \mathrm{d}v \;\big|\; U_{(1)} = u, U_{(1)} \neq U_{(2)}\big) \cdot \mathbb{P}\big(U_{(1)} \in \mathrm{d}u \;\big|\; U_{(1)} \neq U_{(2)}\big). \end{aligned}$$

We calculate the involved probabilities and (conditional) probability functions in the following two lemmas.

Lemma 7. *Let C be an upper semilinear copula with equal diagonals $\delta_j \equiv \delta$, $j \in \{2, \dots, d\}$. Furthermore, let $\boldsymbol{U} \sim C$ and define the corresponding order-statistic by $U_{(1)} \leq \dots \leq U_{(d)}$. Then, for u, $v \in [0, 1]$ with $u < v$, we have*

$$\mathbb{P}\big(U_{(1)} \leq u, U_{(1)} = U_{(2)}\big) = \delta(u) - d \int_0^u \frac{x - \delta(x)}{1-x}\,\mathrm{d}x,$$

$$\mathbb{P}\big(U_{(1)} \le u, U_{(2)} \le v, U_{(1)} \neq U_{(2)}\big) = d\left(\frac{v-u}{1-u}\big(u-\delta(u)\big) + \int_0^u \frac{x-\delta(x)}{1-x}\,\mathrm{dx}\right).$$

Proof. We prove the two identities with results from measure theory. In particular, we use additivity and σ-continuity of the corresponding probability measure. For both identities, let $u>0$ and let $\{\mathcal{Z}_n\}_{n\in\mathbb{N}}$ be a refining partition of $[0,u]$ defined by

$$\mathcal{Z}_n := \{0 = x_{1,n} < x_{2,n} < \ldots < x_{n-1,n} < x_{n,n} = u\}$$

such that $\mathrm{Mesh}(\mathcal{Z}_n) := \max_{k\le n} |x_{k,n} - x_{k-1,n}| \to 0$ for $n\to\infty$.

For the first identity, consider that

$$\dot{\bigcup_{k=1}^{n}} (x_{k-1,n}, x_{k,n}]^d \downarrow \big\{\boldsymbol{x}\in[0,u]^d : x_1 = \ldots = x_d\big\}.$$

Then, we use additivity and σ-continuity as well as exchangeability and the identities from Lemma 6 to establish

$$\begin{aligned}
&\mathbb{P}\big(U_{(1)} \le u, U_{(1)} = U_{(2)}\big)\\
&= \lim_{n\to\infty} \sum_{k=1}^{n} \delta(x_{k,n}) - \delta(x_{k-1,n}) - d\,\frac{x_{k,n}-x_{k-1,n}}{1-x_{k-1,n}} \cdot (x_{k-1,n} - \delta(x_{k-1,n}))\\
&= \delta(u) - d\int_0^u \frac{x-\delta(x)}{1-x}\,\mathrm{dx}.
\end{aligned}$$

For the second identity, consider that it holds for all $v\ge u$ that

$$\begin{aligned}
&\dot{\bigcup_{k=1}^{n}} \dot{\bigcup_{i=1}^{d}} (x_{k,n}, v]^{i-1} \times (x_{k-1,n}, x_{k,n}] \times (x_{k,n}, v]^{d-i}\\
&\quad\uparrow \big\{\boldsymbol{x}\in[0,1]^d : \exists i \text{ s.t. } x_i \le u, x_j \in (x_i, v]\ \forall j\neq i\big\}.
\end{aligned}$$

Thus, we use the same techniques as in the previous identity and integration-by-parts to show

$$\begin{aligned}
&\mathbb{P}\big(U_{(1)} \le u, U_{(2)} \le v, U_{(1)} \neq U_{(2)}\big)\\
&= d\lim_{n\to\infty} \sum_{k=1}^{n} (v - x_{k,n})\left(\frac{x_{k,n}-\delta(x_{k,n})}{1-x_{k,n}} - \frac{x_{k-1,n}-\delta(x_{k-1,n})}{1-x_{k-1,n}}\right)\\
&= d\int_0^u (v-x)\,\mathrm{d}\left(\frac{x-\delta(x)}{1-x}\right)\\
&= d\left(\frac{v-u}{1-u}\cdot(u-\delta(u)) + \int_0^u \frac{x-\delta(x)}{1-x}\,\mathrm{dx}\right).
\end{aligned}$$

Note that we require $v(u) = (u-\delta(u))/(1-u)$ to be of bounded variation, which is fulfilled if v is non-decreasing, such that this limit can be interpreted as a Riemann–Stieltjes Integral. □

Lemma 8. *Let C be an upper semilinear copula with equal diagonals $\delta_j \equiv \delta$, $j\in\{2,\ldots,d\}$. Define $p := d\int_0^1 \big(x-\delta(x)\big)/(1-x)\,\mathrm{dx}$. Furthermore, let $\boldsymbol{U}\sim C$ and define the corresponding order-statistic by $U_{(1)}\le\ldots\le U_{(d)}$. Then, for $u, v\in[0,1]$ with $u<v$, we have*

$$\mathbb{P}\big(U_{(1)} \neq U_{(2)}\big) = p = d\int_0^1 \frac{x-\delta(x)}{1-x}\,\mathrm{dx}.$$

If $p\neq 1$, we have

$$\mathbb{P}\big(U_{(1)} \le u \mid U_{(1)} = U_{(2)}\big) = \frac{1}{1-p}\left(\delta(u) - d\int_0^u \frac{x-\delta(x)}{1-x}\,\mathrm{dx}\right).$$

Furthermore, if C is not equal to the comonotonicity copula, i.e. $p\neq 0$, we furthermore have

$$\begin{aligned}
&\mathbb{P}\big(U_{(1)} \le u, U_{(2)} \le v \mid U_{(1)} \neq U_{(2)}\big) = \frac{d}{p}\left(\frac{v-u}{1-u}\big(u-\delta(u)\big) + \int_0^u \frac{x-\delta(x)}{1-x}\,\mathrm{dx}\right),\\
&\mathbb{P}\big(U_{(2)} \le v \mid U_{(1)} = u, U_{(1)} \neq U_{(2)}\big) = \frac{v-u}{1-u}.
\end{aligned}$$

Proof. The first three identities follow directly from Lemma 7, where we use the fact that $p\neq 1$ in the second equation and that $p\neq 0$ in the third equation.

For the remaining identity, consider the following argument. Let $\boldsymbol{U}$ be defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\Omega' := \{\omega \in \Omega\colon U_{(1)}(\omega) \neq U_{(2)}(\omega)\}$, $\mathbb{P}_{\Omega'} := \mathbb{P}(\cdot \cap \Omega')/\mathbb{P}(\Omega')$, and consider the probability space $(\Omega, \mathcal{F}, \mathbb{P}_{\Omega'})$. Consequently, it suffices to show that

$$\mathbb{P}_{\Omega'}\big(U_{(2)} \leq v \,\big|\, U_{(1)} = u\big) = \frac{v-u}{1-u}, \quad \forall u, v \in [0,1],\ u < v.$$

From the third identity of this lemma, we conclude that the push-forward measure of $\boldsymbol{U}$ with respect to $\mathbb{P}_{\Omega'}$ is absolutely continuous with respect to the Lebesgue measure with density

$$f_{U_{(1)},U_{(2)}}(u,v) = \frac{d}{p}\,\frac{\big(1-\delta'(u)\big)(1-u) + \big(u-\delta(u)\big)}{(1-u)^2} \cdot 1_{(u,1]}(v), \quad u, v \in [0,1].$$

Here, we use that δ is Lipschitz continuous and therefore, according to Rademacher's theorem, δ is differentiable with derivative δ' almost everywhere (with respect to the Lebesgue measure). Finally, we conclude that the desired conditional density and conditional distribution function can be written for $u, v \in [0,1]$ as

$$f_{U_{(2)}\big|U_{(1)}=u}(v) = \frac{f_{U_{(1)},U_{(2)}}(u,v)}{\int_0^1 f_{U_{(1)},U_{(2)}}(u,v)\,\mathrm{d}\,v} = \frac{1}{1-u}1_{(u,1]}(v),$$

$$F_{U_{(2)}\big|U_{(1)}=u}(v) = \int_0^v \frac{1}{1-u}1_{(u,1]}(v)\ \mathrm{d}\,x = \frac{\max\{v-u,0\}}{1-u}.$$

This proves the claim. □

Note that the (conditional) distribution function of $[U_{(2)}|U_{(1)} \neq U_{(2)}, U_{(1)} = u]$ corresponds to the uniform distribution on $[u, 1]$.

In conclusion, we have obtained the desired stochastic model for arbitrary upper semilinear copulas C with equal multivariate diagonals, see Algorithm 1.

Algorithm 1 Sampling algorithm for an upper semilinear copula C with equal multivariate diagonals $\delta_j \equiv \delta$, $j \in \{2, \ldots, d\}$.

input An admissible d-diagonal δ.
output A sample from the upper semilinear copula with diagonal δ.
function USLC(δ)
  Draw $I \sim$ Bernoulli(p) with $p = d\int_0^1 \frac{x-\delta(x)}{1-x}\,\mathrm{dx}$.
  if $I = 0$ **then**
    Draw $U_\wedge \sim \Big(\delta(u) - d\int_0^u \frac{x-\delta(x)}{1-x}\,\mathrm{dx}\Big)/(1-p)$.
    Set $U_1 = \ldots = U_d = U_\wedge$.
  else
    Draw $U_\wedge \sim d\Big((u-\delta(u)) + \int_0^u \frac{x-\delta(x)}{1-x}\,\mathrm{dx}\Big)/p$.
    Draw $U_\vee \sim \mathcal{U}_{[U_\wedge,1]}$.
    Draw K uniform from the set $\{1, \ldots, d\}$.
    Set $U_K := U_\wedge$ and $U_j := U_\vee$, $j \neq K$.
  end if
  return $\boldsymbol{U} = (U_1, \ldots, U_d)'$.
end function

Besides the possibility of sampling from those copulas, we can use these results to simplify the conditions from Theorem 4.

Corollary 5. *Let C be an upper semilinear copula candidate function with equal diagonals $\delta_j \equiv \delta$, $j \in \{2, \ldots, d\}$. Then, C is a copula if and only if ν is non-decreasing and bounded by $\delta'(x)/d$ almost everywhere (with respect to the Lebesgue measure), where*

$$\nu : [0,1) \to [0,\infty),\ x \mapsto \frac{x-\delta(x)}{1-x}.$$

Proof. Note that

$$\left(\delta(u) - d\int_0^u \frac{x-\delta(x)}{1-x}\,\mathrm{dx}\right)' = \delta'(u) - d\frac{u-\delta(u)}{1-u} \geq 0 \text{ a.e.}$$

is equivalent to

$$\frac{u-\delta(u)}{1-u} \leq \frac{1}{d}\delta'(u) \text{ a.e.}$$

For the forward direction, assume that C is a copula. The non-decreasingness of ν follows from Theorem 4. With Lemma 7, we conclude that $\delta(u) - d\int_0^u \nu(x)\,\mathrm{d}x$ is non-decreasing as well, which implies $\nu(u) \le \delta'(u)/d$ a.e.

For the reverse direction, assume that ν is non-decreasing and that $\nu(u) \le \delta'(u)/d$ holds a.e. Combined with the diagonal properties, this implies that $u \mapsto \delta(u) - d\int_0^u \nu(x)\,\mathrm{d}x$ is non-decreasing as well as bounded from below by zero and from above by $\delta(u)$. In particular, p (as defined in the previous lemmas) fulfills $0 \le p \le 1$. With integration-by-parts, we obtain

$$\big(u - \delta(u)\big) + \int_0^u \frac{x - \delta(x)}{1-x}\,\mathrm{dx} = \int_0^u \left(\frac{x-\delta(x)}{1-x}\right)' \cdot (1-x)\,\mathrm{dx}\,.$$

The non-decreasingness of ν implies that this function is also non-decreasing. Furthermore, this implies that it is bounded from below by zero and from above by p. Moreover, due to the continuity of the Riemann-integral as well as the continuity of the diagonals δ, both functions are continuous. In conclusion, we attain that the functions

$$u \mapsto \frac{1}{1-p}\left(\delta(u) - d\int_0^u \frac{x-\delta(x)}{1-x}\,\mathrm{d}x\right),$$
$$u \mapsto \frac{d}{p}\left(\big(u-\delta(u)\big) + \int_0^u \frac{x-\delta(x)}{1-x}\,\mathrm{d}x\right)$$

are continuous distribution functions on $[0,\,1]$ and $p = d\int_0^1 \nu(x)\,\mathrm{d}x$ is a probability, i.e. $p \in [0,\,1]$. Consequently, we can apply Algorithm 1. Therefore, let $\boldsymbol{U}$ be a realisation of Algorithm 1 and let $\boldsymbol{u} \in [0,\,1]$. Since the corresponding copula is exchangeable, we can restrict our proof w.l.o.g. to the special case $u_1 \le \ldots \le u_d$. We have

$$\begin{aligned}
&\mathbb{P}(U_1 \le u_1, \ldots, U_d \le u_d)\\
&\quad= \mathbb{P}\big(U_{(1)} = U_{(2)}\big)\cdot \mathbb{P}\big(U_{(1)} \le u_1 \big| U_{(1)} = U_{(2)}\big)\\
&\quad+ \mathbb{P}\big(U_{(1)} \ne U_{(2)}\big)\Bigg[\mathbb{P}\big(K=1\big|U_{(1)}\ne U_{(2)}\big)\cdot \underbrace{\mathbb{P}\big(U_1\le u_1, U_2 \le u_2 \big| U_{(1)}\ne U_{(2)}, K=1\big)}_{=\mathbb{P}\left(U_{(1)}\le u_1, U_{(2)}\le u_2 \,\middle|\, U_{(1)}\ne U_{(2)}\right)}\\
&\qquad+ \sum_{k=2}^{d}\mathbb{P}\big(K=k\big|U_{(1)}\ne U_{(2)}\big)\cdot \underbrace{\mathbb{P}\big(U_k\le u_1, U_1\le u_1\big|U_{(1)}\ne U_{(2)}, K=k\big)}_{=\mathbb{P}\left(U_{(1)}\le u_1 \,\middle|\, U_{(1)}\ne U_{(2)}\right)}\Bigg]\\
&\quad= (1-p)\cdot\frac{1}{1-p}\left(\delta(u_1) - d\int_0^{u_1}\frac{x-\delta(x)}{1-x}\,\mathrm{dx}\right)\\
&\quad+ p\cdot\frac{1}{d}\cdot\left[\frac{d}{p}\left(\frac{u_2-u_1}{1-u_1}(u_1-\delta(u_1)) + \int_0^{u_1}\frac{x-\delta(x)}{1-x}\,\mathrm{d}x\right) + \sum_{k=2}^{d}\frac{d}{p}\left(\int_0^{u_1}\frac{x-\delta(x)}{1-x}\,\mathrm{d}x\right)\right]\\
&\quad= \delta(u_1) + \frac{u_2-u_1}{1-u_1}(u_1-\delta(u_1))\\
&\quad= \frac{\delta(u_1) - u_1\delta(u_1) + (u_2-u_1)u_1 - \delta(u_1)u_2 + \delta(u_1)u_1}{1-u_1}\\
&\quad= \frac{(1-u_2)\cdot\delta(u_1) + (u_2-u_1)\cdot u_1}{1-u_1} = C(\boldsymbol{u}).
\end{aligned}$$

□

6.2. Examples

We conclude this section with two examples. Firstly, in light of Corollary 5, we call a d-diagonal *C_d-admissible* if ν is non-decreasing and almost everywhere bounded by $\delta'(u)/d$. It is highly evident that the class of C_d-admissible diagonals is closed under point-wise limits and convex combinations.

Example 1. In [2], it was shown that for fixed $d \ge 2$ the lowest C_d-admissible d-diagonal is

$$\delta(u) = \frac{(1-u)^d + du - 1}{d-1}, \quad u \in [0,1].$$

Then

$$\delta'(u) = d/(d-1)\big(-(1-u)^{d-1} + 1\big)$$

and simple calculations show that $p = 1$,

$$\nu(u) = \frac{u - \delta(u)}{1-u} = \frac{1}{d-1} - \frac{(1-u)^{d-1}}{d-1},$$
$$F_{U_\wedge | U_\wedge \neq U_\vee}(u) = 1 - (1-u)^d \text{ and}$$
$$F_{U_\vee | U_\wedge = u, U_\wedge \neq U_\vee}(v) = \frac{1-v}{1-u}.$$

Note that in this case, since δ is assumed to be minimal among all C_d-admissible diagonals, we have $\delta'/d = \nu$.

Example 2. We can consider convex combinations between the lowest C_d-admissible diagonal and the comonotonicity diagonal (which is the highest C_d-admissible diagonal), i.e.

$$\delta(u) = \lambda \frac{(1-u)^d + du - 1}{d-1} + (1-\lambda)u, \quad u \in [0,1].$$

Again, simple calculations show that $p = \lambda$,

$$F_{U_\wedge | U_\wedge = U_\vee}(u) = u,$$
$$F_{U_\wedge | U_\wedge \neq U_\vee}(u) = 1 - (1-u)^d \text{ and}$$
$$F_{U_\vee | U_\wedge = u, U_\wedge \neq U_\vee}(v) = \frac{1-v}{1-u}.$$

7. An extendible subclass

In this section, we present a subclass of extendible upper semilinear copulas. Bear in mind that an upper semilinear copula is extendible (in the class of upper semilinear copulas) if an exchangeable sequence $\{U_i\}_{i\in\mathbb{N}}$ exists such that each finite margin is upper semilinear and $(U_1, \ldots, U_d)' \sim C$.

Extendible subclasses are interesting, because they admit a so-called deFinetti representation, see [1, Chapters 2 and 3]. That means a realisation $\boldsymbol{U} \sim C$ can be represented as

$$\boldsymbol{U} = (F^{\leftarrow}(\tilde{U}_1), \ldots, F^{\leftarrow}(\tilde{U}_d))'$$

for a random distribution function F with generalised inverse $F^{\leftarrow}$ and iid uniform random variables $\tilde{U}_i$, $i \in [d]$, independent of F.[9] Models of deFinetti's kind are interesting, as they can easily be generalised in an efficient way to higher dimensions by considering an iid uniform family $\{\tilde{U}_i\}_{i\in\mathbb{N}}$ instead of $\tilde{U}_i$, $i \in [d]$.

We construct the extendible subclass directly by its deFinetti representation. For this representation, the idea is to distribute two independent uniform random variables to a vector by iid Bernoulli experiments.

Theorem 5. *Let $q \in [1/2, 1]$ and consider the following stochastic model: Let $\boldsymbol{V} \sim \Pi_2$ and $\{J_i\}_{i\in\mathbb{N}}$ an iid family of Bernoulli distributed random variables with success parameter q. Define*

$$U_i := J_i V_1 + (1 - J_i)V_2 = \begin{cases} V_1, & \text{if } J_i = 1, \\ V_2, & \text{if } J_i = 0, \end{cases} \qquad i \in \mathbb{N}.$$

Then for each $d \geq 2$, the random vector $\boldsymbol{U} = (U_1, \ldots, U_d)'$ has the distribution function

$$C_d(\boldsymbol{u}) = u_{(1)} \left(q^d + (1-q)^d + \sum_{j=2}^{d} u_{(j)} \left[q(1-q)^{j-1} + (1-q)q^{j-1} \right] \right), \quad \boldsymbol{u} \in [0,1]^d.$$

Furthermore, C_d is an upper semilinear copula as a distribution function and the corresponding diagonal functions are

$$\delta_k(u) = u \left(q^k + (1-q)^k + u \sum_{j=2}^{k} \left[q(1-q)^{j-1} + (1-q)q^{j-1} \right] \right), \quad u \in [0,1].$$

Proof. Firstly, we calculate the distribution function of $\boldsymbol{U}$. Therefore, let $d \geq 2$ and $\boldsymbol{u} \in [0,1]^d$ and assume w.l.o.g. that $u_1 \leq \ldots \leq u_d$. Then, with the convention $\min \emptyset = 1$, we use the tower property to conclude

$$C_d(\boldsymbol{u}) = \sum_{\emptyset \subseteq L \subseteq \{1,\ldots,d\}} \Pi_2 \left(\min_{i\in L} u_i, \min_{i \notin L} u_i \right) \cdot q^{|L|}(1-q)^{d-|L|}$$

[9] For a distribution function F, we define the generalised inverse $F^{\leftarrow}$ by $F^{\leftarrow}(y) = \inf\{x \in \mathbb{R} : F(x) \geq y\}$, where we use the convention $\inf \emptyset = \sup\{\mathrm{ran} F\}$, see [8].

$$= u_1\left(\sum_{\substack{\emptyset\subseteq L\subseteq\{1,\dots,d\}\\ 1\in L}} \min_{i\notin L} u_i\cdot q^{|L|}(1-q)^{d-|L|} + \sum_{\substack{\emptyset\subseteq L\subseteq\{1,\dots,d\}\\ 1\notin L}} \min_{i\in L} u_i\cdot q^{|L|}(1-q)^{d-|L|}\right)$$

$$= u_1\left(q^d + (1-q)^d + \sum_{j=2}^{d} u_j \sum_{l=0}^{d-j}\binom{d-j}{l} q^{l+1}(1-q)^{d-l-1} + \sum_{j=2}^{d} u_j \sum_{l=0}^{d-j}\binom{d-j}{l} q^{d-l-1}(1-q)^{l+1}\right)$$

$$= u_1\left(q^d + (1-q)^d + \sum_{j=2}^{d} u_j\left[q(1-q)^{j-1} + (1-q)q^{j-1}\right]\right).$$

Furthermore, this implies that the corresponding k-diagonal is

$$\delta_k(u) = u\left(q^k + (1-q)^k + u\sum_{j=2}^{k}\left[q(1-q)^{j-1} + (1-q)q^{j-1}\right]\right).$$

Secondly, we prove that C_d is upper semilinear. Therefore, let $\boldsymbol{v}\in[0,1]^k$ with $v_{(k)}=1$ and $\lambda\in(0,1)$ and assume w.l.o.g. that $v_1\le\ldots\le v_k$. Then

$$\begin{aligned}
&C_k(\lambda v_1\mathbf{1} + (1-\lambda)\boldsymbol{v})\\
&\quad= v_1\left(q^k + (1-q)^k + \sum_{j=2}^{k}[\lambda v_1 + (1-\lambda)v_i]\left[q(1-q)^{j-1} + (1-q)q^{j-1}\right]\right)\\
&\quad= \lambda v_1\left(q^k + (1-q)^k + \sum_{j=2}^{k} v_1\left[q(1-q)^{j-1} + (1-q)q^{j-1}\right]\right)\\
&\qquad + (1-\lambda)v_1\left(q^k + (1-q)^k + \sum_{j=2}^{k} v_i\left[q(1-q)^{j-1} + (1-q)q^{j-1}\right]\right)\\
&\quad= \lambda\delta_k(v_1) + (1-\lambda)C_k(\boldsymbol{v}).
\end{aligned}$$

□

The class of upper semilinear copulas is closed under convex combinations and point-wise limits. This implies for this example, that we can randomise the success parameter q in the previous model and stay in the class of upper semilinear copulas.

Corollary 6. *Let F be a distribution function on $[1/2, 1]$ and consider the following stochastic model. Let $Q\sim F$, $V\sim\Pi_2$, and $\{J_i\}_{i\in\mathbb{N}}$ a conditionally iid family of Bernoulli distributed random variables with random success parameter Q. Define*

$$U_i := J_iV_1 + (1-J_i)V_2, \quad i\in\mathbb{N}.$$

Then, for each $d\ge 2$, the random vector $\boldsymbol{U} = (U_1,\ldots,U_d)'$ has the distribution function (for $\boldsymbol{u}\in[0,1]^d$)

$$C_d(\boldsymbol{u}) = u_{(1)}\left(\mathbb{E}[Q^d] + \mathbb{E}[(1-Q)^d] + \sum_{j=2}^{d} u_{(j)}\mathbb{E}[Q(1-Q)^{j-1}] + \mathbb{E}[(1-Q)Q^{j-1}]\right).$$

Furthermore, C_d is an upper semilinear copula and the corresponding diagonal functions are (for $u\in[0,1]$)

$$\delta_k(u) = u\left(\mathbb{E}[Q^d] + \mathbb{E}[(1-Q)^d] + u\sum_{j=2}^{d}\mathbb{E}[Q(1-Q)^{j-1}] + \mathbb{E}[(1-Q)Q^{j-1}]\right). \tag{14}$$

Remark. In the previous sections, we discussed the two subclasses where either the survival copula is exGMO or all multivariate diagonals are identical. For the former subclass, at most one component of a realisation may differ from the joint minimum. In contrast, for the latter subclass, at most one component of a realisation may differ from the joint maximum. In a way, these subclasses are diametrically opposed to each other in the class of upper semilinear copulas. However, both subclasses are similarly restricted, since their realisations can have at most one component differing from all other ones. With regard to this aspect, the subclass presented in this section is very different. If $q\neq 1$ (resp. $F\neq 1_{[1,\infty)}$), all splits of the components have a positive probability.

We conclude this section by exploiting the stochastic representation to derive some of the properties of this class.

Corollary 7. *The copula C_d from Cololary 6 is radially symmetric, has a lower and upper multivariate tail dependence coefficient of $\mathbb{E}[Q^d + (1-Q)^d]$, and is positively orthant dependent.*

Proof. Firstly, we obtain radial symmetry from the stochastic representation and the radial symmetry of Π_2. Secondly, we use Eq. (14) to show that $\lim_{u\to 0} \delta_k(u)/u = \mathbb{E}[Q^d + (1-Q)^d]$. Consequently, the statement for the upper tail dependence coefficient is implied by the radial symmetry of C_d. Finally, with the following calculation, we prove that C_d is positively orthant dependent:

$$\begin{aligned} C_d(\boldsymbol{u}) &= \sum_{\emptyset\subseteq L\subseteq\{1,\dots,d\}} \underbrace{\Pi_2\left(\min_{i\in L} u_i, \min_{i\notin L} u_i\right)}_{\geq \prod_{i=1}^d u_i} \cdot \mathbb{E}[Q^{|L|}(1-Q)^{d-|L|}] \\ &\geq \Pi_d(\boldsymbol{u}), \quad \forall \boldsymbol{u} \in [0,1]^d. \end{aligned}$$

□

8. Conclusion

We provided necessary and sufficient conditions for a copula to be in the intersection of the families of upper semilinear and survival exchangeable generalised Marshall–Olkin copulas. More precisely, the shocks in the corresponding exchangeable exogenous shock model must be almost surely equal to zero for all index sets missing more than one component. This finding has important consequences. In particular, this implies for $d \geq 3$ that the independence copula is not upper semilinear and that the only copula which is upper semilinear, survival exGMO, and admits a deFinetti representation is the comonotonicity copula.

We derived a sampling algorithm based on conditional sampling for the subclass of upper semilinear copulas with identical multivariate diagonals. Firstly, a Bernoulli experiment determines whether all components are identical. Secondly, the minimum $U_\wedge$ and maximum $U_\vee$ are sampled conditionally, based on the outcome of the Bernoulli experiment. In the last step, the position of $U_\wedge$ is drawn with a uniform distribution on the set $\{1,\dots,d\}$. This corresponds to shuffling the ordered components $(U_\wedge, U_\vee, \dots, U_\vee)$ with a uniform distribution on the set of permutations.

For both aforementioned subclasses, we utilised the stochastic representations to derive simplified characterisation theorems.

Finally, we presented a novel subclass of upper semilinear copulas which are radially symmetric and admit a deFinetti representation. In particular, these copulas are conjunctive aggregation operations.

For the special case of equal multivariate diagonals, it remains an open question as to whether the sampling algorithm can be generalised for a larger subclass of upper semilinear copulas.

Declaration of competing interest

There's no financial/personal interest or belief that could affect the objectivity of the submitted research results. No conflict of interests exist.

Acknowledgements

We thank Matjaž Omladič, Tomaž Košir, and Fabrizio Durante as well as the two anonymous reviewers for their helpful comments.

References

[1] D.J. Aldous, Exchangeability and related topics, in: P.L. Hennequin (Ed.), École d'Été de Probabilités de Saint-Flour XIII — 1983, Springer Berlin Heidelberg, Berlin, Heidelberg, 1985, pp. 1–198.
[2] J.J. Arias-Garcìa, H. De Meyer, B. De Baets, Multivariate upper semilinear copulas, Inf. Sci. 360 (2016) 289–300.
[3] F. Durante, E.P. Klement, R. Mesiar, C. Sempi, Conjunctors and their residual implicators: characterizations and construction methods, Mediterranean J. Math. 4 (3) (2007) 343–356.
[4] F. Durante, A. Kolesárová, R. Mesiar, C. Sempi, Semilinear copulas, Fuzzy Sets Syst. 159 (1) (2008) 63–76.
[5] F. Durante, J.J. Quesada-Molina, M. Úbeda Flores, On a family of multivariate copulas for aggregation processes, Inf. Sci. 177 (24) (2007) 5715–5724.
[6] F. Durante, C. Sempi, Principles of Copula Theory, Chapman and Hall/CRC, 2015.
[7] P. Embrechts, R. Frey, A. McNeil, Quantitative Risk Management, Princeton Series in Finance, 5, revised edition, Princeton University Press, 2015.
[8] P. Embrechts, M. Hofert, A note on generalized inverses, Math. Methods Oper. Res. 77 (3) (2013) 423–432.
[9] P. Embrechts, F. Lindskog, A. McNeil, Modelling dependence with copulas and applications to risk management, in: S.T. Rachev (Ed.), Handbook of Heavy Tailed Distributions in Finance, Handbooks in Finance, Vol. 1, Elsevier/North-Holland, Amsterdam, 2003, pp. 329–384.
[10] P. Jaworski, On copulas and their diagonals, Inf. Sci. 179 (17) (2009) 2863–2871.
[11] H. Joe, Multivariate Models and Multivariate Dependence Concepts, CRC Press, 1997.
[12] X. Li, F. Pellerey, Generalized Marshall–Olkin distributions and related bivariate aging properties, J. Multivariate Anal. 102 (10) (2011) 1399–1409.
[13] J. Lin, X. Li, Multivariate generalized Marshall–Olkin distributions and copulas, Methodol. Comput. Appl. Probab. 16 (1) (2014) 53–78.
[14] J.-F. Mai, Extendibility of Marshall–Olkin distributions via Lévy subordinators and an application to portfolio credit risk, Technische Universität München, 2010 Dissertation.
[15] J.-F. Mai, S. Schenk, M. Scherer, Exchangeable exogenous shock models, Bernoulli 22 (2) (2016) 1278–1299.
[16] J.-F. Mai, M. Scherer, Simulating copulas: stochastic models, sampling algorithms and applications, Series in Quantitative Finance, 6, 2, World Scientific, 2017.
[17] A.W. Marshall, I. Olkin, A multivariate exponential distribution, J. Am. Stat. Assoc. 62 (317) (1967) 30–44, doi:10.2307/2282907.

[18] A.J. McNeil, J. Nešlehová, Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions, Ann. Stat. 37 (5B) (2009) 3059–3097.
[19] R.B. Nelsen, An introduction to copulas, Lecture Notes in Statistics, 139, second ed., Springer-Verlag, New York, 2006.
[20] P. Ressel, Monotonicity properties of multivariate distribution and survival functions — with an application to Lévy-frailty copulas, J. Multivariate Anal. 102 (3) (2011) 393–404, doi:10.1016/j.jmva.2010.10.001.
[21] T. Rychlik, Distributions and expectations of order statistics for possibly dependent random variables, J. Multivariate Anal. 48 (1) (1994) 31–42.
[22] S. Schenk, Exchangeable exogenous shock models, Technical University of Munich, Munich, 2016 Dissertation.
[23] B. Schweizer, A. Sklar, Associative functions and statistical triangle inequalities, Publicationes Mathematicae Debrecen 8 (1961) 169–186.
[24] B. Schweizer, A. Sklar, Associative functions and abstract semigroups, Publicationes Mathematicae Debrecen 10 (1963) 69–81.
[25] B. Schweizer, A. Sklar, Probabilistic Metric Spaces, Elsevier/North-Holland, 1983.
[26] A. Sklar, Fonctions de répartition a n dimensions et leurs marges, Publ. Inst. Statist. Univ. Paris 8, 1959, pp. 229–231.

A.2 Implementing Markovian models for extendible Marshall–Olkin distributions

Implementing Markovian models for extendible Marshall–Olkin distributions

Henrik Sloot

The article [2] derives a novel stochastic representation for exchangeable Marshall–Olkin distributions: they correspond to ordered death times associated with certain Markovian death-counting processes, shuffled uniformly at random. Furthermore, it provides a numerically stable approximation of the corresponding Markov generators for the extendible subclass. In summary, it proposes a simulation algorithm for extendible Marshall–Olkin distributions, provides detailed implementation instructions, and analyzes its efficiency.

The introduction outlines the research gap and the article's contribution. For this, I highlight the challenges of simulating multivariate random vectors, particularly in higher dimensions. Following, I summarize existing simulation algorithms for extendible Marshall–Olkin distributions and work out their shortcomings regarding the simulation in higher dimensions. These range from inefficiency due to the *curse of dimensionality* to lack of coverage of all extendible Marshall–Olkin distributions. I conclude with the research question: *Can we find low-parametric Markov-based models for extendible Marshall–Olkin distributions with a numerically stable implementation?*

Section 2 recalls background on Marshall–Olkin distributions, the exchangeable and extendible subclasses, Bernstein functions, and the Lévy-frailty model. In particular, I present several parametric subclasses of extendible Marshall–Olkin distributions and hint at how Marshall–Olkin distributions can be constructed using hierarchical models.

Section 3 investigates death-set processes of Marshall–Olkin distributions to derive their infinitesimal Markov generator matrices. For this, I extend an existing proof of their Markov property and show that they can be represented stochastically as random walks on the semigroup $(\{1, \dots, d\}, \cup)$ subordinated by Poisson processes.

Section 4 proves that exchangeable Marshall–Olkin distributions' death-counting processes are Markov, and it derives their generator matrices. For the proof, I rely on the previous section's separation of transitions and transition times, and I use exchangeability to aggregate transition probabilities of events with matching death counts. Afterward, I obtain a simulation algorithm for exchangeable Marshall–Olkin distributions: simulate their death-counting processes and subsequently shuffle the ordered death times uniformly at random.

Section 5 derives a numerically stable approximation of extendible Marshall–Olkin distribution's generator matrices. This approximation is integral for implementing the previous section's simulation algorithm. First, in a numerical study, I demonstrate that calculating those generator matrices naïvely produces sizeable distortions resulting from loss of significant digits. Second, I derive two integral representations allowing the approximation of generator matrices using numerical integration. Third, I show that the integrands are well-suited for numerical integration except for two boundary cases, for which I provide alternative approximations. Fourth, I obtain a recursion allowing the row-wise calculation of the generator matrices. Finally, I repeat the case study, showing that the proposed approximation works sufficiently well for all cases considered.

Section 6 benchmarks the new simulation algorithm against existing ones. From this, I conclude the following about the new simulation algorithm: First, it is significantly faster and requires less memory than algorithms based on the exogenous shock model and the Arnold model. Second, it is slower than the Lévy-frailty model algorithm for most cases. However, the latter requires tailor-made algorithms for the jump distributions. Moreover, I show that the latter algorithm's runtime explodes for asymptotically increasing probabilities of small jumps. The new algorithm is a generic, reliable simulation algorithm with broad coverage of the extendible subclass.

The article contains appendices on the death-set and death-counting processes' generator matrices for selected subclasses, theoretical runtime boundaries for our simulation algorithm and the Lévy-frailty algorithm, background on statistical testing of implemented simulation algorithms, and Walker's algorithm for sampling discrete random variables.

Statement of individual contribution

I, Henrik Sloot, am the sole author of this article.

Permission to include

H. Sloot. “Implementing Markovian models for extendible Marshall–Olkin distributions”. In: *Dependence Modeling* 10.1 (2022), pp. 308–343. DOI: 10.1515/demo-2022-0151

Reprinted according to the *CC BY 4.0 license*;[2] see human-readable summary included below.

This page is available in the following languages:

Creative Commons License Deed

Attribution 4.0 International (CC BY 4.0)

This is a human-readable summary of (and not a substitute for) the license.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material

for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

[2] https://creativecommons.org/licenses/by/4.0/.

Research Article

Henrik Sloot*

Implementing Markovian models for extendible Marshall–Olkin distributions

https://doi.org/10.1515/demo-2022-0151
received May 20, 2022; accepted November 2, 2022

Abstract: We derive a novel stochastic representation of exchangeable Marshall–Olkin distributions based on their death-counting processes. We show that these processes are Markov. Furthermore, we provide a numerically stable approximation of their infinitesimal generator matrices in the extendible case. This approach uses integral representations of Bernstein functions to calculate the generator's first row, and then uses a recursion to calculate the remaining rows. Combining the Markov representation with the numerically stable approximation of corresponding generators allows us to sample extendible Marshall–Olkin distributions with a flexible simulation algorithm derived from known Markov sampling strategies. Finally, we benchmark an implementation of this Markov-based simulation algorithm against alternative simulation algorithms based on the Lévy frailty model, the Arnold model, and the exogenous shock model.

Keywords: Marshall–Olkin distribution, sampling algorithm, Markov processes, exchangeability, lack of memory, multivariate survival analysis

MSC 2020: 60G09, 60J28, 62-08, 62H05

1 Introduction

Simulating a multivariate distribution can be complex: even if a stochastic model is known, implementing a model into a simulation algorithm requires considering and overcoming numerical problems, methodological challenges, and efficiency obstacles. A typical numerical problem is aggregation operations over large or infinite sets, requiring approximations and introducing a bias. A methodological challenge is choosing between general algorithms covering a large family of distributions or specialized algorithms targeting smaller subfamilies. General algorithms can be valuable for a general-purpose application or cover cases for which a specialized simulation algorithm does not exist. However, specialized simulation algorithms can be less complex, faster, or have a smaller memory footprint. Examples of this are the canonical implementations of the natural models for independent and comonotonic multivariate distributions. An efficiency obstacle often arises from implementation details, e.g., storage requirements and limitations or algorithm choices for subtasks. Thus, developing, selecting, and implementing general or specialized simulation algorithms for a multivariate distribution is an intriguing and challenging research area.

The Marshall–Olkin distribution is a multivariate exponential distribution. It was identified in [19] as the unique class of multivariate distributions whose margins possess the following multivariate generalization of the lack of memory property: given joint survival up to a particular time, the excess-time distribution equals the original distribution. A familiar characterizing property was derived in [5]: the Marshall–Olkin distribution is the unique class of multivariate distributions whose margins' death-set processes have the (homogeneous) Markov property.

* **Corresponding author: Henrik Sloot,** Technical University of Munich, Parkring 11, 85748 Garching-Hochbrück, Germany, e-mail: henrik.sloot@tum.de

The Markov representation allows using simulation algorithms for Markov processes, for example, as described in [4, p. 493], to simulate Marshall–Olkin distributions, given the corresponding Markov generator. However, general models for the entire family of Marshall–Olkin distribution share the *curse of dimensionality*, meaning that the number of parameters grows exponentially with the dimension. Even for low-parametric subclasses, the first row of the Markov generator has $2^d - 1$ off-diagonal parameters, and a simple calculation shows that the generator matrix has up to $3^d - 1$ nonzero parameters. Considering that the maximum index value for implementations of statically typed programming languages is often one of the large unsigned integer values $2^{16} - 1$, $2^{32} - 1$, or $2^{64} - 1$, the difficulties of implementing such a model in high dimensions become evident.

A large number of parameters also poses a methodological issue, reducing model interpretability and risking over-parametrization. A specialized simulation algorithm for exchangeable Marshall–Olkin distributions, introduced in [15], addresses this problem by reducing the number of parameters to the dimension. Another specialized simulation algorithm for certain low-parametric subclasses of exchangeable Marshall–Olkin distributions was suggested in [13]. The focus on exchangeable distributions is supported by the hierarchical models presented in [14, Sec. 5.2] and [16, Sec. 4.2], which use the exchangeable subclass as building blocks. However, [15] and [13] leave significant practical issues unaddressed. The first algorithm requires the on-the-fly calculation of submodel parameters, having numerical issues in higher dimensions similar to those discussed in Section 5. The second algorithm requires the simulation of Lévy processes; this can only be done without bias for the compound Poisson case and requires specialized sampling algorithms for the corresponding jump distribution. In particular, both algorithms have limited applicability to extendible Marshall–Olkin distributions in higher dimensions without further research or specialized implementations of the Lévy processes' jump distributions.

We asked ourselves: *Can we find low-parametric Markov-based models for extendible Marshall–Olkin distributions with a numerically stable implementation?*

This article is structured as follows: Section 2 recalls prerequisites about the Marshall–Olkin, exchangeable Marshall–Olkin, and extendible Marshall–Olkin distributions, including Bernstein functions. In particular, we present existing simulation algorithms and provide several extendible parametric subfamilies. Following, we provide an alternative proof that the death-set process of Marshall–Olkin distributed random vectors is Markov in Section 3, which allows identifying the corresponding Markov generator matrix. In Section 4, we derive the first primary result of this article. It proves that the death-counting process of exchangeable Marshall–Olkin distributed random vectors is Markov. Furthermore, it shows that we can construct an exchangeable Marshall–Olkin distributed random vector from a sample of the corresponding death-counting process by applying an independent random shuffling. Section 5 discusses the second primary result, a numerically stable calculation of the death-counting process' Markov generator Matrix of extendible Marshall–Olkin distributions using integral representations. Finally, before concluding this article, Section 6 benchmarks the runtime of various simulation algorithms for extendible Marshall–Olkin distributions against the algorithm using our new Markov death-counting model.

2 Background

This section recalls the required background information on the Marshall–Olkin distribution. In particular, we discuss the exchangeable and extendible subclass and compare existing stochastic models.

We assume the reader is familiar with stochastic vectors and processes, particularly exchangeability and Markov processes. We also presume some familiarity with the Marshall–Olkin distribution, and we will only discuss the material required in later sections for deriving further results in the following. For the interested reader, we refer to [17] for a detailed discussion of random vectors and Marshall–Olkin distributions, [4] for Markov processes, and [1] for exchangeable random variables. For Poisson and Lévy processes, we refer to [25] and [27], respectively. Finally, we refer to [9] for numerical mathematics and numerical integration. Throughout this article, we use Δ to denote the (*finite*) *forward difference* on a sequence or a

function, i.e., $\Delta x_n := x_{n+1} - x_n$ and $\Delta f(x) = f(x+1) - f(x)$, and Δ^i to denote the (finite) forward differences of the order i.

2.1 Marshall–Olkin distributions

The exponential distribution takes a central position among positive, continuous univariate distributions due to its *lack of memory* (*LOM*) *property*: As a consequence of Cauchy's exponential functional equation, the exponential distribution is the unique continuous distribution on the nonnegative half-line for which knowledge of survival up to a given time does not change the law of the *excess time distribution*. In mathematical notation, we write:

$$\mathbb{P}(\tau > t + s | \tau > t) = \mathbb{P}(\tau > s), \quad t, s \geq 0. \tag{1}$$

This property has a fundamental parallel: the *death-indicator process*, defined by $t \in [0, \infty) \mapsto \mathbb{1}_{\{\tau \leq t\}}$, is a (homogeneous) Markov process in the finite state space $\{0, 1\}$. Hence, the LOM property corresponds to the distribution of the time-to-death being independent of the already elapsed time.

The Marshall–Olkin (MO) distribution was introduced in [19] as the unique multivariate exponential distribution whose margins fulfill the following generalization of the univariate LOM property: Given the joint survival until a particular time, the excess time distribution equals the original distribution. In mathematical notation, we write

$$\mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I + t | \boldsymbol{\tau}_I > t) = \mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I), \quad \boldsymbol{s}_I := (s_i : i \in I), \quad t \geq 0, \quad I \subseteq [d] := \{1, \ldots, d\}. \tag{2}$$

It was shown in [5] that this property also has a parallel: the marginal death-set processes of MO distributed random vectors $\boldsymbol{\tau}$, defined for $\varnothing \neq I \subseteq [d]$ by $t \in [0, \infty) \mapsto \{i \in I : \tau_i \leq t\}$, are (homogeneous) Markov processes.

MO distributions have for $2^d - 1$ nonnegative parameters $\boldsymbol{\lambda} = \{\lambda_I : \varnothing \neq I \subseteq [d]\}$, subsequently called *shock-arrival intensities*, the survival function

$$\mathbb{P}(\boldsymbol{\tau} > \boldsymbol{t}) = \exp\left\{ - \sum_{\varnothing \neq I \subseteq [d]} \lambda_I \cdot \max_{i \in I} t_i \right\}, \quad \boldsymbol{t} \geq 0,$$

where $\sum_{I : I \ni i} \lambda_I > 0 \ \forall i \in [d]$. The latter condition ensures that all components of $\boldsymbol{\tau}$ are almost surely finite and can be relaxed if this is not required.

In the following, we present three stochastic models for MO distributions. Our intention for this is twofold: First, we want to use these representations to derive different results and properties. Second, we want to implement these models for the runtime comparison in Section 6.

The exogenous shock model

A natural stochastic model for MO distributions known as the exogenous shock model (ESM) was proposed in [19]. It assumes independent exponential random variables, called *shock-arrival times*, with respective shock-arrival intensities. Subsequently, it defines the elements of the random vector as the respective minima of component-related shock-arrival times.

Theorem 2.1. [19] *Let $\{E_I : \varnothing \neq I \subseteq [d]\}$ be independent exponential random variables with nonnegative rates $\boldsymbol{\lambda} = \{\lambda_I : \varnothing \neq I \subseteq [d]\}$ and define the vector $\boldsymbol{\tau}$ by*

$$\tau_i := \min\{E_I : I \ni i\}, \quad i \in [d]. \tag{3}$$

Then, $\boldsymbol{\tau}$ has an MO distribution with shock-arrival intensities $\boldsymbol{\lambda}$.

The ESM is a straightforward stochastic model; theoretical aspects of its implementation are discussed in [17, Algo. 3.1]. The aforementioned reference makes two critical findings: First, the index sets $\varnothing \neq I \subseteq [d]$ have the binary representation equal to $(\mathbb{1}_{\{i \in I\}}, i \in [d]) \in \{0, 1\}^d$. Second, the total number of shock-arrival times, equal to the number of parameters $2^d - 1$, grows exponentially and quickly becomes an issue for every implementation.

The Arnold model

Another stochastic model for MO distributions, subsequently called the Arnold model (AM), was derived in [2]. It defines the random vector $\boldsymbol{\tau}$ as first hitting times onto the components of a sequence of iid set-valued random variables with iid exponentially distributed inter-arrival times. The shock-arrival probabilities are proportional to the shock-arrival intensities, and the inter-arrival rate equals the sum over all shock-arrival intensities.

Theorem 2.2. [2, Sec. 4] and [17, Lem. 3.4] *Let $\{W_j : j \in \mathbb{N}\}$ be iid exponential random variables with rate $\lambda > 0$ and, independent thereof, $\{Y_j : j \in \mathbb{N}\}$ be iid discrete random variables with values in the power set of $[d]$ and probability vector $\boldsymbol{p} = \{p_I : \varnothing \neq I \subseteq [d]\}$. Furthermore, define the vector $\boldsymbol{\tau}$ by*

$$\tau_i \coloneqq \sum_{j=1}^{\min\{k \in \mathbb{N} : Y_k \ni i\}} W_j, \quad i \in [d]. \tag{4}$$

Then, $\boldsymbol{\tau}$ has an MO distribution with shock-arrival intensities $\boldsymbol{\lambda} \coloneqq \lambda \cdot \boldsymbol{p}$.

An interesting aspect of this model proven in [2, Sec. 4] is that the ESM is recovered as follows: Assume that $\boldsymbol{\tau}$ has the representation in Eq. (4) and define

$$E_I \coloneqq \sum_{j=1}^{\min\{k \in \mathbb{N} : I = Y_k\}} W_j, \quad \varnothing \neq I \subseteq [d].$$

Then, $\{E_I : \varnothing \neq I \subseteq [d]\}$ are independent with rates $\{\lambda_I : \varnothing \neq I \subseteq [d]\}$ and $\boldsymbol{\tau}$ has the ESM representation from Eq. (3). Thus, the AM samples the shock-arrival times of the ESM in ascending order by sampling their inter-arrival times and corresponding affected components until all components are dead. However, two distinctions between both models are crucial for their comparison: First, the AM allows us to ignore all shocks arriving after the time when all components are dead, requiring potentially considerably less than $2^d - 1$ shocks. Second, the AM can have *looping steps*, whose shocks Y_j contain only already dead elements. The looping steps have no counterpart in the ESM.

Theoretical aspects of this model's implementation are discussed in [17, Algo. 3.3]. A nontrivial choice of any implementation is the sampling method for iid sequences of discrete laws. The aforementioned reference proposes a version of the probability integral transform for discrete random variables with a divide-and-conquer search for the (left) inversion of the probability distribution function. We suggest considering the *alias method* proposed by [31] since, after an initial parameter transformation, it only requires a uniform discrete sample, a uniform real sample, two reference operations, and a comparison per sample; see also Appendix E.

The Markov death-set model

Finally, a last stochastic model for MO distributions is the Markov death-set model (MDSM). For every death-set process $t \in [0, \infty) \mapsto Z_t \coloneqq \{i \in [d] : \tau_i \leq t\}$, we have

$$\tau_i = \inf\{t \geq 0 : i \in Z_t\}, \quad i \in [d].$$

Recall that the death-set process of an MO distributed random vector is Markov. Consequently, given the Markov generator matrix $Q \in \mathbb{R}^{2^d \times 2^d}$ of Z, where the row indices correspond to the binary representation of the states, we can directly sample Z using models for finite-state Markov processes as follows.

Theorem 2.3. [4, p. 493] *Let $Q = (q_{ij})_{i,j\in[n]} \in \mathbb{R}^{n\times n}$ be a Markov generator matrix and let $q_i := -q_{ii}$ for $i \in [n]$. Define the transition matrix $K = (k_{ij})_{i,j\in[n]}$ as follows*:

$$k_{ij} = \begin{cases} \dfrac{q_{ij}}{q_i} & i \neq j, \\ 0 & i = j \quad \textit{and} \quad q_i > 0, \\ 1 & \textit{else}. \end{cases}$$

Let $\tilde{Z}$ be a discrete Markov chain with transition matrix K and the inter-arrival times of the transitions $\{W_k : k \in \mathbb{N}\}$, conditioned on $\tilde{Z}$, be independent exponentially distributed with rates $\{q_{\tilde{Z}_{k-1}} : k \in \mathbb{N}\}$. Then, Z, defined in the following, is a continuous Markov process with generator matrix Q:

$$Z_t = \tilde{Z}_{\max\left\{k\in\mathbb{N}_0 : \sum_{l=1}^{k} W_l \leq t\right\}}, \quad t \in [0, \infty).$$

We call $\tilde{Z}$ the embedded Markov chain of Z.

Given death-set process Z has the representation from Theorem 2.3 with $n = 2^d$, we conclude that,

$$\tau_i = \inf\{t \geq 0 : i \in Z_t\} = \sum_{j=1}^{\min\{k\in\mathbb{N} : \tilde{Z}_k \ni i\}} W_j, \quad i \in [d].$$

However, to use this algorithm to sample from MO distributions, we require the mapping $\boldsymbol{\lambda} \to Q$, which is not described in the literature to the best of our knowledge.

2.2 Exchangeable Marshall–Olkin distributions

The general MO distribution, as previously discussed, has the problem of an exponentially increasing number of parameters. We can alleviate the issue by restricting ourselves to elements of the exchangeable subclass, whose number of parameters is equal to the dimension. An MO distribution is exchangeable if and only if its shock-arrival intensities are equal among those that represent shocks of the same cardinality, see [14, Lem. 3.1.1], i.e.,

$$|I| = |J| \;\Leftrightarrow\; \lambda_I = \lambda_J \quad \forall \varnothing \neq I,\ J \subseteq [d].$$

For exchangeable Marshall–Olkin (exMO) distributions, we write for their *exchangeable shock-arrival intensities* $\lambda_i \equiv \lambda_I$, $i = |I|$, and define the *exchangeable shock-size-arrival intensities* $\boldsymbol{\eta} = \{\binom{d}{i}\lambda_i : i \in [d]\}$, i.e.,

$$\eta_i = \sum_{\varnothing\neq I\subseteq[d] : |I|=i} \lambda_I, \quad i \in [d],$$

is the rate of $\min\{E_I : |I| = i\}$, the smallest *shock-arrival time* E_I of all shocks I with cardinality i.

2.3 Extendible Marshall–Olkin distributions

A restriction to the exchangeable subclass dramatically reduces the number of parameters from $2^d - 1$ to d. However, the number is still sizeable for higher dimensions. For this reason, many examples of high-dimensional MO distributions are based on low-parametric, extendible subfamilies. Extendible subfamilies of the MO distribution have a stochastic representation as the finite margin of a countably infinite exchangeable sequence with MO-distributed margins. We will not discuss the details of extendible

Marshall–Olkin (extMO) distributions and exchangeable sequences in this article and refer to [14] and [1] for the details on these; however, we require some results and will recall them in the following.

Before discussing extMO distributions, we have to make a short excursus into *Bernstein functions*. A *Bernstein function* is a nonnegative, nondecreasing function $\psi : [0, \infty) \to [0, \infty)$ that has infinitely many derivatives with alternating signs. An extensive monograph on those functions is [28]. We highlight four important properties:

- Every Bernstein function ψ is uniquely linked to a *Lévy–Khintchine triplet* (a, b, ν) such that
$$\psi(x) = a + bx + \int_0^\infty (1 - e^{-ux})\nu(\mathrm{d}u), \quad x \geq 0, \tag{5}$$
where $a, b \geq 0$ and ν is a *Lévy measure* on $(0, \infty)$, i.e., $\int_0^\infty (1 \wedge u)\mathrm{d}u < \infty$, see [28, Thm. 3.2].
- In combination with [3, Prop. 6.12], the previous property implies that a nonnegative function $\psi : [0, \infty) \to [0, \infty)$ is a Bernstein function if and only if it is nondecreasing and the (negative) finite forward differences are of alternating signs, i.e.,
$$(-1)^{i-1}\Delta^i\psi(x) \geq 0 \quad \forall x \geq 0, \quad i \in \mathbb{N}.$$
- A Bernstein function whose *Lévy measure* has a completely monotone density with respect to Lebesgue measure, i.e., the density is nonnegative, nonincreasing with infinitely many derivatives of alternating sign, is called a *complete Bernstein function*. Every complete Bernstein function ψ is uniquely linked to a *Stieltjes triplet* (a, b, σ) such that
$$\psi(x) = a + bx + \int_0^\infty \frac{x}{x + u}\sigma(\mathrm{d}u), \quad x \geq 0, \tag{6}$$
where $a, b \geq 0$ and σ is a *Stieltjes measure* on $(0, \infty)$, i.e., $\int_0^\infty (1 + u)^{-1}\sigma(\mathrm{d}u) < \infty$, see [28, Thm. 6.2].
- Bernstein functions are the *Laplace exponents* of (*killed*) *Lévy subordinators*, see [28, Chp. 5]: For each Bernstein function ψ, we can create a probability space supporting a nonnegative, nondecreasing (killed) Lévy process Λ on $[0, \infty]$ such that
$$\mathbb{E}[e^{-x\Lambda_t}] = e^{-t\psi(x)}, \quad x, t \geq 0. \tag{7}$$
Conversely, if Λ is a nonnegative, nondecreasing Lévy process, we can find a Bernstein function ψ such that Eq. (7) holds. Recall that Lévy subordinators are nondecreasing, stochastically continuous, càdlàg processes that start in zero and have stationary and independent increments. For more details, we refer to [27].

Remark. [27, cf. Chp. 4] and [28, Chp. 5] For a Lévy–Khintchine triplet (a, b, ν), we have the representation
$$\Lambda_t = \infty \cdot \mathbb{1}_{\{\varepsilon \leq a \cdot t\}} + b \cdot t + J_t,$$
where ε is a unit exponential random variable and J is a pure-jump Lévy process, independent of ε, with Lévy measure ν. If ν is finite, J is a compound Poisson process with intensity $\nu((0, \infty))$ and jump distribution $B \mapsto \nu(B)/\nu((0, \infty))$. Hence, we call a the *killing rate*, b the *drift*, and, if ν is finite, $\nu((0, \infty))$ the *jump rate*, and $B \mapsto \nu(B)/\nu((0, \infty))$ the *jump distribution*.

The following stochastic model for extMO distributions, called the Lévy frailty model (LFM), outlines the relationship between extMO distributions and Bernstein functions:

Theorem 2.4. [13, Thm. 3.3] and [14, Thm. 3.4.1] *There exists a bijection between the distributions of extMO sequences and Bernstein functions. Moreover, given a Lévy subordinator Λ with Bernstein function ψ and, independent thereof, unit exponential barrier values $\{E_i : i \in [d]\}$, we can define*
$$\tau_i := \inf\{t \geq 0 : \Lambda_t \geq E_i\}, \quad i \in [d]. \tag{8}$$

Then, $(\tau_1, \ldots, \tau_d)^\top$ *has an extMO distribution with exchangeable shock-arrival intensities*

$$\lambda_i = (-1)^{i-1}\Delta^i\psi(d-i), \quad i \in [d]. \tag{9}$$

The nontrivial mapping from Bernstein functions to exchangeable shock-arrival intensities in Eq. (9) already indicates that the parametrization via Bernstein functions, albeit methodologically convenient, imposes challenges on the implementation.

2.4 Examples of parametrized families

This section introduces selected Bernstein function families that we implemented for our simulation studies. An extensive list of additional Bernstein functions is compiled in [28, Chp. 16]; whenever possible, we provide the corresponding number of Bernstein functions in that list.

- *Armageddon shock*: Let $\alpha, \beta \geq 0$ and consider the Lévy triplet $(\beta, \alpha, 0)$ corresponding to the Bernstein function $\psi(x) = \beta + \alpha x$, $x \geq 0$. The name reflects that the corresponding ESM representation has only individual shocks, arriving with rate α, and a global shock, killing all components, arriving with rate β.
- *Poisson*: Let $\eta > 0$ and consider the Lévy triplet $(0, 0, \delta_\eta)$ corresponding to the Bernstein function $\psi(x) = (1 - e^{-\eta x})$, $x \geq 0$, where $B \mapsto \delta_\eta(B)$ is the Dirac measure. This family corresponds to Lévy subordinators that are compound Poisson processes with intensity 1 and deterministic jump size η.
- *Exponential* ([28, Chp. 16, No. 4]): Let $\eta > 0$ and consider the Lévy triplet $(0, 0, \nu)$ with a completely monotone Lévy density $\nu(\mathrm{d}u) = \eta \cdot e^{-\eta u}\,\mathrm{d}u$ and Stieltjes measure $\sigma = \delta_\eta$ corresponding to the Bernstein function $\psi(x) = \frac{x}{x+\eta}$, $x \geq 0$. This family corresponds to Lévy subordinators that are compound Poisson processes with intensity 1 and exponential jumps with rate η.
- *Pareto*: Let $\alpha \in (0, 1)$, $x_0 > 0$ and consider the Lévy triplet $(0, 0, \nu)$ with Lévy density $\nu(\mathrm{d}u) = \alpha x_0^\alpha \cdot u^{-\alpha-1}\mathbb{1}_{\{u>x_0\}}\,\mathrm{d}u$. This family corresponds to compound Poisson processes with intensity 1 and jumps from a Pareto distribution. The Pareto family is used in [7, Sec. 5.3] to approximate the subsequently defined *Alpha-stable family*.
- *Alpha-stable* ([28, Chp. 16, No. 1]): Let $\alpha \in (0, 1)$ and consider the Lévy triplet $(0, 0, \nu)$ with a completely monotone Lévy density $\nu(\mathrm{d}u) = \alpha[\Gamma(1-\alpha)]^{-1} \cdot u^{-1-\alpha}\mathrm{d}u$ and Stieltjes density $\sigma(\mathrm{d}u) = \pi^{-1}\sin(\alpha\pi) \cdot u^{\alpha-1}\mathrm{d}u$ corresponding to the Bernstein function $\psi(x) = x^\alpha$, $x \geq 0$. This family corresponds to Lévy subordinators with infinite activity.
- *Gamma* ([28, Chp. 16, No. 26]): Let $a > 0$ and consider the Lévy triplet $(0, 0, \nu)$ with a completely monotone Lévy density $\nu(\mathrm{d}u) = e^{-au}u^{-1}\mathrm{d}u$ and Stieltjes density $\sigma(\mathrm{d}u) = u^{-1}\mathbb{1}_{(a,\infty)}(u)\mathrm{d}u$ corresponding to the Bernstein function $\psi(x) = \log(1 + x/a)$, $x \geq 0$. This family corresponds to Lévy subordinators with infinite activity.
- *Inverse Gaussian* ([17, p. 309] and [28, Chp. 16, No. 2]): Let $\eta \geq 0$ and consider the Lévy triplet $(0, 0, \nu)$ with a completely monotone Lévy density $\nu(\mathrm{d}u) = [\sqrt{2\pi u^3}\,]^{-1}e^{-\frac{1}{2}\eta^2 u}\,\mathrm{d}u$ and Stieltjes density $\sigma(\mathrm{d}u) = \sin(\pi/2)[\pi u]^{-1}\sqrt{2u - \eta^2}\,\mathbb{1}_{(\eta^2/2,\infty)}(u)\mathrm{d}u$ corresponding to the Bernstein function $\psi(x) = \sqrt{2x + \eta^2} - \eta$, $x \geq 0$. This family corresponds to Lévy subordinators with infinite activity.

It can be practical to normalize a Bernstein function ψ to ψ^* such that $\psi^*(1) = 1$: we can do that in multiple ways, for example, by

- *adding a constant part* with $\psi^*(x) = [1 - \psi(1)] + \psi(x)$, $x > 0$,
- *adding a linear part* with $\psi^*(x) = [1 - \psi(1)]x + \psi(x)$, $x > 0$, or
- *scaling* with $\psi^*(x) = \psi(1)^{-1} \cdot \psi(x)$, $x > 0$.

Note that the first two normalization methods require $\psi(1) \leq 1$.

2.5 Hierarchical Marshall–Olkin distributions

We want to conclude the background section with a short note on hierarchical Marshall–Olkin (hMO) distributions. There are multiple approaches to define nonexchangeable hMO distributions with exMO distributions as building blocks, e.g., [14, Sec. 5.2] or [16, Sec. 4.2]. An approach similar to the latter has been proposed in [30, Sec. 4.1]. We do not discuss these representations further in this article. Still, we want to highlight a consequence of their existence: *efficient sampling algorithms for hMO distributions can be built using efficient sampling algorithms for extMO distributions as building blocks*. For this reason, we decided to place our primary focus on the development of such an algorithm for the extendible subclass.

3 The Markovian MO death-set model

Similar to [5], the following section proves that the death-set processes Z of MO distributions are Markovian. In addition, we also provide a formula for the infinitesimal generator matrix Q. This explicit formula is used to analyze the slightly more complicated death-counting process of the exchangeable subclass in Section 4.

Theorem 3.1. *Let $d \geq 2$.*

(*a*) *Let Z be the death-set process of a d-dimensional MO-distributed random vector $\boldsymbol{\tau}$ with shock-arrival intensities $\boldsymbol{\lambda}$. Then, Z is a continuous-time, homogeneous Markov process on the power set of $[d]$ with infinitesimal generator matrix $Q = (q_{IJ} : I, J \subseteq [d])$ defined by*

$$q_{IJ} \coloneqq \begin{cases} -\sum\limits_{I \subsetneq K \subseteq [d]} \sum\limits_{L \subseteq I} \lambda_{L \cup (K \setminus I)} & I = J, \\ \sum\limits_{L \subseteq I} \lambda_{L \cup (J \setminus I)} & I \subsetneq J, \\ 0 & \text{else}. \end{cases} \tag{10}$$

(*b*) *Let Z be a continuous-time, homogeneous Markov process on the power set of $[d]$ with infinitesimal generator matrix Q as in Eq. (10) and define the random vector $\boldsymbol{\tau}$ as follows*:

$$\tau_i \coloneqq \inf\{t > 0 : i \in Z_t\}, \quad i \in [d]. \tag{11}$$

Then, $\boldsymbol{\tau}$ is MO distributed with shock-arrival intensities $\boldsymbol{\lambda}$, and Z is the death-set process of $\boldsymbol{\tau}$.

The infinitesimal generator matrix Q in Eq. (10) often takes a simpler form for MO distribution subclasses. Appendix A simplifies the infinitesimal generator Eq. (10) for exMO distributions, extMO distributions, and the Armageddon-shock family.

Nevertheless, the representation in Eq. (10) imposes enormous challenges for possible implementations of the MDSM: the implementation either stores the generator matrix in memory or calculates its entries on-the-fly during the simulation. We argue that both options are hardly feasible in higher dimensions: On the one hand, a straightforward but tedious calculation yields up to $3^d - 1$ nonzero entries of Q. Consequently, storing this matrix is nontrivial for larger d, even using sparse matrix designs. On the other hand, using on-the-fly calculations requires repeatedly populating large vectors and prevents utilizing optimizations for discrete sampling that require nontrivial setup activities.

Before proving the result, we highlight that the MDSM does not translate well into a simulation algorithm in higher dimensions for two reasons: First, implementing the model becomes challenging because of the exponentially growing number of parameters. Second, nontrivial MO distributions are often specified via a Bernstein function, using Eq. (9), often from a low-parametric family. Consequently, higher-order iterated differences and large sums make calculating the generator matrix numerically challenging, compared with Section 5.

The proof of Theorem 3.1 is separated into three steps: First, we identify the shock-arrival times of the AM with a Poisson process. Second, we show that the sequence of cumulative unions of shock sets from the AM is a random walk on the semigroup of the power set of $[d]$ with the union as a conjunction. Third, we conclude that this makes the death-set indicator process a random walk subordinated by a Poisson process; hence, it is also a Markov process. This representation makes it simple to derive the corresponding infinitesimal generator and embedded transition matrix.

Lemma 3.2. [25, Ex. 3.3.7] *Let $W_1, W_2, \dots$ be iid exponentially distributed with rate $\lambda > 0$. Then, the set of their cumulative sums $\Gamma_1, \Gamma_2, \dots$, defined by*

$$\Gamma_j := \sum_{k=1}^{j} W_k, \quad j \in \mathbb{N},$$

is a (homogeneous) Poisson point process with intensity λ and defines the (homogeneous) Poisson process N with intensity λ by

$$N : [0, \infty) \to \mathbb{N}_0, \quad t \mapsto \sum_{j=1}^{\infty} \mathbb{1}_{\{\Gamma_j \le t\}}.$$

Lemma 3.3. *Let $d \ge 2$, $\boldsymbol{p} = \{p_I : I \subseteq [d]\}$ be a probability vector for the iid sequence $Y_1, Y_2, \dots$ on the power set of $[d]$, and define the discrete-time process*

$$\tilde{Z} : \mathbb{N}_0 \to \mathcal{P}([d]) := \{I : I \subseteq [d]\}, \quad n \mapsto \bigcup_{j=1}^{n} Y_j,$$

where $\bigcup_{j=1}^{0} Y_j := \varnothing$. Then, $\tilde{Z}$ is a discrete-time Markov chain with transition matrix $K = (k_{IJ} : I, J \subseteq [d])$ defined by

$$k_{IJ} := \begin{cases} \sum_{L \subseteq I} p_{L \cup (J \setminus I)} & I \subseteq J, \\ 0 & \text{else}. \end{cases} \tag{12}$$

Proof. Note that $\tilde{Z}$ is a random walk on the finite semigroup $(\mathcal{P}([d]), \cup)$ and consequently also a Markov chain. Let $k \in \mathbb{N}$ and $I, J \subseteq [d]$. First, since $\tilde{Z}_0, \tilde{Z}_1, \dots$ is a nondecreasing sequence of (random) sets, we have

$$\mathbb{P}(\tilde{Z}_{k+1} = J | \tilde{Z}_k = I) = 0,$$

whenever $I \cap J \neq I$. Second, we have for $I \subseteq J$ that

$$\mathbb{P}(\tilde{Z}_{k+1} = J | \tilde{Z}_k = I) = \mathbb{P}(Y_{k+1} \setminus I = J \setminus I) = \mathbb{P}(Y_1 \setminus I = J \setminus I) = \sum_{L \subseteq [d] : L \setminus I = J \setminus I} p_L = \sum_{L \subseteq I} p_{L \cup (J \setminus I)}. \qquad \square$$

Lemma 3.4. *Let $d \ge 2$, N be a (homogeneous) Poisson process with intensity $\lambda > 0$ and, independent thereof, let $\tilde{Z}$ be the Markov chain from* Lemma 3.3 *for probability vector $\boldsymbol{p}$. Define the continuous-time process*

$$Z : [0, \infty) \to \mathcal{P}([d]), \quad t \mapsto \bigcup_{j=1}^{N_t} Y_j.$$

Then, Z is a continuous-time Markov process with infinitesimal generator matrix $Q = (q_{IJ} : I, J \subseteq [d])$ as in Eq. (10) *with $\boldsymbol{\lambda} = \lambda \cdot \boldsymbol{p}$.*

Proof. Note that, due to Lemma 3.3, $\tilde{Z}$ is a Markov chain with transition matrix $K = (k_{IJ} : I, J \subseteq [d])$ defined in Eq. (12). Hence, Z represents a discrete-time Markov chain with transition matrix K, subordinated by a homogeneous Poisson process with intensity λ. We conclude with [4, Expl. 13.2.8] that Z is a Markov process with the infinitesimal generator matrix $Q = \lambda \cdot (K - \mathbb{1})$, where $\mathbb{1}$ is the suitable matrix identity. While the second and third cases of Eq. (10) follow trivially, we obtain the first case from the following calculation:

$$q_{II} = \lambda \cdot (k_{II} - 1) = -\lambda \cdot (1 - k_{II}) = -\lambda \cdot \left[\sum_{I \subseteq L \subseteq [d]} k_{IL} - k_{II} \right] = -\lambda \cdot \sum_{I \subsetneq L \subseteq [d]} k_{IL} = - \sum_{I \subsetneq K \subseteq [d]} \sum_{L \subseteq I} \lambda_{L \cup (K \setminus I)}. \qquad \square$$

Proof of Theorem 3.1. We can assume for both parts without loss of generality that $\boldsymbol{\tau}$ and Z are defined by the AM, compared with Theorem 2.2.

First, for Theorem 3.1(a), the claim becomes a trivial corollary of Lemma 3.4 after considering the AM representation of $\boldsymbol{\tau}$ and subsequently Z.

Second, for Theorem 3.1(b), also considering the AM representation of Z, we have the following calculation:

$$\begin{aligned} \tau_i &= \inf\{t > 0 : i \in Z_t\} = \inf\left\{ t > 0 : i \in \bigcup_{j=1}^{N_t} Y_j \right\} = \inf\left\{ t > 0 : \sum_{j=1}^{N_t} \mathbb{1}_{\{i \in Y_j\}} > 0 \right\} \\ &= \inf\left\{ t > 0 : \sum_{j=1}^{\infty} \mathbb{1}_{\{i \in Y_j, \quad W_1 + \cdots + W_j \le t\}} > 0 \right\} = \inf\left\{ t > 0 : \sum_{j=1}^{\min\{k \in \mathbb{N} : i \in Y_k\}} W_j \le t \right\} = \sum_{j=1}^{\min\{k \in \mathbb{N} : i \in Y_k\}} W_j. \end{aligned}$$

We obtain the claim by comparing the aforementioned equation with Eq. (4). $\square$

4 The Markovian exMO death-counting model

This section shows that the *death-counting process* of exMO distributions inherits the Markov property from the death-set process introduced in Section 3. Furthermore, we derive the Markov death-counting model (MDCM) for exMO distributions. This model requires significantly fewer parameters than the MDSM.

We define the *death-counting process* of a d-variate random vector $\boldsymbol{\tau}$ with death-set process Z by

$$Z^* : [0, \infty) \to [d]_0 := [d] \cup \{0\}, \quad t \mapsto |Z_t| = |\{i : \tau_i \le t\}|.$$

In the case of MO distributions, we again assume an AM representation. Hence, we have

$$Z_t^* = \left| \bigcup_{j=1}^{N_t} Y_j \right|, \quad t \ge 0.$$

The following result shows that the death-counting process of exMO distributions is Markov. Moreover, it shows that exMO distributions are represented by shuffled transition times of suitable Markovian death-counting processes.

Theorem 4.1. *Let $d \ge 2$.*

(*a*) *Let Z^* be the death-counting process of a d-dimensional exMO distributed random vector with exchangeable shock-arrival intensities $\boldsymbol{\lambda}$. Then, Z^* is a continuous-time, homogeneous Markov process on $[d]_0$ with infinitesimal generator matrix $Q^* = (q_{ij}^* : i, j \in [d]_0)$ defined by*

$$q_{ij}^* = \begin{cases} -\sum_{l=1}^{d-i} \binom{d-i}{l} \sum_{k=0}^{i} \binom{i}{k} \lambda_{k+l} & i = j, \\ \binom{d-i}{j-i} \sum_{k=0}^{i} \binom{i}{k} \lambda_{k+(j-i)} & i < j, \\ 0 & \text{else}. \end{cases} \tag{13}$$

(*b*) *Let Z^* be a continuous-time, homogeneous Markov process on $[d]_0$ with infinitesimal generator matrix Q^* as in Eq.* (13). *Furthermore, let Π, independent thereof, be a uniform random permutation on $[d]$ and define the random vector $\boldsymbol{\tau}$ by*

$$\tau_i := \inf\{t > 0 : \Pi(i) \le Z_t^*\}, \quad i \in [d]. \tag{14}$$

Then, $\boldsymbol{\tau}$ has an exMO distribution with exchangeable shock-arrival intensities $\boldsymbol{\lambda}$.

Again, the infinitesimal generator matrix Q^* in Eq. (13) often takes a simpler form for exMO distribution subclasses. Appendix B simplifies the infinitesimal generator Q^* for a reparametrization of exMO distributions, extMO distributions, and the Armageddon shock family.

The MDSM's generator matrix has 4^d entries and up to $3^d - 1$ nonzero entries. In contrast, the MDCM has only $(d+1)^2$ entries and up to $[d+1]d/2 + d$ nonzero entries. Two immediate consequences of this reduction of model parameters are as follows: First, the parameter number grows only quadratically instead of exponentially with the dimension. Second, the parameter number does not exceed technical maximums of integer binary representations for reasonable dimension sizes.

We also separate the proof of Theorem 4.1 into several incremental steps: First, we show that the sequence of death-count transitions from the AM is a discrete-time Markov chain. Second, we conclude that the death-count process is Markov and calculate its infinitesimal generator. Third, we use that the law of exchangeable random vector remains the same after ordering and subsequently shuffling its components uniform at random. Therefore, while we cannot recover the original random vector from the death-counting process alone, we can construct another random vector with the same law by applying a uniform random permutation.

Lemma 4.2. *Let $d \geq 2$, $\boldsymbol{p} = \{p_I : I \subseteq [d]\}$ be an exchangeable probability vector for the exchangeable iid sequence $Y_1, Y_2, \ldots$ on the power set of $[d]$, and define the discrete-time process*

$$\tilde{Z}^* : \mathbb{N}_0 \to [d]_0, \quad n \mapsto |\tilde{Z}_n| = \left| \bigcup_{j=1}^{n} Y_j \right|.$$

Then, $\tilde{Z}^*$ is a discrete-time Markov chain with transition matrix $K^* = (k^*_{ij} : i, j \in [d]_0)$ defined by

$$k^*_{ij} = \binom{d-i}{j-i} \cdot \begin{cases} \sum_{l=0}^{i} \binom{i}{l} p_{l+(j-i)} & i \leq j, \\ 0 & \text{else}. \end{cases} \tag{15}$$

Proof. We know from Lemma 3.3 that $\tilde{Z}$ is a Markov chain with transition matrix K from Eq. (12). We perform a simple calculation using exchangeability, grouping shock-arrival intensities associated with the same cardinality, to obtain

$$k_{IJ} = \sum_{l=0}^{i} \binom{i}{l} p_{l+(j-i)} =: k_{ij}, \quad I \subseteq J, \quad |I| = i, \quad |J| = j,$$

where we write $p_i \equiv p_I$ and denote $k_{ij} \equiv k_{IJ}$ for $|I| = i$ and $|J| = j$.

First, since $\tilde{Z}^*_0, \tilde{Z}^*_1, \ldots$ is a nondecreasing sequence of integer-valued random variables, we have for all sequences $i_1, \ldots, i_{n-1}, i, j$ that are not nondecreasing that

$$\mathbb{P}(\tilde{Z}^*_{n+1} = j | \tilde{Z}^*_n = i, \quad \tilde{Z}^*_l = i_l, \quad \forall l < n) = 0.$$

Second, for nondecreasing sequences of integer values $0 \leq i_1 \leq \cdots \leq i_{n-1} \leq i \leq j$, we denote the set of all nondecreasing set-sequences of the length $n+1$ with corresponding cardinalities $i_1, \ldots, i_{n-1}, i, j$ by $\mathcal{I}(i_1, \ldots, i_{n-1}, i, j)$ and have

$$\begin{aligned}
&\mathbb{P}(\tilde{Z}^*_{n+1} = j | \tilde{Z}^*_n = i, \tilde{Z}^*_l = i_l, \ \forall l < n) \\
&= \frac{\sum_{(I_1, \ldots, I_{n-1}, I, J) \in \mathcal{I}(i_1, \ldots, i_{n-1}, i, j)} \mathbb{P}(\tilde{Z}_{n+1} = J, \quad \tilde{Z}_n = I, \quad \tilde{Z}_l = I_l, \ \forall l < n)}{\sum_{(I_1, \ldots, I_{n-1}, I) \in \mathcal{I}(i_1, \ldots, i_{n-1}, i)} \mathbb{P}(\tilde{Z}_n = I, \quad \tilde{Z}_l = I_l, \ \forall l < n)} \\
&= \frac{\binom{d}{i_1}\binom{d-i_1}{i_2-i_1} \cdots \binom{d-i_{n-1}}{i_n-i_{n-1}}\binom{d-i}{j-i} \cdot k_{0\, i_1} k_{i_1 i_2} \cdots k_{i_{n-1} i} k_{ij}}{\binom{d}{i_1}\binom{d-i_1}{i_2-i_1} \cdots \binom{d-i_{n-1}}{i_n-i_{n-1}} \cdot k_{0\, i_1} k_{i_1 i_2} \cdots k_{i_{n-1} i}} \\
&= \binom{d-i}{j-i} \cdot k_{ij}.
\end{aligned}$$

□

Lemma 4.3. *Let $d \geq 2$, N be a (homogeneous) Poisson process with intensity $\lambda > 0$ and, independent thereof, let $\tilde{Z}^*$ be the Markov chain from* Lemma 4.2 *for an exchangeable probability vector $\boldsymbol{p}$. Define the continuous-time processes*

$$Z^* : [0, \infty) \to [d]_0, \quad t \mapsto \tilde{Z}^*_{N_t} = \left| \bigcup_{j=1}^{N_t} Y_j. \right|$$

Then, Z^* is a continuous-time Markov process with infinitesimal generator matrix $Q^* = (q^*_{ij} : i, j \in [d]_0)$ defined by Eq. (13) *with $\boldsymbol{\lambda} = \lambda \cdot \boldsymbol{p}$.*

Proof. We know from Lemma 4.2 that $\tilde{Z}^*$ is a Markov chain with transition matrix K^* from Eq. (15). Z^* thus represents a discrete-time Markov chain with transition matrix K^*, subordinated by a Poisson process with intensity λ. As in the proof of Lemma 3.4, we use [4, Expl. 13.2.8] to conclude that Z^* is a continuous-time Markov process with infinitesimal generator $Q^* = \lambda \cdot (K^* - \mathbb{1})$, where $\mathbb{1}$ is the suitable matrix identity. While the second and third cases of Eq. (13) follow trivially, we obtain the first case using the following calculation:

$$q^*_{ii} = \lambda \cdot (k^*_{ii} - 1) = -\lambda \left[\sum_{l=i}^{d} k^*_{il} - k^*_{ii} \right] = -\sum_{l=1}^{d-i} \binom{d-i}{l} \sum_{r=0}^{i} \binom{i}{r} \lambda_{r+l}. \qquad \square$$

Proof of Theorem 4.1. We assume without loss of generality that Z^* has a representation of the death-counting process for an exMO distributed random vector $\boldsymbol{\tau}$, which is generated by the AM (see Theorem 2.2). The corresponding death-set process is denoted by Z.

First, for Theorem 4.1(a), the claim becomes a trivial corollary of Lemma 4.3 after considering the AM representation of $\boldsymbol{\tau}$, Z, and Z^*.

Second, for Theorem 4.1(b), let $\hat{\Pi}$ be the (random) permutation of $[d]$ such that

$$\tau_{\hat{\Pi}(1)} \leq \cdots \leq \tau_{\hat{\Pi}(d)}.$$

Then, by using Theorem 3.1, we have

$$\begin{aligned} \tau_{\hat{\Pi}(i)} &= \inf\{t > 0 : \hat{\Pi}(i) \in Z_t\} \\ &= \inf\{t > 0 : \{\hat{\Pi}(j) : j \leq i\} \subseteq Z_t\} \\ &= \inf\{t > 0 : i \leq Z^*_t\}, \quad i \in [d]. \end{aligned}$$

Consequently, we obtain the claim with [17, Lem. 3.8]; the result used states that the law of an exchangeable random vector does not change under a (possibly dependent) reordering (in this case $\hat{\Pi}$) and a subsequent reordering according to an independent uniform permutation (e.g., Π). $\square$

Finally, based on Theorem 4.1, we can sketch a simple simulation algorithm for exMO distributions using the Markov process representation from Theorem 2.3, see Algorithm 1. The algorithm requires simulation algorithms for exponential and (finite-spaced) discrete random variables and a random permutator.

Algorithm 1. Sample from exMO distributions with the MDCM (see Theorem 2.3 and 4.1).

Require: (Vector) arguments `rates` and `probs` with state-dependent parameters for the sampling algorithms `sample_exponential` and `sample_discrete`, which sample from an exponential or discrete distribution (taking values $0, 1, \ldots, k-1$ for probabilities $p_0, p_1, \ldots, p_{k-1}$ and some $k \in \mathbb{N}$). Additionally, a method `shuffle`, which permutates vectors uniform at random.

```
1 : procedure SAMPLE_MDCM(d, rates, probs)
2 :   time = 0
3 :   i = 0
4 :   x = zeros(d)
```

```
    while i < d do
      i0 = i
      time += sample_exponential(rates[i])          ▷ waiting time
      i += 1 + sample_discrete(probs[i])            ▷ additional default-count
      for i0 ≤ j < d do
        x[j] = time
      end for
    end while
    shuffle(x)
    return x
end procedure
```

Remark. [15, Algo. 1] proposed a similar model: First, sample the order-statistic of $\boldsymbol{\tau}$ by recursively sampling shock sizes of marginal models using the representation from Theorem 2.2. Second, apply a permutation that is uniform at random. A close inspection yields that both representations' transition rates and probabilities are equal. From an implementation point of view, corresponding algorithms of both models differ in whether transition rates and probabilities are calculated once or repeatedly on-the-fly during simulation. Note, in particular, that this representation shares all numerical issues of the MDCM regarding mapping the input parameters to simulation parameters, which is discussed in Section 5.

5 Numerically stable approximation of extMO generators

Section 4 derived a novel stochastic representation for exMO distributions based on shuffling *Markovian* death-counting processes' transition times uniformly at random. However, for high-dimensional applications, the possible distributions are often restricted to one or a combination of multiple families of extMO distributions. Hence, we need to calculate the infinitesimal generator matrices of extMO distributions' death-counting processes, denoted by Q^*, to use the aforementioned representation for high-dimensional sampling. This section aims to answer the following question: *How can we approximate Q^* efficiently and numerically stable for a given Bernstein function ψ and dimension $d \geq 2$ to a satisfying accuracy*?

Consider an extMO distribution with Bernstein function ψ and dimension $d \geq 2$. Recall that the exchangeable shock-arrival intensities fulfill

$$\lambda_i = (-1)^{i-1}\Delta^i\psi(d-i), \quad i \in [d]. \tag{9 revisited}$$

Furthermore, we can deduce from Eq. (13) by a tedious but straightforward calculation, see Remark B.2, that

$$q^*_{ij} = \binom{d-i}{j-i} \cdot \begin{cases} -\psi(d-i) & i = j, \\ (-1)^{j-i-1}\Delta^{j-i}\psi(d-j) & i < j, \\ 0 & \text{else}. \end{cases} \tag{16}$$

Hence, a natural, *naïve* method to calculate Q^* is recursively calculating finite forward differences and appropriately changing the signs of the results.

This section shows in a numerical study that this naïve numerical computation of the infinitesimal generator matrix Q^* using Eq. (16) leads to sizeable distortions of transition intensities and probabilities. Subsequently, we propose an improved, staged numerical approximation of Q^*: First, approximate its first row, i.e., the shock-size-arrival intensities, with a numerical integration routine using one of two integral representations of Q^* with nonnegative integrands. Using these representations ensures nonnegativity and allows us to control the approximation's accuracy using a suitable numerical integration method with error bounds. Under weak additional assumptions, these integrals are well suited for numerical integration since their integrands are bounded, except for boundary cases. For the exceptions, we provide suitable alternatives. Second, calculate the remaining rows of Q^* using a simple recursion. Finally, we repeat our study for our proposed approach and demonstrate that it does not suffer from similar distortions as the naïve method.

Special cases with closed-form formulas for Q^*

Before discussing the naïve approach, we present two special cases with a *closed-form* representation of the death-counting process' infinitesimal generator matrix Q^*. We call a representation *closed form* if it can be evaluated quickly with high accuracy, e.g., if it involves only simple algebraic expressions and standard functions such as the exponential, gamma, or beta function. We will use these special cases for the validation of approximations of Q^*.

The first special case is the Poisson family and requires the following auxiliary result, which can be verified via induction.

Lemma 5.1. [3, pp. 92 and 130] *Let* $\psi(x) = 1 - \exp\{-\eta x\}$, $x \geq 0$, $\eta > 0$, *then*

$$(-1)^{i-1}\Delta^i\psi(x) = e^{-\eta x}[1 - e^{-\eta}]^i, \quad x \geq 0, \quad i \in \mathbb{N}.$$

By using this result and Eq. (13), we obtain a closed-form expression of Q^* for the Poisson family:

Proposition 5.2. *Consider an extMO distribution from the Poisson family with Bernstein function* $\psi(x) = [1 - \exp\{-\eta x\}]$, $x \geq 0$, $\eta > 0$. *Then*,

$$q^*_{ij} = \binom{d-i}{j-i} \cdot \begin{cases} -[1 - e^{-\eta(d-i)}] & i = j, \\ e^{-\eta(d-j)}[1 - e^{-\eta}]^{j-i} & i < j, \\ 0 & \textit{else}. \end{cases}$$

The second special case is the Armageddon family, i.e., (almost) affine-linear Bernstein functions with a possible jump after zero. Their generators can be represented in closed form after a straightforward calculation since the second-order finite forward differences of corresponding Bernstein functions are zero, except in zero, see Remark B.3.

Proposition 5.3. *Consider an extMO distribution from the Armageddon family with Beinstein function* $\psi(x) = \beta\mathbb{1}_{\{x>0\}} + \alpha x$, $x \geq 0$, *for* $\alpha, \beta \geq 0$ *with* $\alpha + \beta > 0$. *Then*,

$$q^*_{ij} = \binom{d-i}{j-i} \cdot \begin{cases} -\beta - (d-i)\alpha & i = j < d, \\ \alpha & i+1 = j < d, \\ \alpha + \beta & i+1 = j = d, \\ \beta & i+1 < j = d, \\ 0 & \textit{else}. \end{cases}$$

A simple calculation shows that the mapping $(\psi, d) \mapsto Q^*$ is linear in ψ, particularly, the infinitesimal generator of a convex combination of Bernstein functions is the convex combination of the corresponding infinitesimal generators. Hence, we can separately calculate the infinitesimal generator of the killing and drift part and the infinitesimal generator of the pure-jump part and aggregate them afterward.

Validation of numerical approximations of Q^*

For the cases without a closed form, numerically stable representation of Q^*, any nontrivial approximation for the mapping $(\psi, d) \mapsto Q^*$, especially in higher dimensions d, must be evaluated on its numerical stability. Therefore, we propose two validation criteria:

- A fundamental property of Markov generator matrices is that row sums are equal to zero. Furthermore, the diagonal entries $q^*_{ii} = -\psi(d-i)$, $i \in [d]$, can be calculated with the same accuracy as ψ. Hence, we can measure the total error of transition intensities by comparing the relative differences of the approximations of $-q^*_{ii}$ and $\sum_{j>i} q^*_{ij}$, $i \in [d]_0$: for a numerical approximation $Q^\dagger$ of Q^*, define

$$\varepsilon_1(d) \coloneqq \max_{i\in[d-1]_0} \left| \frac{\sum_{j>i} q^{\dagger}_{ij}}{-q^{\dagger}_{ii}} - 1 \right|.$$

– For families with closed-form formulas, we can calculate Q^* with high accuracy. Consequently, we can compare two numerically approximated versions of Q^*, approximated with the method in question and its closed-form counterpart, respectively, using the subsequently defined metric. For each row i, calculate the total variation distance (TVD) between the approximated transition distributions based on the weights $(\sum_{k>i} q^*_{ik})^{-1} \cdot q^*_{ij}$, $j > i$. In particular, for numerical approximations $Q^{\dagger}$ and $Q^{\ddagger}$ of Q^*, where the latter is calculated with the closed-form representation, define

$$\varepsilon_2(d) \coloneqq \max_{i\in[d-1]_0} \mathrm{TVD}\left(\left\{ \frac{q^{\dagger}_{ij}}{\sum_{k>i} q^{\dagger}_{ik}} : j > i \right\}, \left\{ \frac{q^{\ddagger}_{ij}}{\sum_{k>i} q^{\ddagger}_{ik}} : j > i \right\} \right).$$

Remark. (Calculating the TVD). Our subsequent numerical studies in `R` use the `TotalVarDist` method for `DiscreteDistribution` classes from the packages `distr` and `distrEx`, see [26], for calculating the TVDs between transition distributions.

The two validation criteria measure the distortions of waiting times and transition distributions. While the second validation criterion is only applicable for those families for which Q^* can be calculated with high accuracy using another method, e.g., the Poisson or Armageddon family, the first validation criterion is applicable for every family.

A critical validation aspect is the choice of test cases. Unfortunately, it is impossible to validate a method for every Bernstein function. Therefore, we need to choose a small subset that is large enough to allow meaningful conclusions about a method's numerical stability for all Bernstein functions. In the following, we outline the reasoning for our subjective selection.

– We consider four families, which can be viewed as *extreme points* of important (convex) subclasses: The *Armageddon family*, representing all (almost) affine-linear Bernstein functions; the *Poisson family*, extreme points of the entire Bernstein function class, see Eq. (5) and [28, Chp. 3]; the *Exponential family*, extreme points of complete Bernstein functions, see Eq. (6) and [28, Chp. 6]; and the *α-stable family*, extreme points of Laplace exponents of *completely self-decomposable* laws, see [28, Chp. 5].
– Using normalization by adding a linear part such that $\psi(1) = 1$, see page 314, all involved families have a unique mapping to (bivariate) lower-tail dependence coefficients (LTDCs) in $(0, 1)$. In particular, the *LTDC* of the bivariate margins is $\mathrm{LTDC} = 2 - \psi(2)/\psi(1)$, see [22, Example 5.21]. Thus, to cover a broad spectrum of each family, we choose the members corresponding to low, mid, and high bivariate lower-tail dependence parameters: $\mathrm{LTDC} \in \{0.05, 0.5, 0.95\}$.

Naïve numerical calculation of Q^*

Equation (16) suggests a naïve numerical calculation of Q^*. For this, set lower-triangular values of Q^* to zero, set the diagonal values to $-\psi(d-i)$, $i \in [d]_0$, and calculate the upper-triangular values for $j > i$ in three steps: First, recursively apply finite forward differences and appropriately change the result's sign to calculate $(-1)^{j-i-1}\Delta^{j-i}\psi(d-j)$ numerically. Second, set these values to zero if they are negative. This modification is necessary because these numerical calculations can be negative, albeit theoretically being nonnegative. Third, multiply the result with the binomial coefficient $\binom{d-i}{j-i}$.

Remark. (Binomial coefficients in `R`, see [24]) Binomial coefficients $\binom{n}{k}$, for nonnegative integers $0 \le k \le n$, have exact representations as 64-bit double-precision binary floating-point numbers, in the following

called *binary64* numbers, see [12], for $n \leq 50$. In R, the method `base::choose` calculates binomial coefficients as a *binary64* number via

$$\binom{n}{k} = \begin{cases} 1 & n = k, \\ n \cdot \prod_{j=2}^{k} \frac{n-j+1}{j} & k < 30, \\ n \cdot \prod_{j=2}^{n-k} \frac{n-j+1}{j} & n-k < 30 \quad \text{and} \quad k \geq 30, \\ \exp\{-\log(n+1) - \operatorname{logBeta}(n-k+1, k+1)\} & \text{else}. \end{cases}$$

Expressions are evaluated from left to right and from inner to outer, and logBeta, the logarithm of the beta function, is evaluated using the dedicated numerical routine `lbeta`. The latter case uses the Beta function's binomial coefficient representation

$$\binom{n}{k} = \frac{\Gamma(n+1)}{\Gamma(k+1)\Gamma(n-k+1)} = \frac{1}{(n+1)\cdot B(n-k+1, k+1)}.$$

The results in Figures 1 and 2 show that the naïve approach produces significant differences for low double-digit dimensions and distorts waiting time intensities and transition distributions beyond recognition for mid double-digit dimensions. More precisely, we observe the following:

- The maximum relative error ε_1 for the sum of transition intensities increases exponentially with the dimension once intensities are floored to zero. The range of dimensions requiring no flooring depends on the specific Bernstein function and parameter choice.
- The maximum TVD ε_2 also increases exponentially with the dimension until the maximum TVD, which is 1.
- Note that for the Armageddon family and LTDC = 50%, all values of ψ on $[d]_0$ have an exact representation as binary64 numbers. Hence, the naïve method calculates Q^* exactly for this exceptional circumstance.

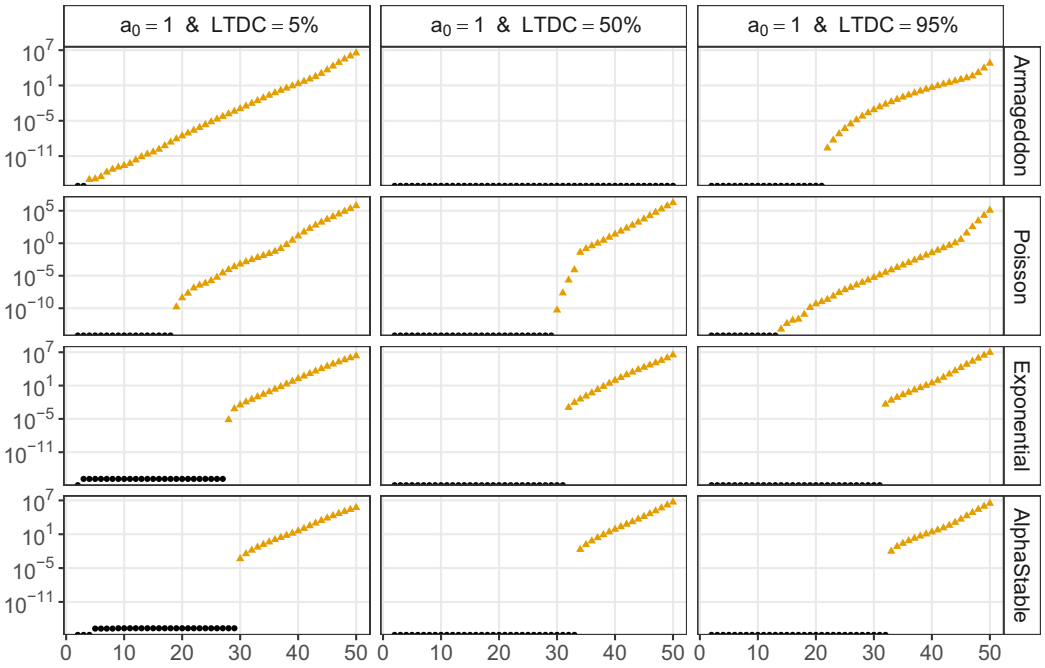


Figure 1: The maximum relative differences $\varepsilon_1(d)$ of the off-diagonal row sums for the naïve approximation of Q^* for $2 \leq d \leq 50$; yellow coloring and triangle shape indicate whether flooring was required.

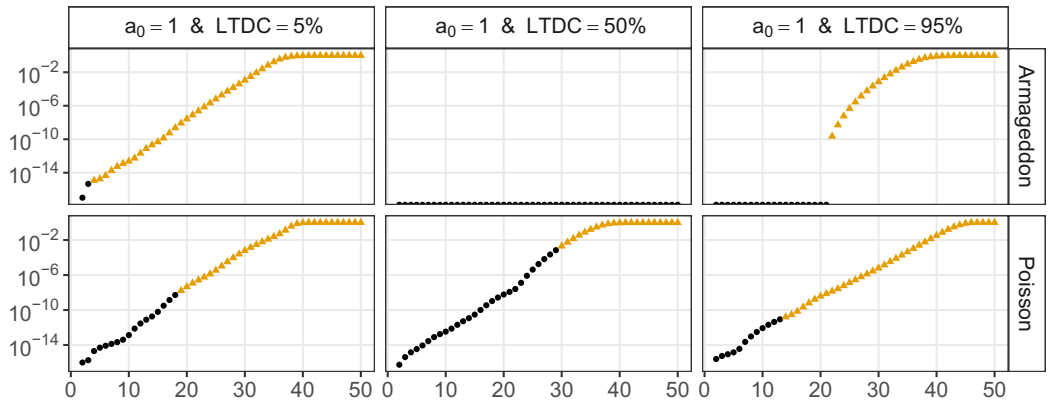


Figure 2: The maximum total variation distance $\varepsilon_2(d)$ of the transition distributions of naïve approximation and high-precision calculation of Q^* for $2 \le d \le 50$; yellow coloring and triangle shape indicate whether flooring was required.

Integral representations of Q^*

Our proposed approximation of Q^* uses two integral representations for the exchangeable *shock-size-arrival intensities*, which will be derived in the following. The idea is simple: Consider a Bernstein function with the representation

$$\psi(x) = \int_I \tilde{\psi}_u(x)\mu(\mathrm{d}u), \quad x \ge 0,$$

where μ is a measure on a set I and $\{\tilde{\psi}_u : u \in I\}$ is a family of Bernstein functions for which the (negative) alternating finite forward differences have a simple, closed-form representation. Note that such a representation exists for every Bernstein function with a drift and killing rate equal to zero due to the *Lévy–Khintchine representation* from Eq. (5) and Lemma 5.1. Then, by using the linearity of the integral, we obtain

$$(-1)^{i-1}\Delta^i\psi(x) = \int_I \underbrace{(-1)^{i-1}\Delta^i\tilde{\psi}_u(x)}_{\ge 0}\mu(\mathrm{d}u), \quad x \ge 0,\ i \in \mathbb{N}.$$

Assuming this integral can be numerically approximated, e.g., because μ is a discrete measure with finite support or if μ has a suitable density with respect to the Lebesgue measure, this representation has the advantage that the approximation is always nonnegative. Furthermore, numerical integration methods, such as Gauss–Kronrod quadrature, allow utilizing error estimates. Consequently, we deem it preferable to approximate directly the integrals

$$q^*_{ij} = \binom{d-i}{j-i} \cdot (-1)^{j-i-1}\Delta^{j-i}\psi(d-j) = \int_I \binom{d-i}{j-i} \cdot (-1)^{j-i-1}\Delta^{j-i}\tilde{\psi}_u(d-j)\mu(\mathrm{d}u).$$

The first integral representation for Q^* uses the *Lévy–Khintchine representation* from Eq. (5) and holds for all Bernstein functions:

Theorem 5.4. *Let $d \ge 2$ and consider a d-variate extMO distribution with Bernstein function ψ and Lévy triplet (a, b, ν) such that Eq. (5) is fulfilled. Then, for $0 \le i < j \le d$, we have*

$$\lambda_i = a\mathbb{1}_{\{i=d\}} + b\mathbb{1}_{\{i=1\}} + \int_0^\infty e^{-u(d-i)}[1-e^{-u}]^i\nu(\mathrm{d}u) \tag{17a}$$

and

$$q^*_{ij} = \binom{d-i}{j-i} \cdot [a\mathbb{1}_{\{j=d\}} + b\mathbb{1}_{\{j-i=1\}}] + \int_0^\infty \binom{d-i}{j-i} \cdot e^{-u(d-j)}[1-e^{-u}]^{j-i}\nu(\mathrm{d}u). \tag{17b}$$

Proof. The result is a direct corollary of Eq. (9), Eq. (16), Lemma 5.1, and Proposition 5.3. Note that the formula for the exchangeable shock-arrival intensities in Eq. (17a) had previously been stated in [17, p. 149]. □

The second integral representation for Q^* uses the *Stieltjes representation* from Eq. (6), which exists for all complete Bernstein functions.

Lemma 5.5. *Let ψ be a complete Bernstein function with Stieltjes triplet $(0, 0, \sigma)$ such that Eq. (6) is fulfilled. Then,*

$$(-1)^{i-1}\Delta^i\psi(x) = \int_0^\infty u \cdot B(1+i, x+u)\sigma(\mathrm{d}u), \quad x \geq 0,\ i \in \mathbb{N},$$

where $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$, $x, y > 0$, is the Beta function.

Proof. Due to the linearity of the integral, it suffices to show that

$$(-1)^{i-1}\Delta^i \frac{x}{x+u} = u \cdot B(1+i, x+u), \quad x \geq 0,\ u > 0,\ i \in \mathbb{N}.$$

We prove this by induction. For this, note that we have

$$\begin{aligned}\Delta\frac{x}{x+u} &= \frac{x+1}{x+1+u} - \frac{x}{x+u} = \frac{(x+1)(x+u)}{(x+u+1)^2} - \frac{x(x+u+1)}{(x+u+1)^2} = u \cdot \frac{1}{(x+u+1)^2} = u \cdot \frac{\Gamma(2)\Gamma(x+u)}{\Gamma(x+u+2)} \\ &= u \cdot B(1+1, x+u).\end{aligned}$$

Now, assume that the claim is true for $i-1$. Then,

$$\begin{aligned}(-1)^{i-1}\Delta^i\frac{x}{x+u} &= -\Delta\left[(-1)^{i-2}\Delta^{i-1}\frac{x}{x+u}\right] \overset{\text{ind.}}{=} -\Delta[u \cdot B(1+i-1, x+u)] \\ &= -\Delta\left[u \cdot \frac{\Gamma(i)\Gamma(x+u)}{\Gamma(x+u+i)}\right] = -\Delta\left[u \cdot \frac{(i-1)!}{(x+u+i-1)^{\underline{i}}}\right] \\ &= u(i-1)! \cdot \left[\frac{1}{(x+u+i-1)^{\underline{i}}} - \frac{1}{(x+u+i)^{\underline{i}}}\right] \\ &= u(i-1)! \cdot \left[\frac{(x+u+i)-(x+u)}{(x+u+i)^{\underline{i}}(x+u)}\right] = u \cdot \frac{i!}{(x+u+i)^{\underline{i+1}}} \\ &= u \cdot \frac{\Gamma(i+1)\Gamma(x+u)}{\Gamma(x+u+i+1)} = u \cdot B(1+i, x+u).\end{aligned}$$

□

Theorem 5.6. *Let $d \geq 2$ and consider a d-variate extMO distribution with complete Bernstein function ψ and Stieltjes triplet (a, b, σ) such that Eq. (6) is fulfilled. Then, for $0 \leq i < j \leq d$,*

$$\lambda_i = a\mathbb{1}_{\{i=d\}} + b\mathbb{1}_{\{i=1\}} + \int_0^\infty u \cdot B(1+i, d-i+u)\sigma(\mathrm{d}u) \tag{18a}$$

and

$$q^*_{ij} = \binom{d-i}{j-i} \cdot [a\mathbb{1}_{\{j=d\}} + b\mathbb{1}_{\{j-i=1\}}] + \int_0^\infty \binom{d-i}{j-i} \cdot u \cdot B(1+j-i, d-j+u)\sigma(\mathrm{d}u), \tag{18b}$$

where B is the Beta function.

Proof. The result is a direct corollary of Eq. (9), Eq. (16), Lemma 5.5, and Proposition 5.3. □

Numerical approximation of Q^* based on integral representations

We want to use the integral representations from Eqs. (17) and (18) to approximate Q^*. However, even if integrals are finite in theory, numerical integration routines can fail for various reasons. In the following, we provide further insight into using numerical integration methods for the two representations.

Remark 5.7. (Numerical integration in R with QUADPACK algorithms). We can perform numerical integration in R with the `stats::integrate` routine. The method uses the well-known *QUADPACK* algorithms *QAGI* (indefinite integrals) and *QAGS* (definite integrals), see [23,24]. These algorithms are also used in the *GNU scientific library* (*GSL*), see [8]. They are adaptive numerical integration algorithms based on a 15-point and 21-point *Gauss–Kronrod quadrature* with the convergence acceleration technique *Wynn's ε-algorithm* for a limit extrapolation, respectively. The former method transforms indefinite integrals with lower bound $a \in \mathbb{R}$ as follows:

$$\int_a^\infty f(x)\mathrm{d}x = \int_0^1 f\left(a + \frac{1-y}{y}\right) y^{-2}\mathrm{d}y.$$

Theoretically, the limit of the *m-panel*, n-point Gauss–Kronrod quadrature, which applies the n-point quadrature to m equidistant subintervals of the integration domain, converges to the integral for $m \to \infty$ if f is bounded and Riemann-integrable, see [23, Thm. 2.6]. This approximation also converges for some functions with singularities, e.g., $f(x) = x^\beta$, $\beta > -1$ for $x \in (0, 1)$; however, the convergence is very slow in this case if $\beta < 0$, see [23, p. 42]. A low convergence rate can mislead implementations to falsely detect divergence.

First, note that the integral is the inner product of weights and integrand values if the Lévy measure, or Stieltjes measure, is discrete with finite support.

In the remainder of this subsection, assume that the Lévy measure, or Stieltjes measure, has a continuous density with respect to the Lebesgue measure on $(0, \infty)$. While there are examples for which this is not the case, e.g., the Stieltjes density of no. 5 in [28, p. 304], it is fulfilled for all examples with nondiscrete representation measures discussed in this article.

The following boundary conditions for continuous Lévy and Stieltjes densities allow identifying the cases for which numerical representations of q_{ij}^* have singularities.

Lemma 5.8. *Let f be a continuous Lévy density, i.e., a continuous function on $(0, \infty)$ such that $\int_0^\infty (1 \wedge u) f(u)\mathrm{d}u < \infty$. Then,*

$$\lim_{x \to 0} x^2 f(x) = 0 \tag{19a}$$

and

$$\lim_{x \to \infty} x f(x) = 0. \tag{19b}$$

Proof. First, for Eq. (19a), note that $\int_0^\infty (1 \wedge u) f(u)\mathrm{d}u < \infty$ implies by

$$0 = \lim_{y \to 0} \int_0^y u f(u)\mathrm{d}u = \lim_{y \searrow 0} y \cdot \zeta(y) f(\zeta(y)) \geq \lim_{x \searrow 0} x^2 f(x) \geq 0,$$

where $\zeta(y) \in [0, y]$ are determined by the *mean value theorem for integration* using continuity of f. Second, for Eq. (19b), using the substitution $u = (1 - t)t^{-1}$, we have

$$\infty > \int_0^\infty (1 \wedge u) f(u) \mathrm{d}u \geq \int_0^\infty \frac{u}{1+u} f(u) \mathrm{d}u = \int_0^1 \frac{1-t}{t^2} f((1-t)t^{-1}) \mathrm{d}t.$$

Hence,

$$\begin{aligned} 0 &= \lim_{z \to 0} \int_0^z \frac{1-t}{t^2} f((1-t)t^{-1}) \mathrm{d}t = \lim_{z \to 0} z \cdot \frac{1-\zeta(z)}{\zeta(z)^2} f((1-\zeta(z))\zeta(z)^{-1}) \\ &\geq \lim_{y \to 0} y \cdot \frac{1-y}{y^2} f((1-y)y^{-1}) = \lim_{x \to \infty} x f(x) \geq 0, \end{aligned}$$

where $\zeta(z) \in [0, z]$ are again determined by the *mean value theorem for integration*. □

Lemma 5.9. *Let f be a continuous Stieltjes density, i.e., a continuous function on $(0, \infty)$ such that $\int_0^\infty (1+u)^{-1} g(u) \mathrm{d}u < \infty$. Then,*

$$\lim_{x \to 0} x g(x) = 0 \tag{20a}$$

and

$$\lim_{x \to \infty} g(x) = 0. \tag{20b}$$

Proof. First, for Eq. (20a), note that $\int_0^\infty (1+u)^{-1} g(u) \mathrm{d}u < \infty$ implies

$$0 = \lim_{y \to 0} \int_0^y (1+u)^{-1} g(u) \mathrm{d}u = \lim_{y \to 0} y(1+\zeta(y))^{-1} g(\zeta(y)) \geq \lim_{x \to 0} x g(x) \geq 0,$$

where $\zeta(y) \in [0, y]$ are determined by the *mean value theorem for integration* using continuity of g. Second, for Eq. (20b), consider that, using the substitution $u = (1-t)t^{-1}$, we have

$$\infty > \int_0^\infty (1+u)^{-1} g(u) \mathrm{d}u = \int_0^1 g((1-t)t^{-1}) t^{-1} \mathrm{d}t.$$

Hence,

$$\begin{aligned} 0 &= \lim_{z \to 0} \int_0^z g((1-t)t^{-1}) t^{-1} \mathrm{d}t = \lim_{z \to 0} z \cdot g((1-\zeta(z))\zeta(z)^{-1}) \zeta(z)^{-1} \\ &\geq \lim_{y \to 0} g((1-y)y^{-1}) = \lim_{x \to \infty} g(x) \geq 0, \end{aligned}$$

where $\zeta(z) \in [0, z]$ are again determined by the *mean value theorem for integration*. □

By using the boundary results from Lemma 5.8, we obtain for a continuous Lévy density f and numbers $y > 0$, $x \geq 0$, and $k \in \mathbb{N}$

$$\lim_{t \to 0} y \cdot e^{-x(1-t)t^{-1}} [1 - e^{-(1-t)t^{-1}}]^k f((1-t)t^{-1}) t^{-2} = y \cdot \lim_{u \to \infty} e^{-xu} u^2 f(u) \begin{cases} \text{unknown} & x = 0, \\ = 0 & x > 0 \end{cases}$$

and

$$\lim_{t \to 1} y \cdot e^{-x(1-t)t^{-1}} [1 - e^{-(1-t)t^{-1}}]^k f((1-t)t^{-1}) t^{-2} = y \cdot \lim_{u \to 0} u^k f(u) \begin{cases} \text{unknown} & k = 1, \\ = 0 & k \geq 2. \end{cases}$$

Furthermore, by using the boundary results from Lemma 5.9, we obtain for a continuous Stieltjes density g and numbers $y > 0$, $x \geq 0$, and $k \in \mathbb{N}$

$$\lim_{t\to 0} \gamma \cdot (1-t)t^{-3}B(1+k, x+(1-t)t^{-1})g((1-t)t^{-1}) = \gamma \cdot \lim_{u\to\infty} \frac{u^3 \cdot k! \cdot g(u)}{(x+u+k)^{\underline{k+1}}} \begin{cases} \text{unknown} & k=1, \\ =0 & k \geq 2, \end{cases}$$

and

$$\lim_{t\to 1} \gamma \cdot (1-t)t^{-3}B(1+k, x+(1-t)t^{-1})g((1-t)t^{-1}) = \gamma \cdot \lim_{u\to 0} \frac{u \cdot k! \cdot g(u)}{(x+u+k)^{\underline{k+1}}} \begin{cases} \text{unknown} & x=0, \\ =0 & x>0. \end{cases}$$

Thus, for a continuous Lévy density, or Stieltjes density, the integrand in Eq. (17), or Eq. (18), after a substitution $u = (1-t)t^{-1}$, tends to zero at both ends whenever $j < d$ and $j - i > 1$. Consequently, under these assumptions, the *QUADPACK* algorithms are well suited for approximating the integral Eq. (17), or Eq. (18), if $j < d$ and $j - i > 1$.

Improved numerical approximation of Q^*

The previous subsection outlined that, given a continuous Lévy density, or Stieltjes density, we can expect *QUADPACK* algorithms to work well for the respective integral representations of q^*_{ij} whenever $j - i > 1$ and $j < d$. That does not imply that they will fail if $j - i = 1$ or $j = d$. However, if they do, we can use that these exceptions are boundary cases:

- For $j - i = 1$, we have

$$q^*_{i,i+1} = (d-i)\cdot[\psi(d-i) - \psi(d-i-1)].$$

  Hence, we can calculate $q^*_{i,i+1}$ with the naïve approach as the result of a single finite forward difference. This calculation is sufficiently exact and does not suffer from the numerical issues of recursively calculating finite differences.
- For $j = d$, we have multiple options. Here, the problem is a possible singularity for the integrand. The first option is to use the continuity of the integrand by replacing the expression $d - j$ iteratively with an increasing sequence of real numbers $d - j + \varepsilon_k$, $k \in \mathbb{N}$, until the numerical integration succeeds. The second option is to use that row sums of Q^* have to be equal to zero. Thus, this condition implicitly defines an approximation of q^*_{id}, given approximations of all other values, flooring it to zero if the resulting approximation is negative. Of both approaches, we suggest using the latter over the former since it is not iterative and flooring should only be necessary if the value is already close to zero, given that the remaining values are sufficiently accurate.

To avoid rounding issues, we also suggest using the following recursive representation for products with binomial coefficients for $0 \leq k \leq n$:

$$\left[\binom{n}{k} \cdot x\right] = \begin{cases} \frac{n}{k} \cdot \left[\binom{n-1}{k-1} \cdot x\right] & 1 \leq k \leq \lfloor n/2 \rfloor, \\ \left[\binom{n}{n-k} \cdot x\right] & k > \lfloor n/2 \rfloor, \\ x & k = 0. \end{cases}$$

The following result allows using a recursion for numerically calculating the remaining rows from the first row of Q^*. Consequently, it is sufficient to approximate only the first row with numerical integration.

Theorem 5.10. *Let $d \geq 2$ and consider a d-variate extMO distribution with Bernstein function ψ and infinitesimal generator matrix Q^* for the corresponding law of the death-counting process. Then, for $i < j$,*

$$q^*_{i+1,j+1} = \frac{d-j}{d-i} \cdot q^*_{i,j} + \frac{j+1-i}{d-i} \cdot q^*_{i,j+1}.$$

Proof. We have

$$\begin{aligned}(-1)^{j-i-1}\Delta^{j-i}\psi(d-j-1) &= (-1)^{j-i-1}\Delta^{j-i}[\psi(d-j)+\psi(d-j-1)-\psi(d-j)]\\ &= (-1)^{j-i-1}\Delta^{j-i}\psi(d-j)+(-1)^{j-i}\Delta^{j-i+1}\psi(d-j-1).\end{aligned}$$

Consequently,

$$q^*_{i+1,j+1} = \frac{\binom{d-i-1}{j-1}}{\binom{d-i}{j-i}}q^*_{i,j} + \frac{\binom{d-i-1}{j-i}}{\binom{d-i}{j+1-i}}q^*_{i,j+1} = \frac{d-j}{d-i}\cdot q^*_{i,j} + \frac{j+1-i}{d-i}\cdot q^*_{i,j+1}. \qquad \square$$

The following remark summarizes the proposed approximation of Q^*.

Remark 5.11. Assuming a continuous Lévy density, or Stieltjes density, we can approximate Q^* as follows:
(1) Set all of the lower-triangular values to zero.
(2) Numerically calculate the diagonal values as $-\psi(d-i)$, $i \in [d]_0$.
(3) Approximate $q^*_{0,j}$, $1 < j < d$, using numerical integration, e.g., with the *QUADPACK* algorithms, based on one of the integral representations Eqs. (17) and (18).
(4) For $j = 1$, calculate $q^*_{0,1}$ as $d[\psi(d) - \psi(d-1)]$ numerically.
(5) For $j = d$, set $q^*_{0,d}$ to the value, equating the first-row sum of Q^* to zero, and then floor this value by zero.
(6) Calculate the remaining rows of Q^* numerically using the recursion from Theorem 5.10.

We repeated the previous numerical studies for the proposed approach using the relative tolerance $\sqrt{\varepsilon} \approx 1.490116 \times 10^{-8}$ for the accuracy of the numerical integration, where $\varepsilon = 2^{-52}$ is the *machine epsilon* for a *binary64* number. We can see in Figures 3 and 4 that $\varepsilon_1(d)$ and $\varepsilon_2(d)$ stay below or around $\sqrt{\varepsilon}$ up to $d = 125$.

6 Benchmarks of extMO simulation algorithms

Previously, we derived the Markov death-counting model (MDCM) for exMO distributions in Section 4. Furthermore, we have shown how their inputs, the infinitesimal Markov generators of their death-counting processes, can be approximated in a numerically stable way from Bernstein functions, parametrizing extMO distributions, in Section 5. In combination with Algorithm 1, we obtain a novel simulation algorithm for high-dimensional extMO distributions. This section aims to compare the runtime of this algorithm with that of alternative sampling algorithms for extMO distributions, summarized in Table 1. We also investigate the algorithm's setup activities' proportion of the overall runtime.

We performed the presented benchmarks on a Windows consumer laptop. In addition, we also repeated them on other devices and operating systems. If not mentioned otherwise, the presented findings were representative of all machines. However, we want to stress that our intention with the subsequent analysis is merely to indicate how the runtimes compare; hardware and software changes can significantly impact such measurements.

The following is a high-level summary of our key findings from theoretical considerations and subsequent benchmarks.

- The general MO simulation algorithms ESM and AM appear ill suited for simulating extMO distributions in higher dimensions without tweaking the algorithms to special cases. In particular, both have relatively large memory requirements for $d \approx 30$, and the former shows significantly larger runtimes up from $d \approx 8$ than all other benchmarked algorithms.
- The MDCM was slower than the LFM for most tested configurations. However, the former required a large proportion of its runtime for the initial setup. Consequently, the gaps between both algorithms' runtimes decreased when increasing the sample size. In addition, we found a way to choose the parameters so that

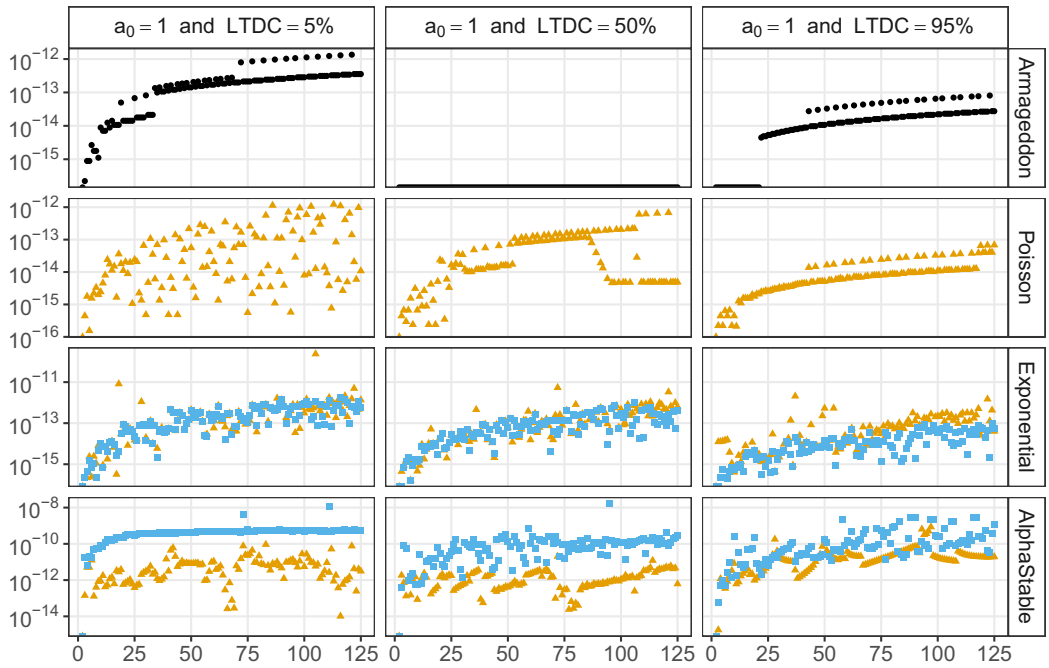


Figure 3: The maximum relative differences $\varepsilon_1(d)$ of the off-diagonal row-sums for the proposed calculation of Q^* for $2 \leq d \leq 50$; coloring and shape indicate the used integral representation with none (black, round), i.e., closed-form, Lévy (yellow, triangle), and Stieltjes (blue, square).

the LFM's (expected) runtime becomes arbitrarily large — such a scenario does not exist for the MDCM as its dimension bounds the transition number.

Overall, we conclude that the MDCM is a viable option for simulating extMO distributions in dimensions at least up to $d = 128$. The LFM is an alternative for suitable extMO distributions. However, it has issues with its runtime explosions in corner cases, and it has limited applicability. In particular, exact simulation using the LFM requires the underlying Lévy subordinator to be a compound Poisson process and being able to simulate from its jump distribution. In contrast, if the Lévy subordinator is not compound Poisson, the corresponding extMO distribution cannot be sampled exactly with the LFM but requires approximations similar to those in [7, Sec. 5.3]. The MDCM also relies on numerically integrating Q^*. However, the corresponding approximation errors can be bound with an appropriate numerical integration method, as

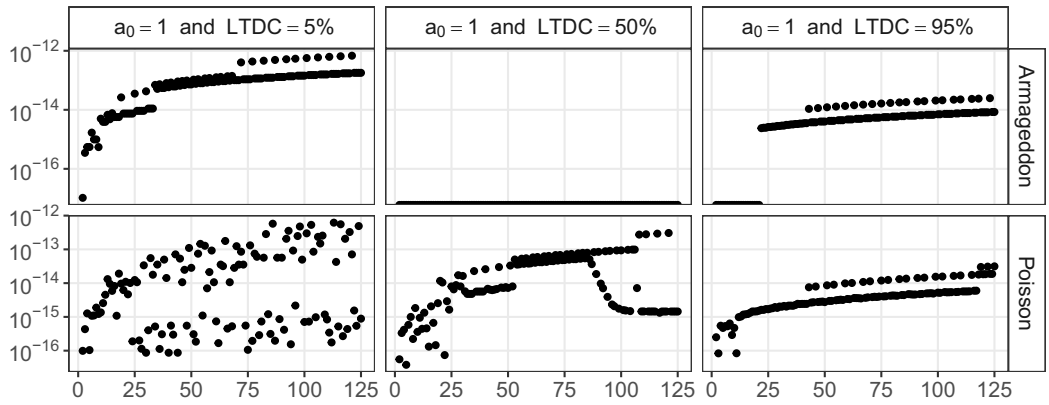


Figure 4: The maximum total variation distance $\varepsilon_2(d)$ of the transition distributions of the proposed and high-precision calculation of Q^* for $2 \leq d \leq 50$.

Table 1: An overview of our implementations for the runtime comparison; for more details on the software used, see our statement following the conclusion

Algorithm	Input[e]	Restrictions	
		Subfamily	Max. dimension[f]
ESM[a]	$\{\lambda_I : I \subseteq [d]\}$	MO	$\log_2(n_{\max} + 1) - 1$
AM[b]	$\{\lambda_I : I \subseteq [d]\}$	MO	$\log_2(n_{\max} + 1) - 1$
MDCM[c]	$\{\eta_i : i \in [d]\}$	exMO	$n_{\max}$
LFM[d]	a, b, $\nu(0, \infty)$, and jump dist. param.	extMO[d]	$n_{\max}$
Mod. ESM[a]	$\alpha = \lambda_1$ and $\beta = \lambda_d$	Armageddon	$n_{\max}$

[a] Optimized implementation of [17, Algo. 3.1], see Section 2.1, skipping a shock at runtime if the shock-arrival intensity is zero and using bit arithmetics for death verification. We modified the algorithm for the Armageddon family to sampling $d + 1$ exponential random variables and performing a loop with d bivariate minimum operations.
[b] Optimized implementation of [17, Algo. 3.3], see Section 2.1, using Walker's alias algorithm for discrete sampling, see [31], and bit arithmetics for death verification.
[c] Implementation of Algorithm 1 using the representation of Theorem 2.3 with Walker's alias algorithm for discrete sampling.
[d] The LFM requires the subordinator to be a compound Poisson process with a feasible jump distribution sampling algorithm. We implemented [17, Algo. 3.7] with deterministic, exponential, and Pareto distributed jumps.
[e] The input parameters for ESM, AM, and MDCM are calculated in `R` using the approach from Section 5.
[f] Our implementation's (technical) maximum dimension is expressed as a function of the maximum vector size $n_{\max}$. E.g., in the case of the vector size being a 32-bit unsigned integer, we have $n_{\max} = 2^{32} - 1$; `R` with `LongVector`-support has a technical maximum of $n_{\max} = 2^{52}$, but considering a 2^{30}-length `binary64` double vector requires roughly $8.6GB$ memory, we believe that general MO distributions are infeasible long before that technical maximum is reached. Note that implementations for larger dimensions are possible but require specialized data structures and could be suboptimal for smaller dimensions. In particular, optimizations based on bit arithmetics might not be possible anymore.

discussed in Section 5. Furthermore, statistical tests did not reject any distributional assumption for our implementation of the MDCM for any subfamilies discussed in this section; see Appendix D. Finally, we want to highlight that we can use the MDCM for arbitrary extMO distributions with closed-form, continuous Lévy or Stieltjes densities without requiring further specialization.

For the ensuing comparisons, we describe extMO distributions by the properties of their subordinators in the LFM representation from Theorem 2.4, as it is exceptionally well suited for obtaining a basic understanding of the probability law. For this, recall that we can characterize every extMO distribution by a Bernstein function ψ, which defines the law of a (potentially killed) Lévy subordinator. Components corresponding to the extMO distributed random vector are killed once the subordinator passes their individual unit exponential barrier values. For compound Poisson subordinators, we distinguish subordinator laws by their jump intensity and jump size distribution: The jump intensity translates into the overall speed with which the subordinator surpasses the barrier values. Thus, it corresponds to the random vector's marginal rate. The distribution of jump sizes predefines the chances of the subordinator simultaneously surpassing multiple barrier values, and therefore, it corresponds to the random vector's dependence structure. Simply put, a high probability of larger jumps increases the chance of simultaneous deaths, while a high probability of smaller jumps increases the likelihood of individual deaths. This logic culminates into the pure-drift and pure-killing corner cases, corresponding to the independence and comonotonicity, respectively, pure-jump Lévy subordinators with *infinite activity*, which can be approximated by compound Poisson processes, and convex combinations of those above.

We have chosen the exponential family with possible drift and killing as a representative example for our subsequent benchmarks for two reasons: First, its Bernstein function has a simple form. Second, its LFM representation requires the simulation of jumps but does not require potentially expensive simulation techniques such as *rejection sampling*. However, whenever possible, we performed the following benchmarks for all other families from Section 2.4 without observing noteworthy structural differences.

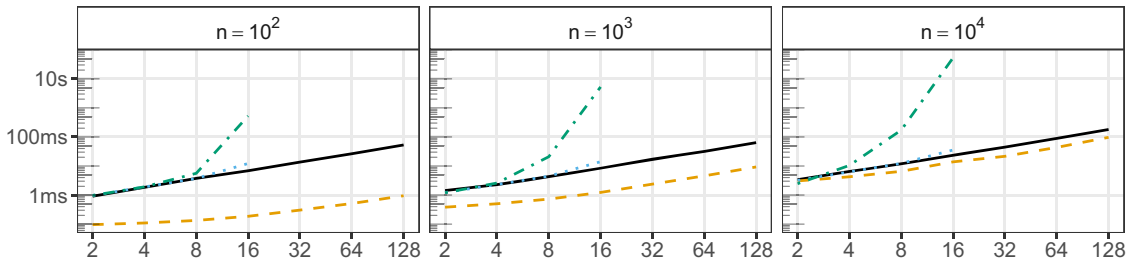


Figure 5: The median runtime of the MDCM (solid, black), LFM (dashed, yellow), AM (dotted, blue), and ESM (dash-dotted, green) algorithms for an extMO distribution from the exponential family with drift and without killing, calibrated to unit exponential margins and a bivariate lower-tail dependence coefficient of 50%. We measured the AM's and ESM's runtimes for $d \leq 16$.

We begin by comparing the MDCM's simulation runtime to that of the LFM, AM, and ESM for the exponential family, using normalization by adding a linear part, see Section 2.4. Furthermore, we fix the bivariate margin to an extMO distribution with unit margins and a lower-tail dependence coefficient of 50%, see [22, Exmpl. 5.21]. We make the following observations from the benchmark results in Figures 5 and 6:

- We conclude that the ESM is ill suited for higher dimensions due to exploding runtimes for increasing dimensions. In contrast, the AM seems to be less problematic: although it has higher setup requirements than the ESM, the overall runtime is slightly higher than that of the MDCM and LFM but significantly smaller than that of the ESM. We attribute this in part to using the highly efficient *Walker's alias method* for discrete sampling and also in part to a binary representation of shock sets, allowing a quick death verification. However, using a binary representation for shock sets makes it challenging to scale this particular implementation to higher dimensions, e.g., our implementation is limited to $d = 30$ with 16 GB memory. Furthermore, Figure 6 shows that the setup activities' proportion of the median runtime is similar for the MDCM and the AM, but diminishes for the ESM with increasing dimensions.
- In this benchmark, the MDCM's median runtime was measurably slower than that of the LFM. However, the gap shrinks significantly when increasing the sample size from $n = 10^2$ to $n = 10^4$.
- In contrast to the LFM, the MDCM requires setup activities, mapping the Bernstein function to the infinitesimal generator, contributing significantly to the overall runtime. We observe from Figure 6 that while the proportional runtime of setup activities is sizeable for small sample sizes, it becomes less pronounced for larger sample sizes. In particular, for $d = 128$, the runtime proportion of the setup decreases significantly by a factor of approximately 3 when increasing the sample size from $n = 10^2$ to $n = 10^4$. Note that the proportional runtime of setup activities could be decreased as, contrary to the sampling algorithm, we implemented them in `R` and not `C++`.

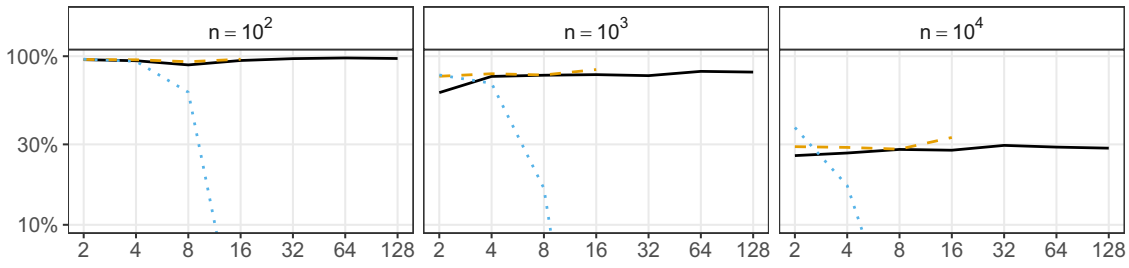


Figure 6: The setup activities' proportion of the median runtime of the MDCM (solid, black), AM (dashed, yellow), and ESM (dotted, blue) algorithms for extMO distributions from the exponential family with drift and without killing, calibrated to unit exponential margins and a bivariate lower-tail dependence coefficient of 50%. We measured the setup activities' median runtime in a separate benchmark. Similar to Figure 5, we measured the AM's and ESM's runtimes for $d \leq 16$.

We continue with a short theoretical comparison of the MDCM and LFM for the case without drift or killing to highlight the possibility of exploding runtimes for the LFM: Both stochastic models have embedded transition processes counting the number of dead components. For the LFM, the compound Poisson subordinator can surpass several barrier values and trigger the death of the corresponding components with each new increment. The counting process accumulating the death toll over these transitions has a similar property to the AM: it can loop in the same state for several transitions. In particular, we show in Appendix C that the number of transitions to surpass the barrier value of the first component, subsequently called *barrier-exceedance count*, has a geometric distribution with success probability $p = 1 - \kappa(1)$, where κ is the Laplace transform of the jump distribution. Consequently, the expected number of transitions for the subordinator surpassing all barrier values is at least $1/[1 - \kappa(1)]$ and can become arbitrarily large depending on the jump distribution. Conversely, the death counting process of the MDCM cannot loop in the same state. Thus, the total number of transitions until reaching the absorbing state cannot exceed dimension d. For more details on the expected number of transitions of the MDCM and LFM, see Appendix C.

In the second benchmark, we demonstrate the issue of exploding runtimes for the LFM using a special case of the exponential family: a compound Poisson subordinator with exponentially distributed jumps without drift or killing. By decreasing p, the success probability of the geometric distribution corresponding to the barrier-exceedance count from the previous paragraph, we can produce a particularly adverse parametrization for the LFM. A short calculation yields that $p = 50\%$, $p = 10\%$, and $p = 1\%$, respectively, correspond to rates $\eta = 1$, $\eta = 9$, and $\eta = 99$ of the jump distribution. We conclude from Figure 7 that the benchmarks for smaller success probabilities p highlight an advantage of the MDCM over the LFM: The dimension bounds the number of transitions in the former model. Thus, the median runtime cannot *explode* if the success probability approaches zero as in the LFM.

In the previous benchmark, we chose parameters to highlight a disadvantage of the LFM for small success probabilities of the first component's barrier-exceedance count distribution. Now, we probe a parametrization more to the LFM's advantage and, in particular, to the MDCM's disadvantage. For this, note that we can expect the fastest runtime of the LFM for a pure-drift subordinator, corresponding to an independence distribution: The LFM algorithm samples and sorts the barrier values and iterates over the barrier values, setting the random variables to the barrier values. In particular, it requires no simulation of subordinator increments. Conversely, the independence case constitutes the worst case for the MDCM as the death-counting process will always require d transitions into the absorbing state. Nevertheless, we infer

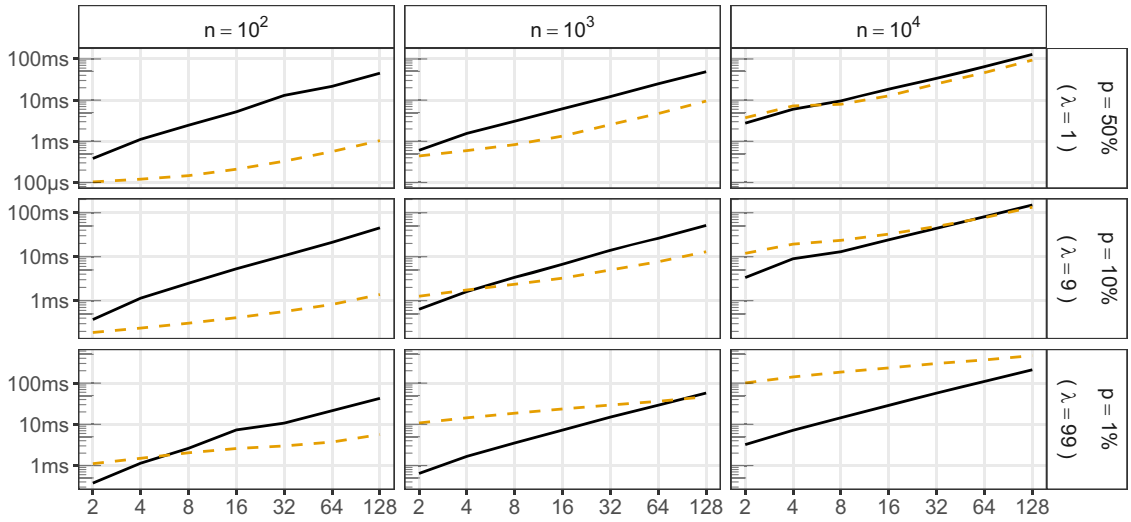


Figure 7: The median runtime of the MDCM (solid, black) and LFM (dashed, yellow) algorithms for extMO distributions from the exponential family without drift or killing. We used three configurations for various success probabilities p from the first component's (geometric) barrier-exceedance count distribution.

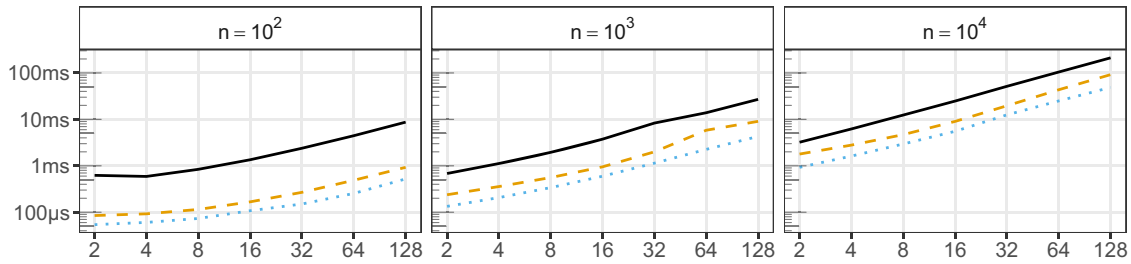


Figure 8: The median runtime of the MDCM (solid, black) and LFM algorithms (dashed, yellow) for the extMO distribution with a pure-drift subordinator (independence case). We included the optimized ESM algorithm for the Armageddon family (dotted, blue) as a reference. It hints at the costs of the sorting operation of the LFM algorithm, which is superfluous for a pure-drift subordinator.

from the benchmark results in Figure 8 that the LFM algorithm is only approximately twice as fast as the MDCM algorithm in this parametrization for larger sample sizes.

7 Conclusion

We started this article with the question: "Can we find low-parametric Markov-based models for extendible Marshall–Olkin distributions with a numerically stable implementation?"

We showed that the death-counting process of exchangeable Marshall–Olkin distributions is Markov. Furthermore, we proved that its shuffled transition times' distribution equals the original one. This result allows us to represent exchangeable Marshall–Olkin distributions directly with Markovian death-counting processes.

Using this Markov representation requires calculating its infinitesimal generator from the distribution's parameters. Many well-known examples of low-parametric subfamilies of extendible Marshall–Olkin distributions are parametrized via Bernstein functions, the Lévy exponents of the subordinators from the Lévy frailty model representation. However, we found that naïvely calculating Markov generators from Bernstein functions is not numerically stable. Therefore, we derived a numerically stable approximation of the generator's first row using integral representations of Bernstein functions and summation identities and calculate its remaining rows using a recursion. We conducted a numerical study to demonstrate this approach's numerical stability for various examples up to dimension $d = 128$.

We proposed a new simulation algorithm by combining extendible Marshall–Olkin distributions' Markov representations with the numerically stable approximation of their generators. However, this algorithm necessitates significant setup activities for calculating the generators from Bernstein functions, requiring numerical integration and recursions. We benchmarked the runtime of our new algorithm to that of alternative sampling algorithms, corresponding to the Lévy frailty model, the Arnold model, and the exogenous shock model, and obtained the following findings: First, we found that the setup activities' proportion of our algorithm's overall runtime was sizeable for small sample sizes, e.g., 10^2, but decreased significantly when moving toward larger sample sizes, e.g., 10^4. Second, we found that both specialized algorithms — our algorithm and the algorithm corresponding to the Lévy frailty model — were faster than the general algorithms, corresponding to the Arnold model and exogenous shock model. Third, we found that our algorithm was, in many cases, slower than the algorithm corresponding to the Lévy frailty model. However, the gap between the measured runtimes decreased when increasing the sample size and, in contrast to its counterpart, our simulation algorithm's runtime is bounded by the dimension. Finally, we analyzed worst-case parametrizations for the Markov death-counting model and the Lévy frailty model. We found that the latter's runtime explodes for particular parameter asymptotics. In contrast, the former's runtime is bound for a fixed dimension.

Apart from runtime differences, the Lévy frailty model and Markov death-counting model have significant methodological differences. The Lévy frailty model comprises specialized sampling algorithms, each depending on a specific jump distribution algorithm. In contrast, the Markov death-counting model is a general sampling algorithm for all extendible Marshall–Olkin distributions, requiring no specialization if it has suitable Lévy or Stieltjes density. The Markov death-counting model requires an approximation of the simulation parameters except for special cases, but the approximation uses numerical integration methods, which can bound the error. Moreover, the Lévy frailty model also requires approximations if the corresponding Lévy subordinator is not of compound Poisson type.

We conclude that our novel Markov-based simulation algorithm is well suited for simulating extendible Marshall–Olkin distributions. In contrast to the Lévy frailty model, our Markov representation covers all extendible Marshall–Olkin distributions and does not require tailored jump distribution algorithms.

Acknowledgement: Many thanks to Lexuri Fernández, Matthias Scherer, and the anonymous referees for their feedback on the article.

Funding information: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of interest: The author states no conflict of interest.

Software used: We used `R` (see [24]) for implementing the MDCM, LFM, AM, and ESM simulation algorithms, the corresponding parameter mapping routines from Section 5, and for performing the benchmarks and numerical studies. We implemented the simulation algorithms in `C++17` (see [11]) using the `C++` interface of the R-package `Rcpp` (see [6]). We used the R-package `bench` (see [10]) for our benchmark studies.

Appendix

A The infinitesimal generator matrix Q for special cases

This appendix contains the MDSM's generator matrix representations for the exchangeable subclass, the extendible subclass, and the Armageddon shock subclass. The following derivations constitute mostly tedious but straightforward calculations. However, we use their results indirectly in the main part of this article, e.g., to derive Eq. (16) or Proposition 5.3.

Remark A.1. For an exMO distribution with exchangeable shock-arrival intensities $\boldsymbol{\lambda} = \lambda \cdot \boldsymbol{p}$, we have

$$k_{IJ} = \begin{cases} \sum_{l=0}^{i} \binom{i}{l} p_{l+(j-i)} & I \subseteq J \quad \text{with} \quad |I| = i, \quad |J| = j, \\ 0 & \text{else}, \end{cases}$$

and

$$q_{IJ} = \begin{cases} -\sum_{k=1}^{d-i}\binom{d-i}{k}\sum_{l=0}^{i}\binom{i}{l}\lambda_{l+k} & I = J, \quad |I| = i, \\ \sum_{l=0}^{i}\binom{i}{l}\lambda_{l+(j-i)} & I \subsetneq J, \ |I| = i, \ |J| = j, \\ 0 & \text{else}. \end{cases}$$

Proof. We obtain both results from Eqs. (10) and (12) by using the exchangeability and by grouping all shock-arrival intensities, or shock-arrival probabilities, belonging to the same cardinalities. □

ExMO distributions are often reparametrized with so-called *d-monotone sequences* instead of the exchangeable shock-arrival intensities. The reparametrization was proposed in [13] and extensively studied by [14, Chp. 3]. The latter serves as a reference for the subsequently summarized results on this reparametrization. We define for the exchangeable shock-arrival intensities

$$a_k := \sum_{j=0}^{d-k-1}\binom{d-k-1}{j}\lambda_{j+1}, \quad k \in [d-1]_0 = \{0, \ldots, d-1\}.$$

Then, the sequence $a_0, a_1, \ldots, a_{d-1}$ is *d-monotone*, i.e.,

$$(-1)^{i-1}\Delta^{i-1}a_{d-i} = \sum_{j=0}^{i-1}(-1)^j\binom{i-1}{j}a_{d-i+j} \geq 0, \quad i \in [d].$$

In particular, [14, Chp. 3] shows that

$$\lambda_i = (-1)^{i-1}\Delta^{i-1}a_{d-i}, \quad i \in [d], \tag{A.1}$$

and, moreover, every d-monotone sequence defines an exMO distribution. Furthermore, [14, Chp. 4] shows that the reparametrization of an extMO distribution with Bernstein function ψ is

$$a_i = \psi(i+1) - \psi(i), \quad i \in [d-1]_0. \tag{A.2}$$

This reparametrization bridges the gap toward the parametrization of extMO distributions via Bernstein functions and has the following property: the first k sequence elements describe the law of k-margins, i.e., the subvector $(\tau_1, \ldots, \tau_k)$, $k \in [d]$, has the reparametrization $a_0, \ldots, a_{k-1}$ and the margin τ_1 is exponentially distributed with rate a_0. We have the following representation of the MDSM's Markov generator matrix in terms of the reparametrization:

Remark A.2. For an exMO distribution with reparametrization $\boldsymbol{a}$, we have

$$k_{IJ} = \frac{1}{\sum_{k=0}^{d-1}a_k}\begin{cases} \sum_{k=0}^{d-1}a_k - \sum_{k=0}^{d-i-1}a_k & I = J, \quad |I| = i, \\ (-1)^{j-i-1}\Delta^{j-i-1}a_{d-j} & I \subsetneq J, \quad |I| = i, \quad |J| = j, \\ 0 & \text{else}, \end{cases}$$

and

$$q_{IJ} = \begin{cases} -\sum_{k=0}^{d-i-1}a_k & I = J, \quad |I| = i, \\ (-1)^{j-i-1}\Delta^{j-i-1}a_{d-j} & I \subsetneq J, \quad |I| = i, \quad |J| = j, \\ 0 & \text{else}. \end{cases}$$

For the proof of Remark A.2, we require the following auxiliary lemma:

Lemma A.3. *Let $d \in \mathbb{N}$, $c_0, c_1, \ldots, c_d$ be a real sequence, and $i, j \in [d]_0$ with $i \leq j$. Then,*

$$\sum_{k=0}^{i}\binom{i}{k}(-1)^{k+j-i-1}\Delta^{k+j-i}c_{d-k-(j-i)} = (-1)^{j-i-1}\Delta^{j-i}c_{d-j}. \tag{A.3}$$

Proof. We first proof the claim for $j - i = 0$, i.e.,

$$\sum_{k=0}^{i}\binom{i}{k}(-1)^{k-1}\Delta^{k}c_{d-k} = -c_{d-i}. \tag{A.4}$$

For this, we have

$$\sum_{k=0}^{i}\binom{i}{k}(-1)^{k-1}\Delta^{k}c_{d-k} \overset{(*)}{=} \sum_{k=0}^{i}\binom{i}{k}\sum_{l=0}^{k}(-1)^{l-1}\binom{k}{l}c_{d-k+l} = \sum_{m=0}^{i} c_{d-m}\underbrace{\sum_{\substack{k\in[i]_0,\ l\in[k]_0\\ k-l=m}}(-1)^{l-1}\binom{i}{k}\binom{k}{l}}_{\overset{(\dagger)}{=}-\mathbb{1}_{\{i=m\}}} = -c_{d-i},$$

where $(*)$ follows from [14, Lem. 2.5.2] and $(\dagger)$ follows with

$$\sum_{\substack{k\in[i]_0,\ l\in[k]_0\\ k-l=m}}(-1)^{l-1}\binom{i}{k}\binom{k}{l} = \sum_{l=0}^{i-m}(-1)^{l-1}\binom{i}{l+m}\binom{l+m}{l} = \sum_{l=0}^{i-m}(-1)^{l-1}\binom{i}{m}\binom{i-m}{l} = -\mathbb{1}_{\{i=m\}}.$$

For the general statement, we have

$$\sum_{k=0}^{i}\binom{i}{k}(-1)^{k+j-i-1}\Delta^{k+j-i}c_{d-k-(j-i)} = \sum_{k=0}^{i}\binom{i}{k}(-1)^{k-1}\Delta^{k}[(-1)^{j-i}\Delta^{j-i}c_{d-k-(j-i)}] \overset{\text{Eq. (A.4)}}{=} -[(-1)^{j-i}\Delta^{j-i}c_{d-j}].$$

□

Proof of Remark A.2. Let $\varnothing \neq I \subseteq J \subseteq [d]$ with $|I| = i$ and $|J| = j$ and recall from Eq. (A.1) that $\lambda_i = (-1)^{i-1}\Delta^{i-1}a_{d-i}$, $i \in [d]$. Hence, we have for $i < j$

$$\begin{aligned}
\lambda k_{IJ} &\overset{\text{Remark (A.1)}}{=} \sum_{k=0}^{i}\binom{i}{k}\lambda_{k+j-i} = \sum_{k=0}^{i}\binom{i}{k}(-1)^{k+j-i-1}\Delta^{k+j-i-1}a_{d-k-(j-i)}\\
&\overset{(*)}{=} \sum_{k=0}^{i}\binom{i}{k}(-1)^{k+j-i-1}\Delta^{k+j-i}\left[\sum_{l=1}^{d-k-(j-i)}a_{l-1}\right]\\
&\overset{\text{Lemma (A.3)}}{=} (-1)^{j-i-1}\Delta^{j-i}\left[\sum_{l=1}^{d-j}a_{l-1}\right]\\
&= (-1)^{j-i-1}\Delta^{j-i-1}a_{d-j},
\end{aligned}$$

where we assume in $(*)$ that the finite forward difference operator is applied to the sequence $0, a_0, a_0 + a_1, \ldots, \sum_{k=1}^{d}a_{k-1}$. Furthermore, we have from [17, p. 134]

$$\lambda = \sum_{k=0}^{d-1}a_k.$$

Finally, we obtain the claim after obtaining with a similar calculation

$$\lambda k_{II} = \sum_{l=0}^{d-1}a_l - \sum_{l=0}^{d-i-1}a_l.\square$$

Remark A.4. For an extMO distribution with Bernstein function ψ, we have

$$k_{IJ} = \frac{1}{\psi(d)}\begin{cases}\psi(d) - \psi(d-i) & I = J,\quad |I| = i,\\ (-1)^{j-i-1}\Delta^{j-i}\psi(d-j) & I \subsetneq J,\quad |I| = i,\quad |J| = j,\\ 0 & \text{else},\end{cases}$$

and

$$q_{IJ} = \begin{cases} -\psi(d-i) & I = J, \quad |I| = i, \\ (-1)^{j-i-1}\Delta^{j-i}\psi(d-j) & I \subsetneq J, \quad |I| = i, \quad |J| = j, \\ 0 & \text{else}. \end{cases}$$

Proof. The claim follows from Remark A.2 with Eq. (A.2). □

Remark A.5. For an extMO distribution from the Armageddon family with parameters $\alpha, \beta \geq 0$ with $\alpha + \beta > 0$, we have

$$k_{IJ} = \frac{1}{\beta + d\alpha} \begin{cases} \beta + d\alpha & I = J, \quad |I| = d, \\ i\alpha & I = J, \quad |I| < d, \\ \alpha & I \subsetneq J, \quad |I| = i, \quad |J| = j, \quad j - i = 1, \quad j < d, \\ \alpha + \beta & I \subsetneq J, \quad |I| = i, \quad |J| = j, \quad j - i = 1, \quad j = d, \\ \beta & I \subsetneq J, \quad |I| = i, \quad |J| = j, \quad j - i > 1, \quad j = d, \\ 0 & \text{else}, \end{cases}$$

and

$$q_{IJ} = \begin{cases} -\beta - (d-i)\alpha & I = J, \quad |I| < d, \\ \alpha & I \subsetneq J, \quad |I| = i, \quad |J| = j, \quad j - i = 1, \quad j < d, \\ \alpha + \beta & I \subsetneq J, \quad |I| = i, \quad |J| = j, \quad j - i = 1, \quad j = d, \\ \beta & I \subsetneq J, \quad |I| = i, \quad |J| = j, \quad j - i > 1, \quad j = d, \\ 0 & \text{else}. \end{cases}$$

Proof. The claim follows for $k \geq 1$ with the following calculations:

$$\psi(x) = \beta \mathbb{1}_{\{x>0\}} + \alpha x, \Delta\psi(x) = \beta \mathbb{1}_{\{x=0\}} + \alpha,$$

and

$$(-1)^{k-1}\Delta^k\psi(x) = \beta \mathbb{1}_{\{x=0\}}.$$ □

B The infinitesimal generator matrix Q^* for special cases

This appendix contains the MDCM's generator matrix representations for the exchangeable reparametrization, the extendible subclass, and the Armageddon shock subclass. The results are all direct corollaries from Appendix A and Theorem 4.1.

Remark B.1. For an exMO distribution with reparametrization $\boldsymbol{a}$, we have

$$k^*_{ij} = \frac{\binom{d-i}{j-i}}{\sum_{k=0}^{d-1} a_k} \cdot \begin{cases} \sum_{k=0}^{d-1} a_k - \sum_{k=0}^{d-i-1} a_k & i = j, \\ (-1)^{j-i-1}\Delta^{j-i-1}a_{d-j} & i < j, \\ 0 & \text{else}, \end{cases}$$

and

$$q^*_{ij} = \binom{d-i}{j-i} \cdot \begin{cases} -\sum_{k=0}^{d-i-1} a_k & i = j, \\ (-1)^{j-i-1}\Delta^{j-i-1}a_{d-j} & i < j, \\ 0 & \text{else}. \end{cases}$$

Remark B.2. For an extMO distribution with Bernstein function ψ, we have

$$k^*_{ij} = \frac{\binom{d-i}{j-i}}{\psi(d)} \cdot \begin{cases} \psi(d) - \psi(d-i) & i = j, \\ (-1)^{j-i-1}\Delta^{j-i}\psi(d-j) & i < j, \\ 0 & \text{else}, \end{cases}$$

and

$$q^*_{ij} = \binom{d-i}{j-i} \cdot \begin{cases} -\psi(d-i) & i = j, \\ (-1)^{j-i-1}\Delta^{j-i}\psi(d-j) & i < j, \\ 0 & \text{else}. \end{cases}$$

Remark B.3. For an extMO distribution from the Armageddon family with parameters $\alpha, \beta \geq 0$ with $\alpha + \beta > 0$, we have

$$k^*_{ij} = \frac{\binom{d-i}{j-i}}{\beta + d\alpha} \cdot \begin{cases} \beta + d\alpha & i = j = d, \\ i\alpha & i = j < d, \\ \alpha & i + 1 = j < d, \\ \alpha + \beta & i + 1 = j = d, \\ \beta & i + 1 < j = d, \\ 0 & \text{else}, \end{cases}$$

and

$$q^*_{ij} = \binom{d-i}{j-i} \cdot \begin{cases} -\beta - (d-i)\alpha & i = j < d, \\ \alpha & i + 1 = j < d, \\ \alpha + \beta & i + 1 = j = d, \\ \beta & i + 1 < j = d, \\ 0 & \text{else}. \end{cases}$$

C Runtime boundaries for the MDCM and LFM

The following section compares the runtime of the MDCM algorithm to that of the LFM algorithm on a theoretical basis. For this, recall that the MDCM samples discrete transitions with conditionally independent exponential waiting times until all components are extinct, determining the order with a random shuffling afterward, and the LFM samples discrete transitions with exponential waiting times until the compound Poisson subordinator surpassed all unit exponential barrier values. We provide formulas to calculate or bound the expected number of sampled waiting times for both models. These confirm that the expected number of waiting times is bounded for the MDCM but can become arbitrarily large for the LFM.

Proposition C.1. *Let $d \geq 2$ and consider the LFM from* Theorem 2.4 *for an extMO distribution with Bernstein function $\psi(x) = c(1 - \kappa(x))$ for $c > 0$ and a completely monotone function κ*, i.e., *ψ is the Lévy exponent of a pure-jump compound Poisson subordinator with intensity c and whose jumps have the Laplace function κ. Furthermore, let K_i be the number of jumps required for surpassing the ith barrier value and K be the number of jumps required for surpassing all barrier values. Then, K_i is geometrically distributed with success probability $1 - \kappa(1)$ and*

$$\frac{1}{1 - \kappa(1)} \leq \mathbb{E}[K] \leq \frac{d}{1 - \kappa(1)}.$$

Proof. We use the notation from Theorem 2.4, and let Λ be defined by

$$\Lambda_t = \sum_{j=1}^{N_t} X_j, \quad t \geq 0,$$

for a Poisson process N and, independent thereof, iid jumps $\{X_j : j \in \mathbb{N}\}$ with Laplace function κ, and define

$$K_i := \min\left\{k \in \mathbb{N} : E_i \leq \sum_{j=1}^{k} X_j\right\}, \quad i \in [d].$$

We can derive with a simple calculation that K_i, $i \in [d]$, have geometric distributions with success probability $1 - \kappa(1)$ by using the tower property and conditioning on $\{X_j : j \in \mathbb{N}\}$ for the survival function of K_i. Consequently, we obtain the claim by using $K = \max\{K_1, \ldots, K_d\}$ and

$$\frac{1}{1-\kappa(1)} = \mathbb{E}[K_1] \leq \mathbb{E}[K] \leq \sum_{i=1}^{d} \mathbb{E}[K_i] = d \cdot \mathbb{E}[K_1] = \frac{d}{1-\kappa(1)}. \qquad \square$$

Proposition C.2. *Let $d \geq 2$ and consider the MDCM from* Theorem 4.1 *for an infinitesimal generator matrix Q^*. Furthermore, let M be the number of transitions until the Markov chain process is absorbed, i.e., until all components are dead. Then*

$$\mathbb{E}[M] = \overrightarrow{e}_0^{\top} S \cdot \mathbf{1} \leq d,$$

where $S = (s_{ij})_{i,j \in [d-1]_0}$ is the fundamental matrix of the embedded Markov chain, (implicitly) defined by

$$(S^{-1})_{ij} = \begin{cases} 1 & i = j, \\ \dfrac{q^*_{ij}}{q^*_{ii}} & i \neq j. \end{cases}$$

Proof. First note that the transition matrix $T = (t_{ij} : i, j \in [d]_0)$ of the embedded Markov process if defined by

$$t_{ij} = \begin{cases} 1 & i = j = d, \\ -\dfrac{q^*_{ij}}{q^*_{ii}} & j > i, \\ 0 & j \leq i, \end{cases}$$

and that d is the only recurrent, absorbing state. By using [4, Thm. 6.2.3], we conclude that $S = [\mathrm{Id} - T]^{-1}$ is the *fundamental matrix* of the embedded chain and obtain the claim using the arguments from [4, Sec. 6.2.1]:

$$\mathbb{E}[M] = \sum_{j=0}^{d-1} \sum_{n=0}^{\infty} (T^n)_{0j} = \sum_{j=0}^{d-1} \left(\sum_{n=0}^{\infty} T^n \right)_{0j} = \sum_{j=0}^{d-1} s_{0j}.$$

We obtain the upper bound from the observation that each transition increases the death count by at least one, limiting the number of transitions until absorption by d. $\square$

We conclude from Propositions C.1 and C.2 that while the MDCM cannot have more than d transitions, the expected number of transitions in the LFM is not bounded and can be significantly larger when the probability of small jumps in the corresponding Lévy subordinator is high. We created Figure A1 to demonstrate these differences for the exponential family. The plot highlights that the MDCM requires significantly fewer expected transitions than the LFM if the expected jump sizes are tiny.

D Statistical tests for extMO distributions

A challenge for implementing simulation algorithms is developing statistical tests to verify their goodness-of-fit. For example, R uses the *Dvoretzky–Kiefer–Wolfowitz inequality* with the tight constant derived in [20] for testing their univariate distribution algorithms. Other popular tests include the *Kolmogorov–Smirnov tests* and the χ^2*-test*. For multivariate distributions, a recent advance is the derivation of a tight constant for the multivariate *Dvoretzky–Kiefer–Wolfowitz inequality*, see [21].

We propose a simple alternative approach using the *min-stability* of MO distribution: Consider iid d-variate extMO distributed random vectors $\boldsymbol{\tau}_k = (\tau_{k,1}, \ldots, \tau_{k,d})$, $k \in \{1, \ldots, n\}$, with corresponding Bernstein function ψ and define

$$U_k \coloneqq 1 - \exp\{-\psi(d) \cdot \min_{i \in [d]} \tau_{k,i}\}, \quad k \in [n].$$

A straightforward calculation shows that the overall minimum of $\boldsymbol{\tau}_1$ has an exponential distribution with rate $\psi(d)$. Hence, $U_1, \ldots, U_n$ is an iid standard uniform sequence. Consequently, we can apply *Kolmogorov–Smirnov tests* for goodness-of-fit testing, see [18] and [29]. We employed this method to extensively test all implementations discussed in Section 6 using a p-value threshold of 1% and a *Bonferroni correction* for the total number of tests.

E Alias method for sampling on finite-state spaces

The following section shortly sketches the *alias method* for sampling discrete random variables on finite-state spaces; see [31] for the details. Let $\{p_k : k \in \{1, \ldots, n\}\}$ be a probability counting measure for the finite-state space $[n] = \{1, \ldots, n\}$. Suppose there exist mappings $f : [n] \mapsto [n]$ and $q : [n] \mapsto [0, 1]$ such that

$$p_k = \frac{1}{n} \cdot \left[(1 - q(k)) + \sum_{j \in [n]} q(j) \mathbb{1}_{\{f(j)=k\}} \right]. \tag{$*$}$$

Consider a probability space supporting the following random variables: a uniform random variable Y on $[n]$ and, independent thereof, a uniform random variable U on $[0, 1]$. Define I by $I = \mathbb{1}_{\{U \le q(Y)\}}$ such that I conditioned on Y has a Bernoulli distribution with success probability $q(Y)$. Furthermore, define X by

$$X \coloneqq \begin{cases} f(Y) & I = 1, \\ Y & I = 0, \end{cases}$$

and a straightforward calculation shows that $\mathbb{P}(X = k) = p_k$.

Note that the mappings f and q fulfilling condition $(*)$ exist but are not unique. We sketch one possibility in Algorithm 2 that uses a strategy that recursively moves probability mass from the element

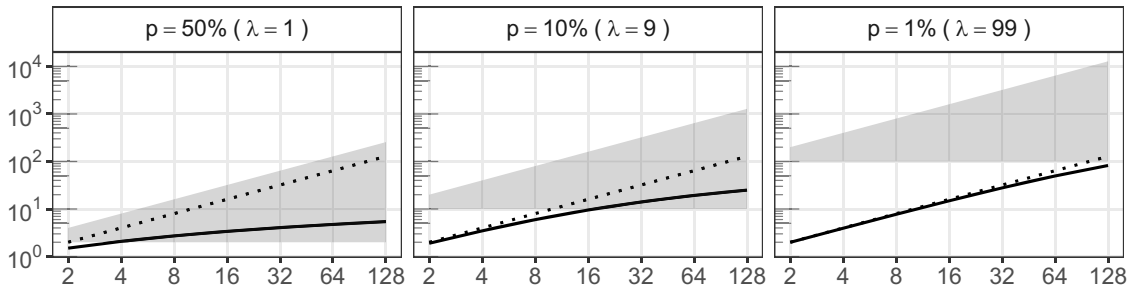


Figure A1: The expected number of transitions in the MDCM (solid line) and expected number of transitions in the LFM (gray area) for the Exponential family; the dotted line marks the identity.

with the largest mass to that with the least mass until all probabilities are equal. By reflecting these probability mass transfers adequately in the success probabilities of the second-order Bernoulli experiment and conditional element-switch, the overall probabilities remain unchanged.

Algorithm 2. Create mappings f and q for the alias method.

```
Require: (Vector) argument p with probabilities
procedure CREATE_ALIAS_MAPPINGS(p)
  n = size(p)
  w = p
  f = sequence(1, n), q = zeros(n)
  while max(w) > 1. / n do
    l = argmin(w), k = argmax(w)
    f[l] = k
    q[l] = (1. / n - w[l]) * n
    w[k] -= (1. / n - w[l])
    w[l] = 1. / n
  end while
  return f, q
end procedure
```

The appeal of the alias method is that, given the mappings f and q, it requires at most the generation of two uniform random variables on $\{1, \ldots, n\}$ and $[0, 1]$, at most two reference operations, and a comparison per sample.

References

[1] Aldous, D. J. (1985). Exchangeability and related topics. *École d'Été de Probabilités de Saint-Flour XIII – 1983*. Berlin/Heidelberg, Germany: Springer. doi: 10.1007/bfb0099421.

[2] Arnold, B. C. (1975). A characterization of the exponential distribution by multivariate geometric compounding. *Sankhyā: The Indian Journal of Statistics, Series A*, *37*(1), 164–173, 0581-572X.

[3] Berg, C., Christensen, J. P. R. & Ressel, P. (1984). *Harmonic analysis on semigroups*. New York, NY: Springer. doi: 10.1007/978-1-4612-1128-0.

[4] Brémaud, P. (2020). *Markov chains* (2nd ed.). Cham, Switzerland: Springer. doi: 10.1007/978-3-030-45982-6.

[5] Brigo, D., Mai, J.-F. & Scherer, M. (2016). Markov multivariate survival indicators for default simulation as a new characterization of the Marshall–Olkin law. *Statistics & Probability Letters*, *114*, 60–66. doi: 10.1016/j.spl.2016.03.013.

[6] Eddelbuettel, D. & François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, *40*(8), 1–18. doi: 10.18637/jss.v040.i08.

[7] Fernández, L. (2015). *Selected topics in financial engineering: First-exit times and dependence structures of Marshall–Olkin kind* (PhD thesis). Bilbao, Spain: University of the Basque Country. https://addi.ehu.es/handle/10810/16067.

[8] Galassi, M., Davies, J., Theiler, J., Gough, B., Jungman, G., Alken, P., ..., Ulerich, R. (2009). GNU scientific library reference manual (3rd ed.). Network theory Ltd. ISBN: 978-0-9546120-7-8.

[9] Gautschi, W. (2012). *Numerical analysis* (2nd ed.). Boston, MA: Birkhäuser. doi: 10.1007/978-0-8176-8259-0.

[10] Hester, J. & Vaughan, D. (2021). *bench: High precision timing of R expressions, R package version 1.1.2*. https://CRAN.R-project.org/package=bench.

[11] ISO C++ Standards Committee. (2017). *ISO/IEC 14882:2017 Programming languages – C*. International Organization for Standardization (ISO).

[12] IEEE-SA Standards Board. (2017). *IEEE standard for floating-point arithmetic*. IEEE Std 754-2019 (Revision of IEEE 754-2008), 1–84. doi: 10.1109/IEEESTD.2019.8766229.

[13] Mai, J.-F. & Scherer, M. (2009). Lévy-frailty copulas. *Journal of Multivariate Analysis*, *100*(7), 1567–1585. doi: 10.1016/j.jmva.2009.01.010.

[14] Mai, J.-F. (2010). *Extendibility of Marshall–Olkin distributions via Lévy subordinators and an application to portfolio credit risk* (PhD thesis). Munich, Germany: Technical University of Munich. https://nbn-resolving.org/urn:nbn:de:bvb:91-diss-20100628-969547-1-8.
[15] Mai, J.-F. & Scherer, M. (2013). Sampling exchangeable and hierarchical Marshall–Olkin distributions. *Communications in Statistics-Theory and Methods*, *42*(4), 619–632. doi: 10.1080/03610926.2011.615437.
[16] Mai, J.-F. (2014). *Multivariate exponential distributions with latent factor structure and related topics* (Habilitation). Munich, Germany: Technical University of Munich. https://nbn-resolving.org/urn:nbn:de:bvb:91-diss-20140220-1236170-0-7.
[17] Mai, J.-F. & Scherer, M. (2017). *Simulating copulas: Stochastic models, sampling algorithms and applications* (2nd ed.). Singapore: World Scientific Publishing. doi: 10.1142/10265.
[18] Marsaglia, G., Tsang, W. W. & Wang, J. (2003). Evaluating Kolmogorov's distribution. *Journal of Statistical Software*, *8*(18), 1–4. doi: 10.18637/jss.v008.i18.
[19] Marshall, A. W. & Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association*, *62*(317), 30–44. doi: 10.1080/01621459.1967.10482885.
[20] Massart, P. (1990). The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality. *The Annals of Probability*, *18*(3), 1269–1283. doi: 10.1214/aop/1176990746.
[21] Naaman, M. (2021). On the tight constant in the multivariate Dvoretzky–Kiefer–Wolfowitz inequality. *Statistics & Probability Letters*, *173*, 1–8. doi: 10.1016/j.spl.2021.109088.
[22] Nelsen, R. B. (2006). *An introduction to copulas* (2nd ed.). New York, NY: Springer. doi: 10.1007/0-387-28678-0.
[23] Piessens, R., de Doncker-Kapenga, E., Überhuber, C. W. & Kahaner, D. K. (1983). *Quadpack*. Berlin/Heidelberg, Germany: Springer. doi: 10.1007/978-3-642-61786-7.
[24] R Core Team. (2021). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.
[25] Resnick, S. I. (1987). *Extreme values, regular variation, and point processes*. New York, NY: Springer. doi: 10.1007/978-0-387-75953-1.
[26] Ruckdeschel, P., Kohl, M., Stabla, T. & Camphausen, F. (2006). S4 classes for distributions. *R News*, *6*(2), 2–6. https://CRAN.R-project.org/doc/Rnews/Rnews_2006-2.pdf.
[27] Sato, K.-I. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge, UK: Cambridge University Press.
[28] Schilling, R. L., Song, R. & Vondracek, Z. (2012). *Bernstein functions* (2nd ed.). Berlin/Boston: De Gruyter. doi: 10.1515/9783110269338.
[29] Simard, R. & L'Ecuyer, P. (2011). Computing the two-sided Kolmogorov-Smirnov distribution. *Journal of Statistical Software*, *39*(11), 1–18. doi: 10.18637/jss.v039.i11.
[30] Sun, Y., Mendoza-Arriaga, R. & Linetsky, V. (2017). Marshall–Olkin distributions, subordinators, efficient simulation, and applications to credit risk. *Advances in Applied Probability*, *49*(2), 481–514. doi: 10.1017/apr.2017.10.
[31] Walker, A. J. (1977). An efficient method for generating discrete random variables with general distributions. *ACM Transactions on Mathematical Software*, *3*(3), 253–256. doi: 10.1145/355744.355749.

B Further articles as primary author

B.1 Exogenous shock models: analytical characterization and probabilistic construction

Exogenous shock models: analytical characterization and probabilistic construction

Matthias Scherer and Henrik Sloot

The article [3] provides a characterization theorem for the survival functions of multivariate failure-time distributions arising from generalized exogenous shock models with independent, nonnegative, unbounded, and continuous shock-arrival times. Generalized exogenous shock models define a random vector via independent shock-arrival times for each subset of elements, destroying all components in that subset not already destroyed by this time. We identify the corresponding survival functions as continuous functions, which are products of their nonincreasingly ordered and individually transformed arguments with order-dependent transformations fulfilling specific monotonicity requirements. The characterization theorem extends an existing result for exchangeable generalized Marshall–Olkin survival copulas. Additionally, we outline applications of linking generalized Marshall–Olkin distributions to the exogenous shock model.

The introduction recalls the *classical* exogenous shock model with exponential shock-arrival times and the associated *classical* Marshall–Olkin distribution. Following, we outline previous research efforts to generalize Marshall–Olkin distributions. Finally, we conclude the introduction with the generalization most relevant for us: allowing independent, nonnegative, unbounded, and continuous shock-arrival times under an exchangeability assumption. For this, previous research identified the corresponding distributions' survival functions as continuous functions, which are the products of their nonincreasingly ordered and individually transformed arguments with order-independent transformations fulfilling specific monotonicity requirements.

Section 2 contains the characterization theorem. We motivate generalizing exponential by nonnegative, unbounded, and continuous shock time distributions with the hazard rate analogy: the former corresponds to a constant hazard rate and the latter virtually to time-dependent hazard rates. A condensed version of the characterization is the following: A continuous function $\bar{F} : \mathbb{R}^d \to \mathbb{R}$ of the form

$$\bar{F}(\boldsymbol{t}) = \prod_{i=1}^{d} g_i^{\pi}(t_{\pi(i)}), \quad \boldsymbol{t} \geq 0,\ \pi \in \mathcal{S}_d,\ t_{\pi(1)} \geq \cdots \geq t_{\pi(d)},$$

is a d-variate survival function if and only if the order-dependent transformations g_i^{π} fulfill specific monotonicity criteria. Moreover, if they do, the survival function corresponds to an generalized exogenous shock model whose shock-arrival time distributions follow from said monotonicity criteria. We conclude this section by explaining the more complex expressions involved in the characterization theorem.

Section 3 presents several applications. They share using the explicit formula for shock-arrival time distributions, allowing to calculate shock-arrival-time-related probabilities for generalized Marshall–Olkin distributed vectors not explicitly specified via an generalized exogenous shock model. Examples of such an alternative specification are previous hierarchical factor models based on the additive frailty model construction. The first application uses partial integration to calculate joint default probabilities. We shortly discuss this approach's viability and practical limits by summarizing a numerical study about implementing these partial integration approaches. The second application links the probability of the first shock's cardinality to *decrement matrices* of *regenerative composition structures* discussed in [70].

Statement of individual contribution

I, Henrik Sloot, am the primary author of this article. I am responsible for the conceptualization, the proofs, the software and the analysis for the validation mentioned in Section 3, and the writing. The role of Matthias Scherer was that of a supervisor, giving feedback on results, the written drafts, related literature, and double-checking results. He also initially suggested extending [25, Theorem 1.1] for the non-exchangeable case. The article originated from my master thesis [87], and Theorem 1 of the article and the corresponding proof, with minor editorial changes, have already been stated therein. However, the applications described

in Section 3 and the mentioned computations were found after completing my Master's thesis and were written specifically for this article. For my doctoral thesis, to explicitly prevent double counting, only those parts of this article that go beyond my Master's thesis are to be accredited.

Addendum

For clarity this thesis, including the preceding summary, use the term "generalized exogenous shock model" while the article uses the term "exogenous shock model".

Erratum

The original publication contained several minor errors.

- Eq. (6) incorrectly equates $\bar{F}(\boldsymbol{t})$ to $\prod_{i=1}^{d} \tilde{g}^{\pi(i,\dots,d),\pi(i)}(t_{\pi(i)})$; the correct version of the latter reads $\prod_{i=1}^{d} \tilde{g}^{\pi(1,\dots,i),\pi(i)}(t_{\pi(i)})$.
- The sentence after Eq. (7) incorrectly requires that "g_1^{π} and $\tilde{g}^{[d],m}$, respectively, are in the respective subclass with no atoms at infinity $\bar{\mathcal{G}}_1$"; the correct version reads "g_1^{π} and $\tilde{g}^{m,m}$, $m \in [d]$, respectively, are in the respective subclass with no atoms at infinity $\bar{\mathcal{G}}_1$".
- Part 3 of Theorem 1 incorrectly requires $I_1 \cap I_2 \neq \emptyset$; the correct requirement reads $I_1 \cap I_2 = \emptyset$.
- The last paragraph of Section 3 incorrectly conflates n and d at several places; the correct version replaces all all n by d for all but the first occurrences.

Permission to include

M. Scherer and H. Sloot. “Exogenous shock models: analytical characterization and probabilistic construction”. In: *Metrika* 82.8 (2019), pp. 931–959. DOI: 10.1007/s00184-019-00715-8

Reprinted according to the author's rights for publishing with Springer;[1] see archived excerpt below.

Springer

Permissions

Get permission to reuse Springer Nature content

Springer Nature is partnered with the Copyright Clearance Center to meet our customers' licensing and permissions needs.

Copyright Clearance Center's RightsLink® service makes it faster and easier to secure permission for the reuse of Springer Nature content to be published, for example, in a journal/magazine, book/textbook, coursepack, thesis/dissertation, annual report, newspaper, training materials, presentation/slide kit, promotional material, etc.

Simply visit SpringerLink and locate the desired content;

Go to the article or chapter page you wish to reuse content from. (Note: permissions are granted on the article or chapter level, not on the book or journal level). Scroll to the botton of the page, or locate via the side bar, the "Reprints and Permissions" link at the end of the chapter or article.

Select the way you would like to reuse the content;

Complete the form with details on your intended reuse. Please be as complete and specific as possible ao as not to delay your permission request;

Create an account if you haven't already. A RightsLink account is different than a SpringerLink account, and is necessary to receive a licence regardless of the permission fee. You will receive your licence via the email attached to your RightsLink receipt;

Accept the terms and conditions and you're done!

For questions about using the RightsLink service, please contact Customer Support at Copyright Clearance Center via phone +1-855-239-3415 or +1-978-646-2777 or email springernaturesupport@copyright.com.

How to obtain permission to reuse Springer Nature content not available online on SpringerLink

Requests for permission to reuse content (e.g. figure or table, abstract, text excerpts) from Springer Nature publications currently not available online must be submitted in writing. Please be as detailed and specific as possible about what, where, how much, and why you wish to reuse the content.

Your contacts to obtain permission for the reuse of material from:

- books: bookpermissions@springernature.com
- journals: journalpermissions@springernature.com

Author reuse

Please check the Copyright Transfer Statement (CTS) or Licence to Publish (LTP) that you have signed with Springer Nature to find further information about the reuse of your content.

Authors have the right to reuse their article's Version of Record, in whole or in part, in their own thesis. Additionally, they may reproduce and make available their thesis, including Springer Nature content, as required by their awarding academic institution. Authors must properly cite the published article in their thesis according to current citation standards.
Material from: 'AUTHOR, TITLE, JOURNAL TITLE, published [YEAR], [publisher - as it appears on our copyright page]'

If you are any doubt about whether your intended re-use is covered, please contact journalpermissions@springernature.com for confirmation.

Self-Archiving

- Journal authors retain the right to self-archive the final accepted version of their manuscript. Please see our self-archiving policy for full details:
https://www.springer.com/gp/open-access/authors-rights/self-archiving-policy/2124

- Book authors please refer to the information on this link:
https://www.springer.com/gp/open-access/publication-policies/self-archiving-policy

[1] https://www.springer.com/gp/rights-permissions/obtaining-permissions/882.

Metrika (2019) 82:931–959
https://doi.org/10.1007/s00184-019-00715-8

Exogenous shock models: analytical characterization and probabilistic construction

Matthias Scherer[1] · **Henrik Sloot**[1]

Received: 16 August 2018 / Published online: 2 May 2019

Abstract
A new characterization for survival functions of multivariate failure-times arising in exogenous shock models with non-negative, continuous, and unbounded shocks is presented. These survival functions are the product of their ordered and individually transformed arguments. The involved transformations may depend on the specific order of the arguments and must fulfill a monotonicity condition. Conversely, every survival function of that form can be constructed using an exogenous shock model with independent and non-homogeneous shocks.

Keywords Exogenous shock model · Fatal shock model · Generalized Marshall–Olkin distribution · Multivariate survival function

1 Introduction

This work is concerned with the analytical characterization and probabilistic construction of multivariate probability laws of random vectors $(\tau_1, \ldots, \tau_d)$ on $\mathbb{R}_+^d$ arising from a fatal shock construction. The seminal model of this kind was presented in Marshall and Olkin (1967). Marshall and Olkin's main objective was to lift the lack-of-memory property to the d-variate case, an ansatz implying a distinct family of survival functions that can be constructed using a fatal shock model involving $2^d - 1$ independent and exponentially distributed shocks. More precisely, the failure time of component $i \in \{1, \ldots, d\} =: [d]$ is defined as

$$\tau_i := \min\{Z_I : i \in I \subseteq [d]\}, \quad i \in [d], \tag{1}$$

Matthias Scherer
scherer@tum.de

Henrik Sloot
henrik.sloot@tum.de

[1] Lehrstuhl für Finanzmathematik, Technische Universität München, Parkring 11, 85748 Garching, Germany

where Z_I, $\emptyset \neq I \subseteq [d]$, are independent exponentially distributed random variables with rates λ_I, $\emptyset \neq I \subseteq [d]$.

Taking the eponymous Marshall–Olkin construction from Eq. (1) as a starting point, various generalizations are possible.[1] Firstly, the operation 'min' might be altered, see Joe (1997, Chp. 4.6) for a general concept for constructing multivariate distributions based on a convolution-closed, infinitely divisible class of univariate distributions, which can be used to construct multivariate normal distributions as well as Marshall–Olkin distributions. Second, the assumption of shocks being independent can be dropped, leading for instance to the class of *Archimax copulas*, also called *scale-mixtures of Marshall–Olkin*, which assume an Archimedean dependence for the Z_I, see Li (2009). Third, and this is the path we pursue, shock distributions other than the exponential law can be considered. This has already been considered for the bivariate case, see Durante et al. (2008) and Li and Pellerey (2011) as well as for the exchangeable d-variate case, see Durante et al. (2007) and Mai et al. (2016). An interesting result, that was derived in Muliere and Scarsini (1987), is that the class of distributions, which is characterized by a modified lack-of-memory property, where the generic addition is replaced by a reducible and associative binary operator, is a subgroup of generalized Marshall–Olkin (GMO) distributions with shock survival functions of the form $\exp\{-\lambda_I H(t)\}$, cf. Eq. (5). In any of the above cases (or combinations thereof), the price to pay for the additional flexibility is a reduction in mathematical tractability. Deriving the survival function of a generalized d-variate fatal shock model and analyzing its properties is a challenging task. Beyond that, the inverse membership-testing problem, i.e. deciding if a given survival function admits a shock-model representation, is much harder. Hence, it is not surprising that the bivariate case was investigated first, see Marshall (1996) and Durante et al. (2008), followed by cases where the complexity is reduced by a reduction in the amount of considered shocks, see Durante et al. (2007), or via some symmetry assumption, see Marshall (1996) and Schenk (2016). In Lin and Li (2014), many properties of generalized Marshall–Olkin distributions, e.g. the corresponding copulae and coefficients of tail-dependence, are derived.

The main achievement of the present manuscript is Theorem 1. It fully characterizes the class of survival functions arising as a particular instance of a fatal shock model with independent shocks. This characterization is analytic on the one hand, translating the tedious d-increasingness property to a more convenient monotonicity property, and probabilistic on the other hand, establishing precisely how the $2^d - 1$ shock distributions must be selected to ultimately arrive at the model under consideration.

Closest to the present work is Schenk (2016), where it is shown that an exchangeable function C mapping $\boldsymbol{u} \in [0,1]^d$ to $[0,1]$, defined via a permutation $\pi \in \mathcal{S}_d$ with $u_{\pi(1)} \leq \cdots \leq u_{\pi(d)}$, of the form

$$C(\boldsymbol{u}) = u_{\pi(1)} \cdot \delta_2(u_{\pi(2)}) \cdot \cdots \cdot \delta_d(u_{\pi(d)}) \tag{2}$$

is a copula if and only if the functions $\{\delta_2, \ldots, \delta_d\}$ fulfill certain monotonicity conditions. This extends the bivariate case treated in Durante et al. (2008). Conversely, all

[1] The functional equation of the lack-of-memory property is another starting point for generalizations, see Muliere and Scarsini (1987).

copulas of the form Eq. (2) admit a stochastic representation as the survival copula of an exchangeable exogenous shock model, i.e. the shock distribution is equal for any two shocks Z_I and Z_J sharing the cardinality of their referencing sets $|I| = |J|$. In our analysis we work with survival functions and restrain ourselves from resorting to copulas, as Sklar's separation, see Sklar (1959), is not as natural in the case of non-exchangeable shock models as it is for exchangeable ones.

To emphasize the relevance of the present study, let us stress that the Marshall–Olkin distribution, mostly due to its embedded lack-of-memory property, arises like a focal point of many inner-mathematical problems. Beyond that, it has been applied in different fields, see Giesecke (2003), Lindskog and McNeil (2003) and Elouerkhaoui (2007), most of the applications having a survival time interpretation/model. For many real-world applications, however, the assumption of exponential shocks needs to be relaxed, see Bielecki et al. (2013) and Klein et al. (1989), and the resulting model is of the form that we classify with Theorem 1.

2 The generalized Marshall–Olkin distribution

The classical d-variate Marshall–Olkin distribution is parametrized by $2^d - 1$ constant *hazard rates*, $\lambda_I \geq 0, \emptyset \neq I \subseteq [d]$. These parameters are used as intensities[2] of the independent exponential shocks in construction Eq. (1), giving rise to the survival function

$$\mathbb{P}(\tau_1 > t_1, \ldots, \tau_d > t_d) = \bar{F}(\boldsymbol{t}) = \exp\left\{ - \sum_{\emptyset \neq I \subseteq [d]} \lambda_I \max_{i \in I} t_i \right\}, \quad \forall \boldsymbol{t} \geq 0. \tag{4}$$

One way of generalizing the Marshall–Olkin distribution is to consider time-dependent shock-intensities $s \mapsto \lambda_I(s)$, i.e.

$$\mathbb{P}(Z_I > t) = \bar{S}_I(t) = \exp\left\{ - \int_0^t \lambda_I(s) \mathrm{d}s \right\}, \quad \forall t \geq 0,$$

where $s \mapsto \lambda_I(s)$ is a non-negative function such that the involved integral is finite for all $t \geq 0$. In the following, this concept is slightly extended by solely demanding that the cumulative hazard rates $H_I(t) := -\log \bar{S}_I(t)$ are strictly positive, non-decreasing, zero in $t = 0$, and continuous. Particularly, atoms at infinity are allowed and the class of considered survival functions is

$$\bar{\mathcal{G}} := \left\{ \bar{S} : \mathbb{R}_+ \to (0, 1] : \bar{S}(0) = 1, \bar{S} \in C^{(0)}(\mathbb{R}_+), \mathrm{d}\bar{S} \leq 0 \right\}.$$

[2] The interpretation $\lambda_I = 0 \Leftrightarrow \mathbb{P}(Z_I = \infty) = 1$ requires the *marginal-finiteness condition*

$$\sum_{I \supseteq \{i\}} \lambda_I > 0, \quad \forall i \in [d], \tag{3}$$

to make the resulting vector $(\tau_1, \ldots, \tau_d)$ well defined.

For a set of survival functions $\bar{S}_I \in \bar{\mathcal{G}}, \emptyset \neq I \subseteq [d]$, with corresponding (cumulative) hazard rate functions H_I, fulfilling the *(generalized) marginal-finiteness condition*

$$\prod_{I \ni i} \bar{S}_I \in \bar{\mathcal{G}}_1 := \left\{ \bar{S} \in \bar{\mathcal{G}} : \lim_{t \to \infty} \bar{S}(t) \to 0 \right\}, \quad \forall i \in [d],$$

the corresponding survival function of a *generalized Marshall–Olkin (GMO) distribution* is

$$\bar{F}(\boldsymbol{t}) = \prod_{\emptyset \neq I \subseteq [d]} \bar{S}_I \left(\max_{i \in I} t_i \right) = \exp \left\{ - \sum_{\emptyset \neq I \subseteq [d]} H_I \left(\max_{i \in I} t_i \right) \right\}. \tag{5}$$

Note, that, with the (generalized) marginal-finiteness condition, the function in Eq. (5) is indeed the survival function of a real, non-negative random vector; this follows if an exogenous shock model with shock survival functions $\bar{S}_I, \emptyset \neq I \subseteq [d]$, is considered.

The survival function in Eq. (5) has an alternative, more compact, representation: Let $\boldsymbol{t} \geq 0$ and $\pi \in \mathcal{S}_d$ be a permutation such that $t_{\pi(1)} \geq \cdots \geq t_{\pi(d)}$; then, by reordering the factors appropriately, it follows that

$$\bar{F}(\boldsymbol{t}) = \prod_{i=1}^{d} g_i^{\pi}(t_{\pi(i)}) = \prod_{i=1}^{d} \tilde{g}^{\pi(\{i,\ldots,d\}),\pi(i)}(t_{\pi(i)}), \tag{6}$$

where for $i \in [d]$ and $\pi \in \mathcal{S}_d$ as well as $\emptyset \neq I \subseteq [d]$ and $m \in I$

$$g_i^{\pi}(t) := \prod_{I : \pi(i) \in I \subseteq \pi(\{i,\ldots,d\})} \bar{S}_I(t) \tag{7a}$$

and

$$\tilde{g}^{I,m}(t) := \prod_{J : I \cap J = \{m\}} \bar{S}_J(t). \tag{7b}$$

Furthermore, it follows that the factors g_i^{π} as well as $\tilde{g}^{I,m}$, respectively, are in the class of admissible survival functions $\bar{\mathcal{G}}$ and g_1^{π} as well as $\tilde{g}^{[d],m}$, respectively, are in the respective subclass with no atoms at infinity $\bar{\mathcal{G}}_1$.

The conclusion from the previous paragraph is, that survival functions of GMO-distributed random vectors are the product of their ordered and individually transformed arguments, i.e. functions of the form as presented in Eq. (6). The following theorem shows, among other things, that a survival function of this kind implies a stochastic representation as an exogenous shock model.[3]

[3] For readability, the necessary conditions on the transformations g_i^{π} are omitted here and the reader is referred to the full statement in Theorem 1.

Theorem 1 *Let $\bar{F} : \mathbb{R}_+^d \to \mathbb{R}$ be a continuous function having a representation as in Eq. (6) for an arbitrary family of functions $\{g_i^\pi : i \in [d], \pi \in \mathcal{S}_d\}$. If additionally*

- $g_1^\pi \in \bar{\mathcal{G}}_1 \ \forall \pi \in \mathcal{S}_d$ *and*
- $g_i^\pi(0) = 1 \ \forall i \in [d], \pi \in \mathcal{S}_d$,

then the following statements are equivalent:

1. *$\bar{F}$ is the survival function of a multivariate random vector $\boldsymbol{\tau} \in \mathbb{R}_+^d$.*
2. *For all $I_1, I_2 \subseteq [d]$ with $I_1 \cap I_2 = \emptyset$ and $I_2 \neq \emptyset$, let $\{\pi_J\}_{J \subseteq I_2} \subseteq \mathcal{S}_d$ be a family of permutations on $[d]$ which fulfills for each $J \subseteq I_2$ the following conditions*

   (a) $\pi_J(\{1, \ldots, |I_1|\}) = I_1$ *(if $I_1 \neq \emptyset$),*
   (b) $\pi_J(\{|I_1| + 1, \ldots, |I_1 \cup J|\}) = J$, *and*
   (c) $\pi_J(\{|I_1 \cup J| + 1, \ldots, |I_1 \cup I_2|\}) = I_2 \backslash J$.

   Define for $s \geq t \geq 0$

$$G_{I_1,I_2}^{\{\pi_J\}_{J \subseteq I_2}}(s,t) := \sum_{J \subseteq I_2} (-1)^{|J|} \prod_{j=1}^{|J|} g_{|I_1|+j}^{\pi_J}(s) \prod_{j=1}^{|I_2 \backslash J|} g_{|I_1 \cup J|+j}^{\pi_J}(t). \tag{8}$$

   Then $G_{I_1,I_2}^{\{\pi_J\}_{j \subseteq I_2}}$ does not depend on the specific family $\{\pi_J\}_{J \subseteq I_2}$ chosen; therefore, write G_{I_1,I_2}. Furthermore, $G_{I_1,I_2}(s,t)$ is non-negative and continuous in s and t.
3. *For all $I_1, I_2 \subseteq [d]$ with $I_1 \cap I_2 \neq \emptyset$ and $I_2 \neq \emptyset$ define for $m \in I_2$*

$$\bar{S}_{I_1,I_2}^m(t) := \prod_{i=1}^{|I_2|} \left(\prod_{\substack{J \subseteq I_2 \\ |J|=i, m \in J}} \tilde{g}^{J \cup I_1, m}(t) \right)^{(-1)^{i-1}}, \quad t \geq 0. \tag{9}$$

   Then $\bar{S}_{I_1,I_2}^m$ does not depend on the choice of m, i.e. $\bar{S}_{I_1,I_2}^m \equiv \bar{S}_{I_1,I_2}$, and $\bar{S}_{I_1,I_2} \in \bar{\mathcal{G}}$.
4. *For all $\emptyset \neq I \subseteq [d]$ and $m \in I$ define*

$$\bar{S}_I^m(t) := \prod_{i=1}^{|I|} \left(\prod_{\substack{J \subseteq I \\ |J|=i, m \in J}} \tilde{g}^{J \cup ([d] \backslash I), m}(t) \right)^{(-1)^{i-1}}, \quad t \geq 0. \tag{10}$$

   Then $\bar{S}_I^m$ does not depend on the choice of m, i.e. $\bar{S}_I^m \equiv \bar{S}_I$, and $\bar{S}_I \in \bar{\mathcal{G}}$.

Remark 1 Let the assumptions of Theorem 1 be fulfilled with $\bar{F}$ being the survival function of a random vector $\boldsymbol{\tau}$. Then $\boldsymbol{\tau}$ has a stochastic representation as an ESM with shock survival functions $\bar{S}_I$, i.e. if the $Z_I \sim \bar{S}_I, \emptyset \neq I \subseteq [d]$, are independent shocks and $\tilde{\boldsymbol{\tau}}$ is defined by Eq. (1), then $\boldsymbol{\tau} \overset{d}{=} \tilde{\boldsymbol{\tau}}$.

Due to the length of the required notation and the complexity of the theorem, giving an intuitive interpretation is appropriate before providing the proof in "Appendix A".

Therefore, the following paragraph provides detailed interpretations for the statements in Theorem 1. To avoid an overflow of phrases like "let …be" or "If …is fulfilled, then …," it is assumed that all objects are used as stated in the theorem and that statement 1. is fulfilled.

The first part of statement 2. was added to avoid confusion over the choice of $\{\pi_J\}_{J\subseteq I_2}$. However, as a direct consequence of $\bar{F}$ having a well-defined representation as in Eq. (6), it is mathematically redundant. The function $G_{I_1,I_2}^{\{\pi_J\}_{J\subseteq I_2}}$ in Eq. (8) has the interpretation of

$$G_{I_1,I_2}^{\{\pi_J\}_{J\subseteq I_2}}(s,t) = \mathbb{P}\left(\tau_i \in [t,s)\ \forall i \in I_2 \mid \tau_i > s\ \forall i \in I_1\right).$$

As it is well-known, see e.g. Schweizer and Sklar (1983), a multivariate function $F : \mathbb{R}^d \to [0,1]$ is a distribution function if and only if it fulfills the three conditions of "having" margins, groundedness, and non-negative F-volume for all *d-boxes* $(\boldsymbol{a}, \boldsymbol{b}]$, $\boldsymbol{a} < \boldsymbol{b}$. The last property guarantees, that all (d-dimensional) rectangles have a non-negative probability, which can be represented with F using the principle of inclusion and exclusion. Particularly, the property reads

$$\sum\nolimits_{\boldsymbol{c}\in\times_{i=1}^{d}\{a_i,b_i\}} (-1)^{|a_i=c_i|} F(\boldsymbol{c}) \geq 0.$$

Moreover, using the principle of inclusion and exclusion, it follows that a function $\bar{F}$ is a (multivariate) survival function if the corresponding (hypothetical) distribution function, which is defined by

$$F(\boldsymbol{x}) = 1 + \sum_{\emptyset\neq I\subseteq[d]} (-1)^{|I|}\, \bar{F}\left(\sum_{i\in I} x_i \vec{e}_i\right),$$

is a proper multivariate distribution function. In that spirit, the second part of statement 2. has the interpretation of an "$\bar{F}$-volume"-condition. Due to the specific form of the survival function, however, it suffices that the $\bar{F}$-volumes of some special sets are non-negative. For the exchangeable case, this aspect was further investigated in Schenk (2016), where an alternative proof of "statement 1. $\Leftrightarrow$ statement 2." was shown on the copula-level: Each rectangle with non-increasing lower boundaries admits a partition into so called d-boxes of the form $\times_{i=1}^{m-1}(t_i, s_i] \times (t,s]^{d-m+1}$ such that $t_1 \geq \cdots \geq t_{m-1} \geq t$ and $t_{m-1} \geq s$. The special form of the representation in Eq. (6) allows to expand each $\bar{F}$-volume of a d-box into the product of the $\bar{F}$-volume of $\times_{i=1}^{m-1}(t_i, s_i] \times \mathbb{R}_+^{d-m+1}$ and $G_{I_1,I_2}(s,t)$, where I_1 and I_2 are arbitrary sets with cardinality $m-1$ and $d-m+1$,[4] respectively:

$$\begin{aligned}&\mathbb{P}(\tau_{\pi(i)} \in (t_i, s_i]\ \forall i \in [m-1], \tau_{\pi(i)} \in (t,s]\ \forall i \geq m)\\ &\quad = \mathbb{P}(\tau_{\pi(i)} \in (t_i, s_i]\ \forall i \in [m-1]) \cdot G_{\pi([m-1]),[d]\setminus\pi([m-1])}(s,t).\end{aligned}$$

[4] This reflects the exchangeability of $\bar{F}$, which is assumed here.

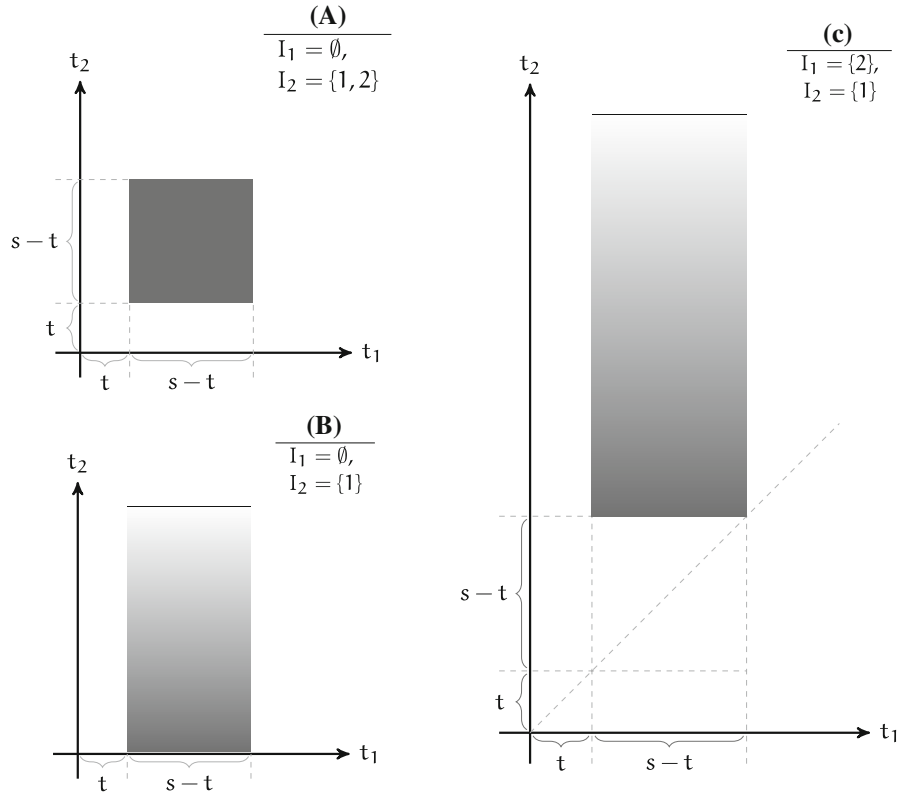


Fig. 1 The reduced set of "test-rectangles" for $d = 2$, which have to be tested for non-negative "2-volume" to verify the validity of a survival function. The three graphs display the three cases, which can be generalized to higher dimensions: **a** squares, which are split in half by the diagonal, **b** infinitely expanding rectangles which touch one axis, and **c** infinitely expanding rectangles which touch the diagonal in one point

Hence, the question of non-negative $\bar{F}$-volume can be reduced inductively to statement 2. For the bivariate case, the remaining sets, which have to be tested for non-negativity, are sketched in Fig. 1. The last part in statement 2. merely reflects the choice of possible shock-distributions, i.e. the class $\bar{\mathcal{G}}$.

Evidently, the statements 3. and 4. are closely linked, as the latter is a special case of the former. The last statement contains the formula, how the survival functions of the original shocks can be retrieved from the multivariate survival function of a GMO distribution. Hence, the implication "statement 4. $\Rightarrow$ statement 1." can be paraphrased as:

If the formula in Eq. (10), *for retrieving the survival functions of the shocks, yields admissible survival functions of class* $\bar{\mathcal{G}}$, *then* $\bar{F}$ *is the survival function of an ESM with shock survival functions* $\bar{S}_I$.

The interpretation of the third statement is a little bit more involved. Given a d-variate model for an ESM and a resulting random vector $\boldsymbol{\tau}$, an important observation, which follows directly from the construction via the min-operator, is that (multivariate)

margins of $\boldsymbol{\tau}$ have a shock model representation, too. Note, that the survival functions of the shocks, corresponding to the marginal model are different, but can be inferred, from those of the full (d-variate) model. To see this, let $\emptyset \neq K \subsetneq [d]$ be a proper subset of $[d]$, preferably with a cardinality bigger than one. Then

$$\tau_i = \min\left\{\min\left\{S_J : J \cap K = I\right\} : i \in I \subseteq K\right\}, \quad i \in K.$$

A calculation, which is very similar to the one used to prove that "statement 4. $\Rightarrow$ statement 1.", yields that

$$\bar{S}_{I_1,I_2}(t) = \prod_{K \subseteq \{1,\ldots,d\} \setminus (I_1 \cup I_2)} \bar{S}_{I_2 \cup K}(t),$$

which is the survival function of $\min\{S_J : J \cap (I_1 \cup I_2) = I_2\}$. Hence, statement 3. requires that statement 4. is fulfilled for every (theoretical) marginal model.

3 Applications and outlook

An *additive subordinator* is a stochastic process $\Lambda = \{\Lambda(t)\}_{t \geq 0}$ on the non-negative real line $[0, \infty]$, which starts at zero, is stochastically continuous, càdlàg, and has independent non-negative increments.[5] Note that this implies that Λ has a.s. non-decreasing path. It can be shown, see Mai et al. (2016), that the distribution of an additive subordinator Λ can uniquely be identified with a family of *Bernstein functions*[6] $\{\psi_t(x)\}_{t \geq 0}$ via $\psi_t(x) = -\log \mathbb{E}[\exp\{-x\Lambda(t)\}]$ and it holds that

(1) $\psi_0(x) = \delta_0(x)$, where δ_0 is the *Dirac-measure* in zero,
(2) $x \mapsto (\psi_s(x) - \psi_t(x))$ is a Bernstein function for all $s > t \geq 0$,
(3) $t \mapsto \psi_t(x)$ is continuous for all $x \geq 0$.

It was shown in Mai et al. (2016) that the random vector $\boldsymbol{\tau}$ belongs to the class of exchangeable generalized Marshall–Olkin distributions which have a stochastic representation as an exchangeable exogenous shock model, where

$$\tau_i := \inf\left\{t > 0 : \Lambda_i(t) > E_i\right\}, \quad i \in [d], \tag{11}$$

$\Lambda_i \equiv \Lambda$ is an additive subordinator, and $\{E_i\}_{i \in [d]}$ are iid unit exponential random variables independent of Λ. Furthmore, if $\psi_t(x) = -\log \mathbb{E}[\exp\{-x\Lambda(t)\}]$, it holds for $\boldsymbol{t} \geq 0$ and $\pi \in \mathcal{S}_d$ with $t_{\pi(1)} \geq \cdots \geq t_{\pi(d)}$ that

$$\mathbb{P}(\boldsymbol{\tau} > \boldsymbol{t}) = \prod_{i=1}^{d} \exp\left\{-\left(\psi_{t_{\pi(i)}}(i) - \psi_{t_{\pi(i)}}(i-1)\right)\right\}. \tag{12}$$

[5] If Λ has also stationary increments, it is called a *Lévy subordinator*.

[6] A *Bernstein function* is a non-negative, infinitely often differentiable function $\psi : [0, \infty) \to [0, \infty)$ with $(-1)^{n+1}\psi^{(n)} \geq 0$. Standard literature, see, e.g., Berg et al. (1984), Schilling et al. (2012), states that the class of Bernstein functions is represented as $\{x \mapsto a1_{(0,\infty)}(x) + bx + \int_{0,\infty}(1 - \exp\{-xs\})\nu(\mathrm{ds}) : a, b \geq 0, \nu$ is a Lévy-measure$\}$.

This model is called *exchangeable additive frailty model (exAFM)* and Theorem 1, or its exchangeable version in Mai et al. (2016), implies that $\boldsymbol{\tau}$ has an alternative representation as an exchangeable exogenous shock model. The exAFM can be generalized to produce non-exchangeable random vectors as the following factor model construction shows: Assume that $\boldsymbol{\tau}$ is defined by Eq. (11), where Λ_i are additive subordinators from the convex cone which is spanned by independent additive subordinators $\Upsilon^{(1)}, \ldots, \Upsilon^{(n)}$ (independent of $E_1, \ldots, E_d$), i.e.

$$\Lambda_i(t) = \boldsymbol{\theta}_i' \boldsymbol{\Upsilon}, \quad i \in [d],$$

for some $n \in \mathbb{N}$ and $\boldsymbol{\theta}_i \in [0, \infty)^n \backslash \{\mathbf{0}\}$, $i \in [d]$. Let $\psi_t^{(k)}(x) = -\log \mathbb{E}[\exp\{-x\Upsilon^{(k)}(t)\}]$, $\boldsymbol{t} \geq 0$. A straightforward calculation, similar to the one in Mai et al. (2016, Prop. 3.1), shows that for $\pi \in \mathcal{S}_d$ with $t_{\pi(1)} \geq \cdots \geq t_{\pi(d)}$

$$\mathbb{P}(\boldsymbol{\tau} > \boldsymbol{t}) = \prod_{i=1}^{d} \prod_{k=1}^{n} \exp\left\{ -\left[\psi_{t_{\pi(i)}}^{(k)} \left(\sum_{j=1}^{i} \Theta_{\pi(i),k} \right) - \psi_{t_{\pi(i)}}^{(k)} \left(\sum_{j=1}^{i-1} \Theta_{\pi(i),k} \right) \right] \right\}, \tag{13}$$

where $\Theta = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n)'$.

This model can be used to define hierarchical models similar to those introduced in Mai (2014). The same model is also described in Sun et al. (2017, Ex. 2). This article also discusses applications to portfolio credit risk. It follows with Theorem 1 that $\boldsymbol{\tau}$ has a generalized Marshall–Olkin distribution, i.e. it has an alternative stochastic representation as an exogenous shock model and the shock distributions can be calculated from the Bernstein functions using the discrete difference operator: Let $s > t \geq 0$ and $\emptyset \neq I \subseteq [d]$ with $I = \{i_1, \ldots, i_{|I|}\}$; then the shock survival function $\bar{S}_I$ fulfills

$$\frac{\bar{S}_I(s)}{\bar{S}_I(t)} = \exp\left\{ (-1)^{|I|} \sum_{k=1}^{n} \Delta_{\Theta_{i_{|I|},k}} \cdots \Delta_{\Theta_{i_1,k}} \left(\psi_s^{(k)} - \psi_t^{(k)} \right) \left(\sum_{j \in [d] \backslash I} \Theta_{j,k} \right) \right\}. \tag{14}$$

This connection between the (hierarchical) additive frailty model and exogenous shock models can be used in multiple ways, e.g., as shown in the following to calculate joint failure probabilities via numerical integration: Let $(t, x) \mapsto \psi_t^{(k)}(x)$ be differentiable w.r.t. t and their partial derivatives w.r.t. t be continuous in x and t. Then

$$\begin{aligned} \mathbb{P}(\tau_1 = \cdots = \tau_d) &= \mathbb{P}\left(Z_{[d]} < \min_{\emptyset \neq I \subsetneq [d]} Z_I \right) \\ &= \mathbb{E}\left[\mathbb{P}\left(Z_{[d]} < \min_{\emptyset \neq I \subsetneq [d]} Z_I \middle| Z_{[d]} \right) \right] = \int_0^\infty \bar{F}(z) \cdot \frac{-\frac{\partial}{\partial z} \bar{S}_{[d]}(z)}{\bar{S}_{[d]}(z)} \, \mathrm{d}\, z \end{aligned}$$

$$= \int_0^\infty \exp\left\{-\sum_{k=1}^n \psi_z^{(k)}\left(\sum_{j=1}^d \Theta_{jk}\right)\right\} \times \left[(-1)^{d+1}\frac{\partial}{\partial z}\sum_{k=1}^n \Delta_{\Theta_{1,k}}\ldots\Delta_{\Theta_{d,k}}\psi_z^{(k)}(0)\right]\mathrm{d}\,z, \tag{15}$$

where $\{Z_I : \emptyset \neq i \subseteq [d]\}$ are independent shocks of a corresponding exogenous shock model and the last step follows with Eqs. (13) and (14). One can also use integration by parts to show that

$$\begin{aligned}
&\mathbb{P}(\tau_1 = \cdots = \tau_d)\\
&= \underbrace{\bar{F}(z)\cdot\left[-\log S_{[d]}(z)\right]\Big|_0^\infty}_{\overset{(\star)}{=}0} + \int_0^\infty \left[\frac{\partial}{\partial z}\bar{F}(z)\right]\cdot\log\bar{S}_{[d]}(z)\,\mathrm{d}\,z\\
&= \int_0^\infty \left[\frac{\partial}{\partial z}\sum_{k=1}^n \psi_z^{(k)}\left(\sum_{j=1}^d \Theta_{jk}\right)\right]\exp\left\{-\sum_{k=1}^n \psi_z^{(k)}\left(\sum_{j=1}^d \Theta_{jk}\right)\right\}\\
&\quad\times\left[(-1)^{d+1}\sum_{k=1}^n \Delta_{\Theta_{1,k}}\ldots\Delta_{\Theta_{d,k}}\psi_z^{(k)}(0)\right]\mathrm{d}\,z,
\end{aligned} \tag{16}$$

where $(\star)$ follows with $\lim_{x\to\infty} x\,\mathrm{e}^{-x} = 0$ and from Eqs. (13) and (14) as well as the Bernstein property of the functions $\psi^{(k)}$, as these imply for $k \in [n]$

$$\begin{aligned}
&\left[(-1)^{d+1}\Delta_{\Theta_{1,k}}\ldots\Delta_{\Theta_{d,k}}\psi_z(0)\right]\\
&= \underbrace{(-1)^{d+1}\Delta_{\Theta_{1,k}}\ldots\Delta_{\Theta_{d-1,k}}\psi_z(\Theta_{d,k})}_{\leq 0} + (-1)^d\Delta_{\Theta_{1,k}}\ldots\Delta_{\Theta_{d-1,k}}\psi_z(0)\\
&\leq (-1)^d\Delta_{\Theta_{1,k}}\ldots\Delta_{\Theta_{d-1,k}}\psi_z(0) \leq \cdots \leq \Delta_{\Theta_{1,k}}\psi_z(0) \leq \psi_z\left(\Theta_{1,k}\right)\\
&\leq \psi_z\left(\sum_{j=1}^d \Theta_{j,k}\right).
\end{aligned}$$

Note that in case the underlying model is exchangeable and the subordinator is a Lévy subordinator with $\psi_t \equiv t\psi_1^{(1)}$ and $\Delta = \Delta_1$, then

$$\begin{aligned}
\mathbb{P}(\tau_1 = \cdots = \tau_d) &\overset{Eq.\,(15)}{=} \int_0^\infty \mathrm{e}^{-z\psi(d)}\cdot\left[(-1)^{d+1}\,\Delta^d\,\psi(0)\right]\mathrm{d}\,z\\
&\overset{Eq.\,(16)}{=} \int_0^\infty \psi(d)\cdot\mathrm{e}^{-z\psi(d)}\cdot\left[(-1)^{d+1}\,z\,\Delta^d\psi(d)\right]\mathrm{d}\,z\\
&= \frac{(-1)^{d+1}\Delta^d\psi(0)}{\psi(d)} = \frac{\sum_{i=0}^d\binom{d}{i}(-1)^{i+1}\psi(i)}{\psi(d)}.
\end{aligned}$$

Equations (13) and (14) have been tested with a simple implementation for the case that $n = 1$, $\Theta = \mathbf{1}$, and $\psi = \psi^{(1)}$ is the Bernstein function of a compound Poisson subordinator with exponentially distributed jumps, i.e. $\psi_t(x) = \mu x t + \beta t \cdot (1 - \eta/(x + \eta))$ for $(\mu, \beta, \eta) \gneq 0$, where exact formulas of the "combined death"-probability are known, see Mai (2010, p. 111 sq.). The three parameter combinations from Mai and Scherer (2017, Fig. 3.6, p. 156 sq.)[7] were used and showed similar results: The exact formula as well as the formula from Eq. (16) perform equally well up to $d \approx 50$ and the formula from Eq. (15) performs well up to $d \approx 25$. The breakdown, which can be detected using the monotonicity properties of the Bernstein function ψ, is due to loss of significant digits in the numerical calculation of the discrete differences. Moreover, for small d the numerical integration formula outperforms a Monte-Carlo estimation of the probabilities w.r.t. error-size as well as runtime.

In case that $n = 1$ and $\Theta = \mathbf{1}$, i.e. if the model is exchangeable, and $\Lambda = \Upsilon^{(1)}$ is a Lévy subordinator, the model can be (uniquely) linked to so called *regenerative composition structures*, see Gnedin and Pitman (2005).[8] In that case, the corresponding shock model is a classical Marshall–Olkin model and the decrement matrix of the corresponding regenerative composition model can be expressed in terms of the exponential rates of the exchangeable MO-distribution $\{\lambda_m^{(n)}, 1 \le m \le n\}$, i.e.

$$q(n : m) = \mathbb{P}\left(\min_{\emptyset \neq I \subseteq [d] : |I| = m} Z_I^{(n)} < \min_{\emptyset \neq I \subseteq [d] : |I| \neq m} Z_I^{(n)} \right) = \frac{\binom{n}{m} \lambda_m^{(n)}}{\sum_{k=1}^{n} \binom{n}{k} \lambda_k^{(n)}},$$

where $\{Z_I^{(n)}\}_{\emptyset \neq I \subseteq [d]}$ are independend exponential random variables with rates $\lambda_I^{(n)} \equiv \lambda_{|I|}^{(n)}$ and

$$\lambda_m^{(n)} = \sum_{j=0}^{m} (-1)^{j+1} \binom{m}{j} \psi(n - m + j), \quad 1 \le m \le n.$$

Theorem 1 can subsequently be used to extend some results from Gnedin and Pitman (2005) for composition structures which fulfill a suitably relaxed notion of regenerativity such that the stochastic process representation uses an additive subordinator instead of a Lévy subordinator.

4 Conclusion

The survival functions of ESM distributions are the product of their ordered and individually transformed arguments. The transformations g_i^{π} are order-dependent if the ESM distribution is not exchangeable. Conversely, if a function of that form is a continuous multivariate survival function, the distribution has a stochastic representation

[7] These are (0.2995, 1.401, 1), (0.2, 2.4, 2), and (0.0151, 0.994749, 0.01).

[8] For a definition of (regenerative) composition structures and an introduction of the notation which is used hereinafter, the interested reader is referred to Gnedin and Pitman (2005).

as an exogenous shock model. Formulas for retrieving the shock survival functions from the transformations g_i^π are given explicitly. Furthermore, the special form of $\bar{F}(\boldsymbol{t}) = \prod_{i=1}^d g_i^\pi(t_{\pi(i)})$ implies a simplified d-volume condition. The attained results generalize the findings from Mai et al. (2016) for the exchangeable subclass.

Acknowledgements We thank the two anonymous reviewers as well as the editor and associate editor who provided helpful comments.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

A Proof of Theorem 1

The theorem will be proven in four steps. Particularly, it is shown that 3. $\Rightarrow$ 4. $\Rightarrow$ 1. $\Rightarrow$ 2. $\Rightarrow$ 3. Proofs of used auxiliary results are deferred to "Appendix B".

Remark 1 Under the assumptions of Theorem 1, particularly the representation of $\bar{F}$ in Eq. (6), the expression

$$g_i^\pi(t) = \frac{\prod_{j=1}^i g_j^\pi(t)}{\prod_{j=1}^{i-1} g_j^\pi(t)}$$

is invariant for different permutations with coinciding images of $[i-1]$ and i. If the first statement of the theorem is fulfilled, then g_i^π has the interpretation of a conditional probability, i.e.

$$g_i^\pi(t) = \mathbb{P}\left(\tau_{\pi(i)} > t \mid \boldsymbol{\tau}_{\pi([i-1])} > t\right).$$

Hence, the function g_i^π only depends on $\pi([i-1])$ and $\pi(i)$ and it is justified to work with $\tilde{g}^{\pi([i]),\pi(i)}$.

Proof of 3. $\Rightarrow$ 4. First observe that 4. is a special case of 3., hence 3. $\Rightarrow$ 4. follows directly. □

Proof of 4. $\Rightarrow$ 1. Let 4. from Theorem 1 be fulfilled and define for independent random variables $Z_I \sim \bar{S}_I$, $\emptyset \neq I \subseteq [d]$, the random vector $\boldsymbol{\tau}$ by

$$\tau_i := \min\{Z_I : i \in I\}, \quad i \in [d].$$

For $\boldsymbol{t} \geq 0$ and $\pi \in \mathcal{S}_d$ with $t_{\pi(1)} \geq \cdots \geq t_{\pi(d)}$, using the independence of the shock variables and reordering the factors, it holds that

$$\mathbb{P}(\boldsymbol{\tau} > \boldsymbol{t}) = \prod_{\emptyset \neq I \subseteq [d]} \mathbb{P}\left(Z_I > \max_{i \in I} t_i\right) = \prod_{i=1}^d \left(\prod_{\substack{I \subseteq \pi(\{i, i+1, \ldots, d\}) \\ \pi(i) \in I}} \mathbb{P}\left(Z_I > t_{\pi(i)}\right) \right).$$

For $i \in [d]$ and $\pi(i) \in I \subseteq \pi(\{i, \ldots, d\})$, by assumption, the survival function $\bar{S}_I \equiv \bar{S}_I^{\pi(i)}$ has a representation as in Eq. (10) with $m = \pi(i)$ and

$$\prod_{\substack{I \subseteq \pi(\{i,i+1,\ldots,d\}) \\ \pi(i) \in I}} \mathbb{P}\left(Z_I > t_{\pi(i)}\right) = \prod_{\substack{I \subseteq \pi(\{i,i+1,\ldots,d\}) \\ \pi(i) \in I}} \left(\prod_{\substack{J \subseteq I \\ \pi(i) \in J}} \left(\tilde{g}^{J \cup ([d] \setminus I), \pi(i)} \left(t_{\pi(i)} \right) \right)^{(-1)^{|J|-1}} \right).$$

Fix $K \subseteq [d]$ with $\pi([i]) \subseteq K$; then $i \leq |K| = k \leq d$ and $1 \leq j \leq k - i + 1$. The expression $\tilde{g}^{K, \pi(i)}(t_{\pi(i)})$ with an exponent of $(-1)^{j-1}$ appears $\binom{k-i}{j-1}$ times, as there are exactly $\binom{k-i}{j-1}$ possible choices for J with $\pi(i) \in J \subseteq K \setminus \pi([i-1])$. Hence, the overall exponent of the expression $\tilde{g}^{K, \pi(i)}(t_{\pi(i)})$ is

$$\sum_{j=1}^{k-i+1} (-1)^{j-1} \binom{k-i}{j-1} = \sum_{j=0}^{k-i} (-1)^j \binom{k-i}{j} = (1-1)^{k-i} = \begin{cases} 1, & k = i \\ 0, & k > i, \end{cases}$$

where the latter expression follows with the binomial formula. Finally, it follows that

$$\mathbb{P}(\boldsymbol{\tau} > \boldsymbol{t}) = \prod_{i=1}^{d} \tilde{g}^{\pi([i]), \pi(i)} \left(t_{\pi(i)} \right) = \prod_{i=1}^{d} g_i^{\pi} \left(t_{\pi(i)} \right).$$

□

In the following, I_1, I_2, $\{\pi_J\}_{J \subseteq I_2}$, s and t (or a subset of these elements) fulfill the *usual conditions* if

1. $s > t \geq 0$,
2. $I_1, I_2 \subseteq [d]$ with $I_1 \cap I_2 = \emptyset$ and $I_2 \neq \emptyset$,
3. for $J \subseteq I_2$ one has
   (a) $\pi_J(\{1, \ldots, |I_1|\}) = I_1$ (if $I_1 \neq \emptyset$),
   (b) $\pi_J(\{|I_1| + 1, \ldots, |I_1 \cup J|\}) = J$,
   (c) $\pi_J(\{|I_1 \cup J| + 1, \ldots, |I_2|\}) = I_2 \setminus J$.

If only a specific permutation π is used, it is assumed that it fulfills this property for $J = I_2$.

Proof of 1. $\Rightarrow$ 2. Let 1. in Theorem 1 be fulfilled and let I_1, I_2, $\{\pi_J\}_{J \subseteq I_2}$, s and t fulfill the usual conditions. First assume that for arbitrary $\pi \in \mathcal{S}_d$ and $i \in [d]$ the functions g_i^{π} are strictly positive on $\mathbb{R}_+$. Then

$$G_{I_1, I_2}^{\{\pi_J\}_{J \subseteq I_2}}(s, t) = \frac{\sum_{J \subseteq I_2} (-1)^{|J|} \prod_{j=1}^{|I_1 \cup J|} g_j^{\pi_J}(s) \prod_{j=1}^{|I_2 \setminus J|} g_{|I_1 \cup J|+j}^{\pi_J}(t)}{\prod_{j=1}^{|I_1|} g_j^{\pi_\emptyset}(s)}, \tag{17}$$

where it is used that by 1. the diagonal of marginal survival functions of $\boldsymbol{\tau}_{I_1}$ can be represented with every π fulfilling $\pi(\{1,\ldots,|I_1|\}) = I_1$. Particularly, it holds that

$$\mathbb{P}(\tau_i > s, i \in I_1) = \prod_{j=1}^{|I_1|} g_j^{\pi_{J_1}}(s) = \prod_{j=1}^{|I_1|} g_j^{\pi_{J_2}}(s),\ J_1, J_2 \subseteq I_2, \quad s \geq 0.$$

Subsequently, the numerator of Eq. (17) can be rewritten using the principle of inclusion and exclusion as

$$\begin{aligned}
&\sum_{i=0}^{|I_2|} (-1)^i \sum_{J \subseteq I_2 : |J| = i} \prod_{j=1}^{|J \cup I_1|} g_j^{\pi_J}(s) \prod_{j=1}^{|I_2 \setminus J|} g_{|I_1 \cup J|+j}^{\pi_J}(t) \\
&\quad = \mathbb{P}\left(A_{\emptyset}^{I_1,I_2}\right) - \sum_{i=1}^{|I_2|} (-1)^{i+1} \sum_{J \subseteq I_2 : |J| = i} \mathbb{P}\left(\bigcap_{j \in J} A_j^{I_1,I_2}\right) \\
&\quad = \mathbb{P}\left(A_{\emptyset}^{I_1,I_2}\right) - \mathbb{P}\left(\bigcup_{i \in I_2} A_i^{I_1,I_2}\right) = \mathbb{P}\left(A^{I_1,I_2}\right),
\end{aligned}$$

where

$$\begin{aligned}
A^{I_1,I_2} &:= \{\tau_i > s\ \forall i \in I_1, \tau_i \in (t,s]\ \forall i \in I_2\}, \\
A_{\emptyset}^{I_1,I_2} &:= \{\tau_i > s\ \forall i \in I_1, \tau_i > t\ \forall i \in I_2\}, \quad \text{and} \\
A_i^{I_1,I_2} &:= \left(\bigcap_{j \in I_1 \cup \{i\}} \{\tau_j > s\}\right) \cap \left(\bigcap_{j \in I_2 \setminus \{i\}} \{\tau_j > t\}\right), \quad i \in I_2.
\end{aligned}$$

It follows that

$$G_{I_1,I_2}^{\{\pi_J\}_{J \subseteq I_2}}(s,t) = \mathbb{P}(\tau_i \in (t,s]\ \forall i \in I_2 \mid \tau_i > s\ \forall i \in I_1)$$

and subsequently that $G_{I_1,I_2}^{\{\pi_J\}_{J \subseteq I_2}}(s,t)$ is non-negative and does not depend on the specific choice of $\{\pi_J\}_{J \subseteq I_2}$.[9]

Now, by induction over i, the strict positivity, continuity, and non-increasingness of g_i^{π} is proven for all $\pi \in \mathcal{S}_d$. This implies that $G_{I_1,I_2}(s,t)$ is continuous in s and t. For $i = 1$ and $\pi \in \mathcal{S}_d$, the assumptions of Theorem 1 imply that g_1^{π} is strictly positive, continuous, and non-increasing. Let the claim be fulfilled for $j < i$, i.e. g_j^{π} is strictly positive, continuous, and non-increasing for $j \leq i-1$ and $\pi \in \mathcal{S}_d$.

[9] The independence of the specific choice of $\{\pi_J\}_{J \subseteq I_2}$ can also be derived without resorting to the probabilistic interpretation by using the assumption that $\bar{F}$ has a well-defined representation as in Eq. (6).

Right-continuity and left-limits It is well known, see, e.g., Schweizer and Sklar (1983, Chp. 6), that copulae are Lipschitz-continuous with constant one. Hence, by exploiting the copula/survival function decomposition, it holds that

$$\left|\bar{F}(s_1,\ldots,s_d)-\bar{F}(t_1,\ldots,t_d)\right| \le \sum_{i=1}^{d}\left|\bar{F}_i(s_i)-\bar{F}_i(t_i)\right|, \quad \forall \boldsymbol{t},\boldsymbol{s}\ge 0$$

and right-continuity as well as left-limits of $\bar{F}$ are inherited from the margins. For $\pi \in \mathcal{S}_d$ the survival function $t \mapsto \mathbb{P}\left(\min_{j\le i}\tau_{\pi(j)} > t\right)$ is right-continuous with left-limits and with

$$g_i^{\pi}(t) = \frac{\prod_{j=1}^{i} g_j^{\pi}(t)}{\prod_{j=1}^{i-1} g_j^{\pi}(t)} = \frac{\mathbb{P}\left(\min_{j\le i}\tau_{\pi(j)} > t\right)}{\prod_{j=1}^{i-1} g_j^{\pi}(t)},$$

right-continuity with left-limits for g_i^{π} follows with the induction hypothesis.

Non-increasingness For $\pi \in \mathcal{S}_d$ and $s \ge t \ge 0$ define the vector $\boldsymbol{u}(s,t)$ by

$$u_{\pi(j)}(s,t) \coloneqq \begin{cases} s, & \forall j < i, \\ t, & j = i, \\ 0, & \forall j > i. \end{cases}$$

Then, by monotonicity of the measure $\mathbb{P}$, one has

$$\begin{aligned} \mathbb{P}(\boldsymbol{\tau} > \boldsymbol{u}(s,s)) &\le \mathbb{P}(\boldsymbol{\tau} > \boldsymbol{u}(s,t)) \\ &\Leftrightarrow g_i^{\pi}(s)\prod_{j=1}^{i-1} g_j^{\pi}(s) \le g_i^{\pi}(t)\prod_{j=1}^{i-1} g_j^{\pi}(s) \\ &\Leftrightarrow g_i^{\pi}(s) \le g_i^{\pi}(t), \end{aligned}$$

where the induction hypothesis, i.e. g_j^{π} is strictly positive for all $j < i$, is used.

Strict positivity Assume for $\pi \in \mathcal{S}_d$ that there exists a finite upper bound $s^{\star}$ for strict positivity of g_i^{π}, i.e. $s^{\star} \coloneqq \inf\{u > 0 : g_i^{\pi}(u) = 0\} < \infty$, and as g_i^{π} is right-continuous and non-increasing we have that $g_i^{\pi}(s^{\star}) = 0$. For $t < s^{\star}$, let $I_1 = \pi(\{1,\ldots,i-2\})$ and $I_2 = \pi(\{i-1,i\})$. Furthermore, let $\tilde{\pi}$ be the permutation which switches the positions of $i-1$ and i in π, i.e. $\tilde{\pi} = \pi(i-1,i)$. Assume w.l.o.g. that $s^{\star} \le u^{\star}$ for $u^{\star} \coloneqq \inf\{u > 0 : g_i^{\tilde{\pi}}(u) = 0\} \in \bar{\mathbb{R}}_+$ (else switch the roles of π and $\tilde{\pi}$ and prove the contradiction for $\tilde{\pi}$ first). Then, with the induction hypothesis it holds that $g_j^{\pi}, g_j^{\tilde{\pi}} > 0\ \forall j < i$ and, for $\pi_{\emptyset} \in \{\pi, \tilde{\pi}\}$, that

$$
\begin{aligned}
0 \overset{\text{IH}}{\leq} G_{I_1,I_2}(s^\star, t) &= \prod_{j=i-1}^{i} g_j^{\pi_\emptyset}(t) - g_{i-1}^{\pi}(s^\star) g_i^{\pi}(t) - g_{i-1}^{\tilde{\pi}}(s^\star) g_i^{\tilde{\pi}}(t) + \prod_{j=i-1}^{i} g_j^{\pi}(s^\star) \\
&= g_{i-1}^{\pi_\emptyset}(t) g_i^{\pi_\emptyset}(t) - g_{i-1}^{\pi}(s^\star) g_i^{\pi}(t) - g_{i-1}^{\tilde{\pi}}(s^\star) g_i^{\tilde{\pi}}(t) \\
&= \begin{cases} \left(g_{i-1}^{\pi}(t) - g_{i-1}^{\pi}(s^\star)\right) g_i^{\pi}(t) - g_{i-1}^{\tilde{\pi}}(s^\star) g_i^{\tilde{\pi}}(t), & \pi_\emptyset = \pi \\ \left(g_{i-1}^{\tilde{\pi}}(t) - g_{i-1}^{\tilde{\pi}}(s^\star)\right) g_i^{\tilde{\pi}}(t) - g_{i-1}^{\pi}(s^\star) g_i^{\pi}(t), & \pi_\emptyset = \tilde{\pi}. \end{cases}
\end{aligned} \tag{18}
$$

The last expression in Eq. (18) becomes negative if t is sufficiently close to $s^\star$:

1. If $u^\star > s^\star$, choose $\pi_\emptyset = \pi$. Then for $t \nearrow s^\star$ Eq. (18) approaches $-g_{i-1}^{\tilde{\pi}}(s^\star) g_i^{\tilde{\pi}}(s^\star-)$.
   As $g_{i-1}^{\tilde{\pi}}(s^\star) > 0$ by the induction hypothesis and $g_i^{\tilde{\pi}}(t) > 0 \, \forall t < u^\star$ with $s^\star < u^\star$ by the assumption made above it holds that

$$0 \leq -g_{i-1}^{\tilde{\pi}}(s^\star) g_i^{\tilde{\pi}}(s^\star-) < 0.$$

2. If $s^\star = u^\star$ and $g_i^{\pi_\emptyset}(s^\star-) > g_i^{\pi_\emptyset}(s^\star) = 0$ for at least one $\pi_\emptyset \in \{\pi, \tilde{\pi}\}$, then for $t \nearrow s^\star$ Eq. (18) approaches $-g_{i-1}^{\pi_\emptyset}(s^\star) g_i^{\pi_\emptyset}(s^\star-)$.
   As $g_{i-1}^{\pi_\emptyset}(s^\star) > 0$ by the induction hypothesis and $g_i^{\pi_\emptyset}(s^\star-) > 0$ by the assumption made above it holds that

$$0 \leq -g_{i-1}^{\pi_\emptyset}(s^\star) g_i^{\pi_\emptyset}(s^\star-) < 0.$$

3. Otherwise, as $g_j^{\pi_\emptyset}$ for $j \in \{i-1, i\}$ have left-limits by the induction hypothesis, for every sequence $t_k \nearrow s^\star$ with $t_k \neq s^\star$, non-negative sequences $\{a_{j,k}^{\pi_\emptyset}\}_{k \in \mathbb{N}}$ with $a_{j,k}^{\pi_\emptyset}(s^\star - t_k) \to 0$ for $k \to \infty$ can be found s.t.

$$g_j^{\pi_\emptyset}(t_k) = g_j^{\pi_\emptyset}(s^\star-) + a_{j,k}^{\pi_\emptyset}(s^\star - t_k), \quad j \in \{i-1, i\}, \quad k \in \mathbb{N}.$$

   By the assumption on $s^\star$, it holds that $a_{i,k}^{\pi_\emptyset} > 0$ for all $k \in \mathbb{N}$ and $\pi_\emptyset \in \{\pi, \tilde{\pi}\}$. If $s^\star = u^\star$ and $g_i^{\pi_\emptyset}(s^\star-) = g_i^{\pi_\emptyset}(s^\star) = 0$ for all $\pi_\emptyset \in \{\pi, \tilde{\pi}\}$, it follows from Eq. (18) and (left-)continuity of $g_{i-1}^{\pi_\emptyset}$ that

$$0 \leq \begin{cases} a_{i-1,k}^{\pi} a_{i,k}^{\pi}(s^\star - t_k)^2 - g_{i-1}^{\tilde{\pi}}(s^\star) a_{i,k}^{\tilde{\pi}}(s^\star - t_k), & \pi_\emptyset = \pi \\ a_{i-1,k}^{\tilde{\pi}} a_{i,k}^{\tilde{\pi}}(s^\star - t_k)^2 - g_{i-1}^{\pi}(s^\star) a_{i,k}^{\pi}(s^\star - t_k), & \pi_\emptyset = \tilde{\pi} \end{cases}$$

   or equivalently

$$0 \leq \begin{cases} a_{i-1,k}^{\pi}(s^\star - t_k) \frac{a_{i,k}^{\pi}}{a_{i,k}^{\tilde{\pi}}} - g_{i-1}^{\tilde{\pi}}(s^\star), & \pi_\emptyset = \pi \\ a_{i-1,k}^{\tilde{\pi}}(s^\star - t_k) \frac{a_{i,k}^{\tilde{\pi}}}{a_{i,k}^{\pi}} - g_{i-1}^{\pi}(s^\star), & \pi_\emptyset = \tilde{\pi}. \end{cases}$$

Now choose k sufficiently large and $\pi_\emptyset$ s.t. the fraction appearing in the upper equation is smaller or equal to 1, then

$$\begin{aligned} 0 \leq & \begin{cases} a^{\pi}_{i-1,k}(s^\star - t_k) - g^{\tilde{\pi}}_{i-1}(s^\star), & a^{\pi}_{i,k} \leq a^{\tilde{\pi}}_{i,k} \\ a^{\tilde{\pi}}_{i-1,k}(s^\star - t_k) - g^{\pi}_{i-1}(s^\star), & a^{\pi}_{i,k} > a^{\tilde{\pi}}_{i,k} \end{cases} \\ & < 0, \end{aligned}$$

where it is used that the respective first summand converges for $k \to \infty$ to 0 and the last summand is negative. Hence, a contradiction is found for each case and therefore $g^{\pi}_i(t) > 0 \; \forall t \in \mathbb{R}_+$.

Left-continuity Let I_1 and I_2 as well as π, $\tilde{\pi}$, and $\pi_\emptyset$ be as above. Then, for all $s > t \geq 0$ the function

$$\mathbb{P}(\tau_i \in (t, s], i \in I_2 \mid \tau_i > s, i \in I_1) = G_{I_1,I_2}(s, t)$$

has left-limits in t. Assume that there exists $s^\dagger \in \mathbb{R}_+^\times$ with $g^{\pi}_i(s^\dagger -) > g^{\pi}_i(s^\dagger)$, then

$$\begin{aligned} 0 &\overset{\text{IH}}{\leq} \lim_{t \nearrow s^\dagger} G_{I_1,I_2}(s^\dagger, t) \\ &= \lim_{t \nearrow s^\dagger} \left(\prod_{j=i-1}^{i} g^{\pi_\emptyset}_j(t) - g^{\pi}_{i-1}(s^\dagger) g^{\pi}_i(t) - g^{\tilde{\pi}}_{i-1}(s^\dagger) g^{\tilde{\pi}}_i(t) + \prod_{j=i-1}^{i} g^{\pi}_j(s^\dagger) \right) \\ &\overset{\pi_\emptyset = \tilde{\pi}, (\star)}{=} \left(g^{\pi}_i(s^\dagger) - g^{\pi}_i(s^\dagger -) \right) g^{\pi}_{i-1}(s^\dagger) < 0, \end{aligned}$$

where it is used in $(\star)$, that the first and third summand cancel out, when using that $g^{\tilde{\pi}}_{i-1}$ is continuous under the induction hypothesis. This is a contradiction—hence g^{π}_i is left-continuous. □

Remark 2 The induction in the second part of the proof can be performed on the basis of statement 2. (instead of 1.) from Theorem 1 if the parts on *right-continuity with left-limits* and *non-increasingness* are replaced by the following lemma (as they rely on the survival function assumption of 1.). In particular, 2. implies $g^{\pi}_i \in \bar{\mathcal{G}}$ for all $i \in [d], \pi \in \mathcal{S}_d$.

Lemma 1 *Let 2. from Theorem 1 be fulfilled and g^{π}_j be right-continuous with left-limits, non-increasing, and strictly positive for all $j \leq i-1$ and $\pi \in \mathcal{S}_d$. Then g^{π}_i is right-continuous with left-limits and non-increasing for all $\pi \in \mathcal{S}_d$.*

Lemma 2 *Assume that statement 2. of Theorem 1 is fulfilled and let I_1 and I_2 fulfill the usual conditions. Then for each $m \in I_2$, $\bar{S}^m_{I_1,I_2}$ is an $\mathbb{R}_+$-valued, positive, and continuous function on $\mathbb{R}_+$. Furthermore, $\bar{S}^m_{I_1,I_2}$ does not depend on $m \in I_2$, i.e.*

$$\bar{S}^{m_1}_{I_1,I_2}(t) = \bar{S}^{m_2}_{I_1,I_2}(t), \quad \forall t \geq 0, m_1, m_2 \in I_2. \tag{19}$$

Lemma 3 *Let I_1 and I_2 fulfill the usual conditions and assume that $\bar{S}^{m_1}_{I_1\cup I_2\setminus J,J} = \bar{S}^{m_2}_{I_1\cup(I_2\setminus J),J} \in \bar{\mathcal{G}}$ for all $\emptyset \neq J \subseteq I_2$ and $m_1, m_2 \in J$. Then for $s > t \geq 0$*

$$G_{I_1,I_2}(s,t) = \mathbb{P}\left(\check{\tau}_i \in (t,s]\ \forall i \in I_2\right),$$

where

$$\check{\tau}_i \coloneqq \min\left\{\check{Z}_J : i \in J \subseteq I_2\right\}, \quad i \in [d]$$

with independent random shocks $\check{Z}_J \sim \bar{S}_{I_1\cup I_2\setminus J,J}$ for $\emptyset \neq J \subseteq I_2$.

The essence of the previous Lemma is the following: Let I_1 and I_2 fulfill the usual conditions, $Z_I \sim \bar{S}_I \in \bar{\mathcal{G}}$, $\emptyset \neq I \subseteq [d]$, $\boldsymbol{\tau}$ be defined as in Eq. (1), and $\check{\boldsymbol{\tau}} \in \mathbb{R}_+^{|I_2|}$ be defined by

$$\check{\tau}_i \coloneqq \min\left\{\min\left\{Z_J : J \cap (I_1 \cup I_2) = I\right\} : i \in I \subseteq I_2\right\}.$$

Then

$$\mathbb{P}\left(\tau_i \in (t,s]\ \forall i \in I_2 \mid \tau_i > s\ \forall i \in I_1\right) = \mathbb{P}\left(\check{\tau}_i \in (t,s]\ \forall i \in I_2\right), \quad \forall s > t \geq 0.$$

Lemma 4 *Let I_1 and I_2 fulfill the usual conditions. Then, for a specific family of permutations, $\{\pi_J\}_{J\subseteq I_2}$, the function $G^{\{\pi_J\}_{J\subseteq I_2}}_{I_1,I_2}$ depends on $g^{\pi_J}_i$, $|I_1|+1 \leq i \leq |I_1\cup I_2|$, $J \subseteq I_2$. Therefore, write*

$$G^{\{\pi_J\}_{J\subseteq I_2}}_{I_1,I_2} \equiv G^{\left\{g^{\pi_J}_{|I_1|+1}, g^{\pi_J}_{|I_1|+2}, \ldots, g^{\pi_J}_{|I_1\cup I_2|}\right\}_{J\subseteq I_2}}_{I_1,I_2}.$$

Assume that $g^{\pi_J}_i$, $|I_1|+1 \leq i \leq |I_1\cup I_2|$, $J \subseteq I_2$ are positive. Then it holds for all $s \geq t \geq 0$ that

$$\begin{aligned} G^{\left\{g^{\pi_J}_{|I_1|+1}, \ldots, g^{\pi_J}_{|I_1\cup I_2|}\right\}_{J\subseteq I_2}}_{I_1,I_2}(s,t) &= \hat{g}^{\pi_\emptyset}_{|I_1|+1}(t) \cdot g^{\pi_\emptyset}_{|I_1|+2}(t) \cdot \cdots \cdot g^{\pi_\emptyset}_{|I_1\cup I_2|}(t) \\ &\quad \times \left(\frac{g^{\pi_\emptyset}_{|I_1|+1}(t)}{\hat{g}^{\pi_\emptyset}_{|I_1|+1}(t)} - \frac{g^{\pi_\emptyset}_{|I_1|+1}(s)}{\hat{g}^{\pi_\emptyset}_{|I_1|+1}(s)}\right) + \frac{g^{\pi_\emptyset}_{|I_1|+1}(s)}{\hat{g}^{\pi_\emptyset}_{|I_1|+1}(s)} \\ &\quad \times G^{\left\{\hat{g}^{\pi_J}_{|I_1|+1}, g^{\pi_J}_{|I_1|+2}, \ldots, g^{\pi_J}_{|I_1\cup I_2|}\right\}_{J\subseteq I_2}}_{I_1,I_2}(s,t) \end{aligned} \tag{20}$$

for an arbitrary function $\hat{g}^{\pi_\emptyset}_{|I_1|+1}$ which is positive on $\mathbb{R}_+$, where

$$\hat{g}^{\pi_J}_{|I_1|+1}(s) \coloneqq \frac{g^{\pi_J}_{|I_1|+1}(s)}{g^{\pi_\emptyset}_{|I_1|+1}(s)} \hat{g}^{\pi_\emptyset}_{|I_1|+1}(s), \quad J \subseteq I_2, s \geq 0,$$

which are by definition positive functions on $\mathbb{R}_+$.

Lemma 5 *For $k \in \mathbb{N}_0$, $j \geq 2$, let the functions $\bar{F}_{1,k}, \ldots, \bar{F}_{j,k} : [0,\infty) \to (0,1]$ as well as $\bar{F}_{1,k+1}, \ldots, \bar{F}_{j-1,k+1} : [0,\infty) \to (0,1]$ be non-increasing with $\bar{F}_{l,k} = \frac{\bar{F}_{l-1,k}}{\bar{F}_{l-1,k+1}}$ for $l \in \{2,\ldots,j\}$. Then it holds that for $s \geq t \geq 0$*

$$0 \leq \bar{F}_{j,k}(t) - \bar{F}_{j,k}(s) \leq \left(\prod_{l=1}^{j-1} \frac{1}{\bar{F}_{l,k+1}(s)} \right) \cdot \left(\bar{F}_{1,k}(t) - \bar{F}_{1,k}(s) \right).$$

Proof of 2. $\Rightarrow$ 3. Let statement 2. in Theorem 1 be fulfilled, then due to Remark 2, Lemmas 1 and 2:

- For $i = 1, \ldots, d$ and $\pi \in \mathcal{S}_d$, it holds that $g_i^\pi \in \bar{\mathcal{G}}$.
- For I_1 and I_2 fulfilling the usual conditions and $m \in I_2$, the function $\bar{S}^m_{I_1,I_2}$ is well-defined as well as positive and continuous. Moreover, it does not depend on the specific $m \in I_2$ chosen, hence write $\bar{S}_{I_1,I_2}$.

It is left to prove that $\bar{S}_{I_1,I_2}$ is non-increasing for all I_1, I_2 fulfilling the usual conditions.

The claim is proven by induction over $|I_2|$. For $I_2 = \{m\}$, let I_1 and I_2 fulfill the usual conditions, then $\bar{S}_{I_1,I_2} = \tilde{g}^{I_1 \cup I_2, m} \in \bar{\mathcal{G}}$. Now let $p > 1$ and assume that for all I_1 and I_2 fulfilling the usual conditions with $|I_2| < p$ it holds that $\bar{S}_{I_1,I_2} \in \bar{\mathcal{G}}$. Let I_1, I_2, $\{\pi_J\}_{J \subseteq I_2}$, s, and t fulfill the usual conditions and $|I_2| = p$ and define the function $\hat{g}^{\pi_\emptyset}_{|I_1|+1} := g^{\pi_\emptyset}_{|I_1|+1} / \bar{S}_{I_1,I_2}$, which is continuous and positive. With Lemma 4 it follows that

$$\begin{aligned} 0 &\leq G^{\left\{g^{\pi_J}_{|I_1|+1}, g^{\pi_J}_{|I_1|+2}, \ldots, g^{\pi_J}_{|I_1 \cup I_2|}\right\}_{J \in I_2}}_{I_1,I_2}(s,t) \\ &= \hat{g}^{\pi_\emptyset}_{|I_1|+1}(t) g^{\pi_\emptyset}_{|I_1|+2}(t) \cdot \cdots \cdot g^{\pi_\emptyset}_{|I_1 \cup I_2|}(t) \cdot \left(\bar{S}_{I_1,I_2}(t) - \bar{S}_{I_1,I_2}(s) \right) + \bar{S}_{I_1,I_2}(s) \\ &\quad \times G^{\left\{\hat{g}^{\pi_J}_{|I_1|+1}, g^{\pi_J}_{|I_1|+2}, \ldots, g^{\pi_J}_{|I_1 \cup I_2|}\right\}_{J \in I_2}}_{I_1,I_2}(s,t), \end{aligned} \tag{21}$$

where $\hat{g}^{\pi_J}_{|I_1|+1} := g^{\pi_J}_{|I_1|+1} / \bar{S}_{I_1,I_2}$ for $J \subseteq I_2$.

In light of Lemma 3, it makes sense to derive an exogenous shock model from

$$\left\{ \hat{g}^{\pi_J}_{|I_1|+1}, g^{\pi_J}_{|I_1|+2}, \ldots, g^{\pi_J}_{|I_1 \cup I_2|} \right\}_{J \in I_2}.$$

Hence one has to check, that for $\emptyset \neq J \subseteq I_2$ if $\bar{\hat{S}}_{I_1 \cup I_2 \setminus J, J} \in \bar{\mathcal{G}}$. Note that

$$\bar{\hat{S}}_{I_1 \cup I_2 \setminus J, J} = \begin{cases} \bar{S}_{I_1 \cup I_2 \setminus J, J}, & \emptyset \neq J \subsetneq I_2 \\ 1, & J = I_2. \end{cases}$$

As $\bar{S}_{I_1 \cup I_2 \setminus J, J} \in \bar{\mathcal{G}}$ by the induction step for $\emptyset \neq J \subsetneq I_2$ and $\bar{\hat{S}}_{I_1,I_2} \equiv 1 \in \bar{\mathcal{G}}$, Lemma 3 can be used. Write for $s > t \geq 0$

$$G^{\left\{\hat{g}^{\pi_J}_{|I_1|+1}, g^{\pi_J}_{|I_1|+2}, \ldots, g^{\pi_J}_{|I_1 \cup I_2|}\right\}_{J \in I_2}}_{I_1,I_2}(s,t) = \mathbb{P}\left(\hat{\tau}_i \in (t,s] \; \forall i \in I_2 \right),$$

where

$$\hat{\tau}_i \coloneqq \min\left\{\hat{Z}_I : i \in I \subseteq I_2\right\}, \quad i \in I_2$$

with independent $\hat{Z}_I \sim \hat{S}_{I_1 \cup I_2 \setminus I, I}$ for $\emptyset \neq I \subseteq I_2$. Let $s > t \geq 0$ and define

$$\hat{A}^{I_1, I_2} \coloneqq \left\{\hat{\tau}_i \in (t, s]\ \forall i \in I_2\right\}.$$

Since $\hat{Z}_{I_2} = \infty$, there are at least two different sets $\emptyset \neq I, J \subsetneq I_2$ for which the respective shocks $\hat{Z}_I$, $\hat{Z}_J$ are minimal for one of their components. Moreover, this implies

$$\hat{A}^{I_1, I_2} \subseteq \bigcup_{\emptyset \neq I, J \subsetneq I_2 : I \neq J} \left\{t < \hat{Z}_I, \hat{Z}_J \leq s\right\}.$$

From the sub-additivity of the probability measure $\mathbb{P}$, it follows that

$$\begin{aligned}
\mathbb{P}(\hat{A}^{I_1, I_2}) &= G_{I_1, I_2}^{\left\{\hat{g}^{\pi_J}_{|I_1|+1}, g^{\pi_J}_{|I_1|+2}, \ldots, g^{\pi_J}_{|I_1 \cup I_2|}\right\}_{J \in I_2}}(s, t) \\
&\leq \sum_{\substack{\emptyset \neq I, J \subsetneq I_2 \\ I \neq J}} \mathbb{P}\left(t < \hat{Z}_I, \hat{Z}_J \leq s\right) \\
&\leq \binom{2^{|I_2|} - 2}{2} \max_{\emptyset \neq I \subsetneq I_2} \left(\bar{S}_{I_1 \cup I_2 \setminus I, I}(t) - \bar{S}_{I_1 \cup I_2 \setminus I, I}(s)\right)^2,
\end{aligned}$$

where we used that for $\emptyset \neq I \subsetneq I_2$

$$\mathbb{P}\left(t < \hat{Z}_I \leq s\right) = \bar{S}_{I_1 \cup I_2 \setminus I, I}(t) - \bar{S}_{I_1 \cup I_2 \setminus I, I}(s).$$

Note that for $\emptyset \neq J \subseteq I \subsetneq I_2$ and $m, n \in J, m \neq n$

$$\begin{aligned}
\bar{S}_{I_1 \cup (I_2 \setminus I), J}(t) &= \bar{S}^m_{I_1 \cup (I_2 \setminus I), J}(t) \\
&= \prod_{\substack{\emptyset \neq L \subseteq J \\ m \in L}} \left(\tilde{g}^{L \cup I_1 \cup (I_2 \setminus I), m}(t)\right)^{(-1)^{|L|-1}} \\
&= \frac{\prod_{\emptyset \neq L \subseteq J \setminus \{n\}} \left(\tilde{g}^{L \cup I_1 \cup (I_2 \setminus I), m}(t)\right)^{(-1)^{|L|-1}}}{\prod_{\substack{\emptyset \neq K \subseteq J \setminus \{n\} \\ m \in K}} \left(\tilde{g}^{K \cup \{n\} \cup I_1 \cup (I_2 \setminus I), m}(t)\right)^{(-1)^{|K|-1}}} \\
&= \frac{\bar{S}^m_{I_1 \cup (I_2 \setminus I), J \setminus \{n\}}(t)}{\bar{S}^m_{I_1 \cup (I_2 \setminus I) \cup \{n\}, J \setminus \{n\}}(t)} = \frac{\bar{S}_{I_1 \cup (I_2 \setminus I), J \setminus \{n\}}(t)}{\bar{S}_{I_1 \cup (I_2 \setminus I) \cup \{n\}, J \setminus \{n\}}(t)}.
\end{aligned}$$

Writing $b \coloneqq \binom{2^{|I_2|}-2}{2}$ and using Lemma 5 for ascending sequences $\emptyset \neq J_1 \subsetneq \cdots \subsetneq J_{|I|} = I \subseteq I_2$ with $|J_I| = |I|$ as well as

1. $\bar{F}_{|J_l|,|I_1 \cup (I_2 \setminus I)|} \equiv \bar{S}_{I_1 \cup (I_2 \setminus I), J_l}$ for $l \in [|I|]$ and
2. $\bar{F}_{|J_l|,|I_1 \cup (I_2 \setminus I) \cup (J_{l+1} \setminus J_l)|} \equiv \bar{S}_{I_1 \cup (I_2 \setminus I) \cup (J_{l+1} \setminus J_l), J_l}$ for $l \in [|I| - 1]$

it follows that

$$\begin{aligned}\mathbb{P}(\hat{A}^{I_1,I_2}) &\leq b \max_{\substack{\emptyset \neq I \subsetneq I_2 \\ \emptyset \neq J_1 \subsetneq \cdots \subsetneq J_{|I|} \\ J_{|I|} = I}} \left(\frac{\bar{S}_{I_1 \cup (I_2 \setminus I), J_1}(t) - \bar{S}_{I_1 \cup (I_2 \setminus I), J_1}(s)}{\prod_{l=1}^{|I|-1} \bar{S}_{I_1 \cup (I_2 \setminus I) \cup (J_{l+1} \setminus J_l), J_l}(s)} \right)^2 \\ &= b \max_{\substack{\emptyset \neq I \subsetneq I_2 \\ \emptyset \neq J_1 \subsetneq \cdots \subsetneq J_{|I|} = I \\ J_1 = \{m\}}} \left(\frac{\tilde{g}^{I_1 \cup (I_2 \setminus I) \cup J_1, m}(t) - \tilde{g}^{I_1 \cup (I_2 \setminus I) \cup J_1, m}(s)}{\prod_{l=1}^{|I|-1} \bar{S}_{I_1 \cup (I_2 \setminus I) \cup (J_{l+1} \setminus J_l), J_l}(s)} \right)^2.\end{aligned}$$

Now let $\emptyset \neq I \subsetneq I_2$, $k = |I_1 \cup (I_2 \setminus I)|$, $J_1 = \{m\}$ and $\pi \in \mathcal{S}_d$ be a permutation fulfilling $\pi(\{1, \ldots, k\}) = I_1 \cup (I_2 \setminus I)$, $\pi(k+1) = m$. Denote with $\tilde{\pi}$ the permutation, which switches the positions of m and $\pi(k)$, i.e. $\tilde{\pi} = \pi(k, k+1)$. Then

$$\begin{aligned}0 &\leq G_{I_1 \cup (I_2 \setminus I) \setminus \{\pi(k)\}, \{m, \pi(k)\}}(s,t) \\ &= \prod_{j=0}^{1} g_{k+j}^{\tilde{\pi}}(t) - g_k^{\pi}(s) g_{k+1}^{\pi}(t) - g_k^{\tilde{\pi}}(s) g_{k+1}^{\tilde{\pi}}(t) + \prod_{j=0}^{1} g_{k+j}^{\pi}(s) \\ &= g_{k+1}^{\tilde{\pi}}(t) \left(g_k^{\tilde{\pi}}(t) - g_k^{\tilde{\pi}}(s) \right) - g_k^{\pi}(s) \left(g_{k+1}^{\pi}(t) - g_{k+1}^{\pi}(s) \right),\end{aligned}$$

which is equivalent to

$$g_{k+1}^{\pi}(t) - g_{k+1}^{\pi}(s) \leq \frac{g_{k+1}^{\tilde{\pi}}(t)}{g_k^{\pi}(s)} \left(g_k^{\tilde{\pi}}(t) - g_k^{\tilde{\pi}}(s) \right).$$

This yields inductively the following inequality

$$g_{k+1}^{\pi}(t) - g_{k+1}^{\pi}(s) \leq \prod_{l=1}^{k} \frac{\tilde{g}^{\pi(\{1,\ldots,l\}) \cup \{m\}, m}(t)}{\tilde{g}^{\pi(\{1,\ldots,l\}), \pi(l)}(s)} \cdot \left(\tilde{g}^{\{m\}, m}(t) - \tilde{g}^{\{m\}, m}(s) \right).$$

Subsequently,

$$\mathbb{P}(\hat{A}^{I_1,I_2}) \leq b p_{I_1,I_2}(s,t) q_{I_2}(s,t)$$

with

$$p_{I_1,I_2}(s,t) \coloneqq \max_{\substack{\emptyset \neq I \subsetneq I_2 \\ \emptyset \neq J_1 \subsetneq \cdots \subsetneq J_{|I|} = I \\ \pi \in \Pi_{I_1,I_2,I} \\ J_1 = \{m\}}} \left\{ \frac{1}{\prod_{l=1}^{|I|-1} \bar{S}_{I_1 \cup (I_2 \setminus I) \cup (J_{l+1} \setminus J_l), J_l}(s)} \right. \\ \left. \times \prod_{l=1}^{|I_1 \cup (I_2 \setminus I)|} \frac{\tilde{g}^{\pi(\{1,\ldots,l\}) \cup \{m\}, m}(t)}{\tilde{g}^{\pi(\{1,\ldots,l\}), \pi(l)}(s)} \right\}^2,$$

where $\Pi_{I_1,I_2,I}$ is the set of permutations fulfilling the conditions stated above and

$$q_{I_2}(s,t) \coloneqq \max_{m \in I_2} \left\{ \tilde{g}^{\{m\},m}(t) - \tilde{g}^{\{m\},m}(s) \right\}^2.$$

Let $s_0 \geq s > t \geq t_0 \geq 0$; the non-increasingness of the involved survival functions $\bar{S}_{I_1 \cup (I_2 \setminus I) \cup (J_{l+1} \setminus J_l), J_l}(s)$, $\tilde{g}^{\pi([[l]]) \cup \{m\}, m}(t)$, and $\tilde{g}^{\pi([[l]]), \pi(l)}(s)$ implies

$$p_{I_1,I_2}(s,t) \leq p_{I_1,I_2}(s_0,t_0), \quad \forall t < s \text{ with } t, s \in [t_0, s_0].$$

Define for $s \geq t \geq 0$

$$\mu_{I_2}(s,t) = \sum_{m \in I_2} \tilde{g}^{\{m\},m}(t) - \tilde{g}^{\{m\},m}(s).$$

As $\tilde{g}^{\{m\},m}, m \in I_2$ are non-negative and non-increasing and $q_{I_2}(s,t) \geq 0$ all summands are non-negative and

$$\mu_{I_2}(s,t) \geq \sqrt{q_{I_2}(s,t)} \geq 0, \ s \geq t \geq 0.$$

Hence

$$\begin{aligned} 0 &\leq G_{I_1,I_2}^{\left\{\hat{g}_{|I_1|+1}^{\pi_J}, \hat{g}_{|I_1|+2}^{\pi_J}, \ldots, \hat{g}_{|I_1 \cup I_2|}^{\pi_J}\right\}_{J \in I_2}}(s,t) \\ &\leq b p_{I_1,I_2}(s_0,t_0) q_{I_2}(s,t) \\ &\leq b p_{I_1,I_2}(s_0,t_0) \mu_{I_2}(s_0,t_0)^2, \quad \forall t < s \text{ with } t, s \in [t_0, s_0]. \end{aligned}$$

Now, the proof proceeds analogously as for copulas in the exchangeable case (see Mai et al. 2016, 1296 sq.) or bivariate exchangeable case (see Durante et al. 2008, 67).

The function $\bar{S}_{I_1,I_2}$ splits in positive and negative powers in the product terms and

$$\begin{aligned}
\bar{S}_{I_1,I_2}(t) &= \prod_{i=1}^{|I_2|}\left(\prod_{\substack{J\subseteq I_2\\ |J|=i, m\in J}} \tilde{g}^{J\cup I_1,m}(t)\right)^{(-1)^{i-1}}\\
&= \frac{\prod_{i=0}^{\lfloor(|I_2|-1)/2\rfloor}\left(\prod_{\substack{J\subseteq I_2\\ |J|=2i+1, m\in J}} \tilde{g}^{J\cup I_1,m}(t)\right)}{\prod_{i=1}^{\lfloor|I_2|/2\rfloor}\left(\prod_{\substack{J\subseteq I_2\\ |J|=2i, m\in J}} \tilde{g}^{J\cup I_1,m}(t)\right)}\\
&\overset{(\star)}{\leq} \frac{\prod_{i=0}^{\lfloor(|I_2|-1)/2\rfloor}\left(\prod_{\substack{J\subseteq I_2\\ |J|=2i+1, m\in J}} \tilde{g}^{J\cup I_1,m}(t_0)\right)}{\prod_{i=1}^{\lfloor|I_2|/2\rfloor}\left(\prod_{\substack{J\subseteq I_2\\ |J|=2i, m\in J}} \tilde{g}^{J\cup I_1,m}(s_0)\right)}\\
&=: p_{\max}^{I_1,I_2}(s_0,t_0),
\end{aligned}$$

where the monotonicity of $\tilde{g}^{I,m}$ is used in $(\star)$. Assume that $\bar{S}_{I_1,I_2}$ is not non-increasing, i.e. there exists $s_0 > t_0 \geq 0$ s.t. $\bar{S}_{I_1,I_2}(s_0) > \bar{S}_{I_1,I_2}(t_0)$.
Case $q_{I_1}(s_0,t_0) = 0$*:* From Eq. (21) we get

$$\begin{aligned}
0 &\leq G_{I_1,I_2}^{\left\{g^{\pi_J}_{|I_1|+1}, g^{\pi_J}_{|I_1|+2},\ldots,g^{\pi_J}_{|I_1\cup I_2|}\right\}_{J\in I_2}}(s_0,t_0)\\
&= \underbrace{\hat{g}^{\pi_\emptyset}_{|I_1|+1}(t_0)g^{\pi_\emptyset}_{|I_1|+2}(t_0)\ldots g^{\pi_\emptyset}_{|I_1\cup I_2|}(t_0)}_{>0}\underbrace{\left(\bar{S}_{I_1,I_2}(t_0)-\bar{S}_{I_1,I_2}(s_0)\right)}_{<0}\\
&< 0
\end{aligned}$$

which is a contradiction.

Case $q_{I_1}(s_0,t_0) > 0$*:* Let

$$a(s_0,t_0) \coloneqq \frac{\bar{S}_{I_1,I_2}(s_0)-\bar{S}_{I_1,I_2}(t_0)}{\mu_{I_2}(s_0,t_0)} > 0$$

then we can write

$$\bar{S}_{I_1,I_2}(t_0)-\bar{S}_{I_1,I_2}(s_0) = -a(s_0,t_0)\mu_{I_2}(s_0,t_0).$$

For all $k \geq 1$, one can find $s_k, t_k \in [t_0, s_0]$ with $s_k > t_k$ and

$$\mu_{I_2}(s_k,t_k) = \frac{\mu_{I_2}(s_0,t_0)}{k} \tag{22}$$

as well as

$$\bar{S}_{I_1,I_2}(t_k)-\bar{S}_{I_1,I_2}(s_k) \leq -a(s_0,t_0)\mu_{I_2}(s_s,t_k).$$

This can be seen by setting $t^{(0,k)} := t_0$, $t^{(k,k)} := s_0$, and

$$t^{(j,k)} := \left(\sum_{m \in I_2} \tilde{g}^{\{m\},m} \right)^{\leftarrow} \left(x^{(j,k)} \right), \quad j \in \{1, \ldots, k-1\},$$

where $\leftarrow$ denotes the generalized inverse for non-increasing functions[10] and for $k \in \{0, \ldots, k\}$

$$x^{(j,k)} := \frac{k-j}{k} \sum_{m \in I_2} \tilde{g}^{\{m\},m}(t_0) + \frac{j}{k} \sum_{m \in I_2} \tilde{g}^{\{m\},m}(s_0).$$

As $\tilde{g}^{\{m\},m}$ are continuous and non-decreasing the generalized inverse is a right-inverse[11] and

$$\begin{aligned} \mu_{I_2}\left(t^{(j,k)}, t^{(j-1,k)}\right) &= \underbrace{\sum_{m \in I_2} \tilde{g}^{\{m\},m}\left(t^{(j-1,k)}\right)}_{=x^{(j-1,k)}} - \underbrace{\sum_{m \in I_2} \tilde{g}^{\{m\},m}\left(t^{(j,k)}\right)}_{=x^{(j,k)}} \\ &= \frac{1}{k} \mu_{I_2}(s_0, t_0). \end{aligned}$$

Assume that for all $j \in \{1, \ldots, k\}$ the following inequality holds

$$\bar{S}_{I_1,I_2}(t^{(j-1,k)}) - \bar{S}_{I_1,I_2}(t^{(j,k)}) > -a(s_0, t_0)\mu_{I_2}(t^{(j,k)}, t^{(j-1,k)}).$$

Then,

$$\begin{aligned} \bar{S}_{I_1,I_2}(t_0) - \bar{S}_{I_1,I_2}(s_0) &= \sum_{j=1}^{k} \bar{S}_{I_1,I_2}(t^{(j-1,k)}) - \bar{S}_{I_1,I_2}(t^{(j,k)}) \\ &> -a(s_0, t_0) \sum_{j=1}^{k} \mu_{I_2}(t^{(j,k)}, t^{(j-1,k)}) = -a(s_0, t_0)\mu_{I_2}(s_0, t_0), \end{aligned}$$

which is a contradiction. Hence, with $t_k = t^{(j-1,k)}$, $s_k = t^{(j,k)}$ for some $j \in \{1, \ldots, k\}$, Eq. (22) is fulfilled and $s_k > t_k$.

Combining Eq. (21) with these results gives for feasible t_k, s_k (chosen as above)

[10] For a non-increasing function f, its generalized inverse is defined by $f^{\leftarrow}(x) := \inf\{x : f(x) \leq y\}$ and for a non-decreasing function f, its generalized inverse is defined by $f^{\leftarrow}(x) := \inf\{y : f(y) \geq x\}$.

[11] If g is a continuous and non-increasing function, then $g^{\leftarrow}(x) = (-g)^{\leftarrow}(-x)$, where the generalized inverse on the l.h.s. is for non-increasing and on the r.h.s. for non-decreasing functions. As $(-g)^{\leftarrow}$ is a right-inverse of $-g$, see Embrechts and Hofert (2013, p.425 sq., prop. 1 (4)), this implies that $g^{\leftarrow}$ is a right-inverse of g.

$$
\begin{aligned}
0 \le\ & G_{I_1,I_2}^{\left\{g_{|I_1|+1}^{\pi_J}, g_{|I_1|+2}^{\pi_J}, \ldots, g_{|I_1\cup I_2|}^{\pi_J}\right\}_{J\in I_2}}(s_k,t_k) \\
=\ & \underbrace{\hat{g}_{|I_1|+1}^{\pi_\emptyset}(t_k)}_{=\frac{g_{|I_1|+1}^{\pi_\emptyset}(t_k)}{\bar{S}_{I_1,I_2}(t_k)}} g_{|I_1|+2}^{\pi_\emptyset}(t_k)\cdot\ldots\cdot g_{|I_1\cup I_2|}^{\pi_\emptyset}(t_k)\cdot \underbrace{\left(\bar{S}_{I_1,I_2}(t_k)-\bar{S}_{I_1,I_2}(s_k)\right)}_{\le -a(s_0,t_0)\frac{\mu_{I_2}(s_0,t_0)}{k}} \\
& + \bar{S}_{I_1,I_2}(s_k) G_{I_1,I_2}^{\left\{\hat{g}_{|I_1|+1}^{\pi_J}, g_{|I_1|+2}^{\pi_J}, \ldots, g_{|I_1\cup I_2|}^{\pi_J}\right\}_{J\in I_2}}(s_k,t_k) \\
\le\ & \frac{g_{|I_1|+1}^{\pi_\emptyset}(s_0)}{p_{\max}^{I_1,I_2}(s_0,t_0)} g_{|I_1|+2}^{\pi_\emptyset}(s_0)\cdot\ldots\cdot g_{|I_1\cup I_2|}^{\pi_\emptyset}(s_0) \\
& \times\left(-a(s_0,t_0)\mu_{I_2}(s_0,t_0)\frac{1}{k}\right) \\
& + b p_{\max}^{I_1,I_2}(s_0,t_0) p_{I_1,I_2}(s_0,t_0)\mu_{I_2}(s_0,t_0)^2\frac{1}{k^2}.
\end{aligned}
$$

In particular, if the latter inequality is multiplied by k and the limit $k\to\infty$ is taken, then

$$
0 \le -\underbrace{\frac{1}{p_{\max}^{I_1,I_2}(s_0,t_0)}}_{>0}\underbrace{a(s_0,t_0)}_{>0}\underbrace{\mu_{I_1,I_2}(s_0,t_0)}_{>0}\underbrace{\prod_{j=1}^{|I_2|} g_{|I_1|+j}^{\pi_\emptyset}(s_0)}_{>0} < 0,
$$

which leads to a contradiction. □

B Proofs of supporting lemmas

Proof of Lemma 1 Let I_1, I_2, and π fulfill the usual conditions with $|I_2| = 2$ and $|I_1| = i-2$ and define $\tilde{\pi} = \pi(i-1,i)$.

Right-continuity Let $s+h > s > t \ge 0$. As $G_{I_1,I_2}(s,t)$ is right-continuous in s it holds that

$$
0 = \lim_{h\searrow 0} G_{I_1,I_2}(s+h,t) - G_{I_1,I_2}(s,t) \overset{\text{IH}}{=} \underbrace{g_{i-1}^{\pi}(s)}_{\overset{\text{IH}}{>0}} \lim_{h\searrow 0}\left(g_i^{\pi}(s+h) - g_i^{\pi}(s)\right),
$$

where it is used that under the induction hypothesis all but two terms cancel out.

Left-limits Let $s > s-h > t \ge 0$. As $G_{I_1,I_2}(s,t)$ and $g_{i-1}^{\pi_\emptyset}(s)$, $\pi_\emptyset \in \{\pi,\tilde{\pi}\}$ have left-limits in s and g_{i-1}^{π} is positive by induction hypothesis it follows that g_i^{π} has left-limits:

$$\lim_{h\searrow 0} g_i^{\pi}(s-h) = \lim_{h\searrow 0}\left(\frac{G_{I_1,I_2}(s-h,t) - g_{i-1}^{\tilde{\pi}}(t)g_i^{\tilde{\pi}}(t)}{g_{i-1}^{\pi}(s-h)} - \frac{-g_{i-1}^{\pi}(s-h)g_i^{\pi}(t) - g_{i-1}^{\tilde{\pi}}(s-h)g_i^{\tilde{\pi}}(t)}{g_{i-1}^{\pi}(s-h)}\right).$$

Non-increasingness Now, let I_1, I_2, and π fulfill the usual conditions with $I_2 = \{\pi(i)\}$ and $I_1 = \pi([i-1])$. As G_{I_1,I_2} is non-negative, it holds for all $s > t \geq 0$ that

$$0 \leq G_{I_1,I_2}(s,t) = g_i^{\pi}(t) - g_i^{\pi}(s).$$

□

Proof of Lemma 2 For $\pi \in \mathcal{S}_d$, due to Remark 2 and Lemma 1, it follows that the functions g_i^{π}, $i = 1,\ldots,d$ are positive, continuous functions on $\mathbb{R}_+$. Hence $\bar{S}^m_{I_1,I_2}$ is an $\mathbb{R}_+$-valued, positive, and continuous function for every I_1, I_2 fulfilling the usual conditions with $m \in I_2$.

In the following, it is proven, by induction over $|I_2|$, that Eq. (19) holds and furthermore, that for all I_1 and I_2 fulfilling the usual conditions

$$\prod_{i=1}^{|I_2|} g_{|I_1|+i}^{\tilde{\pi}}(t) = \prod_{i=1}^{|I_2|} g_{|I_1|+i}^{\hat{\pi}}(t), \quad \forall t \geq 0 \tag{23}$$

for all $\tilde{\pi}, \hat{\pi} \in \mathcal{S}_d$ fulfilling $\pi([|I_1|]) = I_1$ and $\pi([|I_1 \cup I_2|]\backslash[|I_1|]) = I_2$ for $\pi \in \{\tilde{\pi}, \hat{\pi}\}$. For $|I_2| = 1$ both claims are naturally fulfilled. Let both claims be fulfilled for $|I_2| < p$ and let I_1, I_2 as well as π fulfill the usual conditions with $|I_2| = p$, $m \in I_2$ as well as $\pi(|I_1|+1) = m$, then for $t \geq 0$

$$\begin{aligned}\prod_{\emptyset\neq J\subseteq I_2} \bar{S}^{\pi(\min_{j\in J}\pi^{-1}(j))}_{I_1\cup(I_2\backslash J),J}(t) &\overset{(\star)}{=} \prod_{i=1}^{|I_2|} \prod_{\substack{J\subseteq\pi(\{|I_1|+i,\ldots,|I_1\cup I_2|\})\\ \pi(|I_1|+i)\in J}} \bar{S}^{\pi(|I_1|+i)}_{I_1\cup(I_2\backslash J),J}(t)\\ &= \prod_{i=1}^{|I_2|} \prod_{\substack{J\subseteq\pi(\{|I_1|+i,\ldots,|I_1\cup I_2|\})\\ \pi(|I_1|+i)\in J}} \\ &\quad\times \prod_{\substack{L\subseteq J\\ \pi(|I_1|+i)\in L}} \left(\tilde{g}^{L\cup I_1\cup(I_2\backslash J),\pi(|I_1|+i)}(t)\right)^{(-1)^{|L|-1}},\end{aligned}$$

where the factors in $(\star)$ are regrouped in a similar sense as for the alternative representation for the GMO survival function.

Now for $i \in [d]$ fix $\pi(\{1,\ldots,|I_1|+i\}) \subseteq K \subseteq I_1 \cup I_2$ and define $k = |K|$ as well as $1 \leq l \leq k - |I_1| - i + 1$. The expression $\tilde{g}_k^{K,\pi(|I_1|+i)}(t)$ with exponent $(-1)^{l-1}$

appears $\binom{k-|I_1|-i}{l-1}$ times and the overall exponent for $\tilde{g}_k^{K,\pi(|I_1|+i)}$ is

$$\sum_{l=1}^{k-i-|I_1|+1}(-1)^{l-1}\binom{k-i-|I_1|}{l-1}=\begin{cases}1, & k=|I_1|+i\\ 0, & \text{else}\,.\end{cases}$$

Hence, as it holds for $k=|I_1|+i$ that $K=\pi(\{1,\ldots,|I_1|+i\})$ and

$$\prod_{\emptyset\neq J\subseteq I_2}\bar{S}_{I_1\cup(I_2\setminus J),J}^{\pi(\min_{j\in J}\pi^{-1}(j))}(t)=\prod_{i=1}^{|I_2|}g_{|I_1|+i}^{\pi}(t)$$

or equivalently,

$$\bar{S}_{I_1,I_2}^{m}(t)=\frac{\prod_{i=1}^{|I_2|}g_{|I_1|+i}^{\pi}(t)}{\prod_{\emptyset\neq J\subsetneq I_2}\bar{S}_{I_1\cup(I_2\setminus J),J}^{\pi(\min_{j\in J}\pi^{-1}(j))}(t)}. \tag{24}$$

By induction, the factors of the denominator of the r.h.s. in Eq. (24), $\bar{S}_{I_1\cup(I_2\setminus J),J}^{\pi(\min_{j\in J}\pi^{-1}(j))}$, are independent of $\pi(\min_{j\in J}\pi^{-1}(j))$ and subsequently also of m. Moreover, for arbitrary I_1, I_2 and $\{\pi_J\}_{J\subseteq I_2}$ fulfilling the usual conditions and $s\geq 0$

$$\prod_{j=1}^{|I_2|}g_{|I_1|+j}^{\pi_{I_2}}(s)=(-1)^{|I_2|}\left(G_{I_1,I_2}^{\{\pi_J\}_{J\subseteq I_2}}(s,0)-\sum_{J\subsetneq I_2}(-1)^{|J|}\prod_{j=1}^{|J|}g_{|I_1|+j}^{\pi_J}(s)\right).$$

By induction and assumption, the r.h.s. does not depend on the specific family of permutations, $\{\pi_J\}_{J\subseteq I_2}$, chosen, therefore Eq. (23) holds for $|I_2|=p$. In conclusion, the nominator in Eq. (24) does not depend on the specific π, and subsequently m, chosen and Eq. (19) holds for $|I_2|=p$. □

Proof of Lemma 3 As in the proof of 4. to 1. one can derive analogously for $\boldsymbol{t}\geq 0$ and $\pi\in\mathcal{S}_d$ with $t_{\pi(1)}\geq\cdots\geq t_{\pi(d)}$ as well as $\pi(\{1,\ldots,|I_1|\})=I_1$ and $\pi(\{|I_1|+1,\ldots,|I_1\cup I_2|\})=I_2$ that

$$\mathbb{P}\left(\check{\tau}_j>t_j\ \forall j\in I_2\right)=\prod_{j=|I_1|+1}^{|I_1\cup I_2|}g_j^{\pi}\left(t_{\pi(j)}\right)=\prod_{j=1}^{|\check{I}_2|}\check{g}_j^{\check{\pi}}\left(\check{t}_{\check{\pi}(j)}\right),$$

where for $\check{I}_2=\{1,\ldots,|I_2|\}$, $\check{\pi}\in\mathcal{S}_{|I_2|}$ is defined by

$$\pi(|I_1|+j)=i_{\check{\pi}(j)},\quad\forall j\in\check{I}_2,\ I_2=\{i_1,\ldots,i_{|I_2|}\}$$

and $\check{g}_j^{\check{\pi}}:=g_{|I_1|+j}^{\pi}$ as well as $\check{t}_{\check{\pi}(j)}:=t_{\pi(|I_1|+j)}$. Then, it holds for all $0\leq t<s$ that

$$\mathbb{P}\left(\check{\tau}_j \in (t,s]\ \forall j \in I_2\right) = \check{G}_{\emptyset,\check{I}_2}(s,t) = G_{I_1,I_2}(s,t),$$

where $\check{G}_{\emptyset,\check{I}_2}$ corresponds to Eq. (8) w.r.t. $\{\check{g}_j^{\check{\pi}}\}_{j\in\check{I}_2,\check{\pi}\in\mathcal{S}_{|I_2|}}$. □

Proof of Lemma 4 Every summand corresponding to a non-empty interval $\emptyset \neq J \subseteq I_2$ contains a term $g^{\pi_J}_{|I_1|+1}(s)$. Therefore the result follows by multiplying G_{I_1,I_2} with $\frac{g^{\pi_\emptyset}_{|I_1|+1}(s)}{g^{\pi_\emptyset}_{|I_1|+1}(s)}$ and its reciprocal, whereas the first summand in Eq. (20) is a correction term for the summand belonging to $J = \emptyset$. □

Proof of Lemma 5 This is a direct corollary of Mai et al. (2016, lem. B.2 on p. 1295). □

References

Berg C, Christensen JPR, Ressel P (1984) Harmonic analysis on semigroups. Graduate texts in mathematics, vol 100. Springer, New York. https://doi.org/10.1007/978-1-4612-1128-0

Bielecki TR, Cousin A, Crépey S, Herbertsson A (2013) A bottom-up dynamic model of portfolio credit risk: part II: common-shock interpretation, calibration and hedging issues. available at SSRN: https://ssrn.com/abstract=2245130 or https://doi.org/10.2139/ssrn.2245130

Durante F, Quesada-Molina JJ, Úbeda Flores M (2007) On a family of multivariate copulas for aggregation processes. Inf Sci 177(24):5715–5724. https://doi.org/10.1016/j.ins.2007.07.019

Durante F, Kolesárová A, Mesiar R, Sempi C (2008) Semilinear copulas. Fuzzy Sets Syst 159(1):63–76. https://doi.org/10.1016/j.fss.2007.09.001

Elouerkhaoui Y (2007) Pricing and hedging in a dynamic credit model. Int J Theor Appl Finance 10(04):703–731. https://doi.org/10.1142/S0219024907004408

Embrechts P, Hofert M (2013) A note on generalized inverses. Math Methods Oper Res 77(3):423–432

Giesecke K (2003) A simple exponential model for dependent defaults. J Fixed Income 13(3):74–83. https://doi.org/10.3905/jfi.2003.319362

Gnedin A, Pitman J (2005) Regenerative composition structures. Ann Probab 33(2):445–479. https://doi.org/10.1214/009117904000000801

Joe H (1997) Multivariate models and multivariate dependence concepts. CRC Press, Boca Raton

Klein JP, Keiding N, Kamby C (1989) Semiparametric Marshall–Olkin models applied to the occurrence of metastases at multiple sites after breast cancer. Biometrics 45(4):1073–1086

Li H (2009) Orthant tail dependence of multivariate extreme value distributions. J Multivar Anal 100(1):243–256. https://doi.org/10.1016/j.jmva.2008.04.007

Li X, Pellerey F (2011) Generalized Marshall–Olkin distributions and related bivariate aging properties. J Multivar Anal 102(10):1399–1409. https://doi.org/10.1016/j.jmva.2011.05.006

Lin J, Li X (2014) Multivariate generalized Marshall–Olkin distributions and copulas. Methodol Comput Appl Probab 16(1):53–78. https://doi.org/10.1007/s11009-012-9297-4

Lindskog F, McNeil AJ (2003) Common Poisson shock models: applications to insurance and credit risk modelling. Astin Bull 33(2):209–238. https://doi.org/10.1017/S0515036100013441

Mai JF (2010) Extendibility of Marshall–Olkin distributions via Lévy subordinators and an application to portfolio credit risk. Dissertation, Technische Universität München. http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20100628-969547-1-8

Mai JF (2014) Multivariate exponential distributions with latent factor structure and related topics. Habilitation thesis, Technische Universität München

Mai JF, Scherer M (2017) Simulating copulas: stochastic models, sampling algorithms and applications. Series in quantitative finance, vol 6, 2nd edn. World Scientific, Singapore

Mai JF, Schenk S, Scherer M (2016) Exchangeable exogenous shock models. Bernoulli 22(2):1278–1299. https://doi.org/10.3150/14-BEJ693

Marshall AW (1996) Copulas, marginals, and joint distributions. In: Rüschendorf L, Schweizer B, Taylor MD (eds) Distributions with fixed marginals and related topics. Lecture notes, monograph series, vol 28. Institute of Mathematical Statistics, Hayward, pp 213–222. https://doi.org/10.1214/lnms/1215452620
Marshall AW, Olkin I (1967) A multivariate exponential distribution. J Am Stat Assoc 62(317):30–44. https://doi.org/10.2307/2282907
Muliere P, Scarsini M (1987) Characterization of a Marshall–Olkin type class of distributions. Ann Inst Stat Math 39(1):429–441. https://doi.org/10.1007/BF02491480
Schenk S (2016) Exchangeable exogenous shock models. Dissertation, Technical University of Munich, Munich. Available at https://mediatum.ub.tum.de/doc/1278138/1278138.pdf
Schilling RL, Song R, Vondracek Z (2012) Bernstein functions. De Gruyter studies in mathematics, vol 37, 2nd edn. De Gruyter, Berlin
Schweizer B, Sklar A (1983) Probabilistic metric spaces. Elsevier/North-Holland, New York
Sklar A (1959) Fonctions de répartition a n dimensions et leurs marges. Publ Inst Stat Univ Paris 8:229–231
Sun Y, Mendoza-Arriaga R, Linetsky V (2017) Marshall–Olkin distributions, subordinators, efficient simulation, and applications to credit risk. Adv Appl Probab 49(2):481–514

B.2 The deFinetti representation of generalized Marshall–Olkin sequences

The deFinetti representation of generalized Marshall–Olkin sequences

Henrik Sloot

The article [4] provides a representation theorem for exchangeable generalized Marshall–Olkin sequences. According to *de Finetti's theorem*, every non-negative, exchangeable sequence $\tau_1, \tau_2, \ldots$ defines a subordinator Λ associated with its *directing measure*. I prove that these are precisely additive subordinators for exchangeable generalized Marshall–Olkin sequences. Additionally, I show how to extend the probability space to obtain a de Finetti representation of an exchangeable generalized Marshall–Olkin sequence as follows:

$$\tau_i = \inf\left\{t > 0 \,:\, \Lambda_t \geq E_i\right\} \text{ a.s.}, \quad i \in \mathbb{N},$$

where $E_1, E_2, \ldots$ is an iid unit exponential sequence independent of Λ.

The introduction identifies the research gap by recalling past research on exchangeable Marshall–Olkin sequences, exchangeable generalized Marshall–Olkin vectors, and the additive-frailty model. Specifically, the contribution of the article is the backward direction of the main result, abridged as follows: *given an additional independent sequence of iid uniform random variables, an exchangeable sequence has a de Finetti representation as above with an additive subordinator if and only if each finite margin has a generalized Marshall–Olkin distribution.*

Section 2 introduces the required mathematical background of Bernstein functions, additive subordinators, and used notation. Following, Section 3 motivates the generalized Marshall–Olkin distribution by allowing non-constant hazard rates for the shock time rates in the exogenous shock model representation. Furthermore, I recall the required facts about the exchangeable subclass.

Section 4 discusses de Finetti representations of exchangeable sequences of random variables in general, presents the main result, and proves it. My proof for the backward direction can be sketched as follows: In the first part, I use de Finetti's theorem to obtain an almost surely unique random distribution function such that by conditioning on it, the sequence becomes iid. Afterward, I transform the sequence by mapping each element uniformly at random to a value in the random preimage of the element under the corresponding random quantile function. Choosing a value uniformly at random from the random preimage requires the additional iid sequence of uniform random variables. The transformed sequence is iid uniform in the unit interval and independent of the random distribution function, and the subordinator and the iid unit exponential sequence are obtained by simple transformations. In the second part, I use a *uniqueness-in-distribution* argument to show that the subordinator is additive. For this, I use a lemma from the previous section, linking the survival function of d-variate exchangeable generalized Marshall–Olkin distributions to families of d-monotone sequences. Overall, I can derive a family of Bernstein functions for which I subsequently show that they are Laplace exponents of an additive subordinator distribution. Finally, I conclude that, by construction, the distribution of corresponding exchangeable generalized Marshall–Olkin sequences is identical to the original.

In the remainder of the article, I provide a detailed step-by-step example of how to recover the implied subordinator for an exchangeable generalized Marshall–Olkin sequence generated by an exogenous shock model with individual shocks and a global shock. In particular, I outline why the additional, independent sequence of iid uniform random variables is required to obtain a de Finetti representation if the subordinator jumps.

The article contains two appendices, one on exchangeable sequences and de Finetti's theorem and another on Bernstein functions and completely monotone sequences.

Statement of individual contribution

I, Henrik Sloot, am the sole author of this article. The article originated from my master thesis [87] that contained a result, abridged as follows: *the law of an exchangeable sequence has a stochastic de Finetti representation with an additive subordinator if and only if each finite margin is an exchangeable generalized*

Marshall–Olkin distribution. This result is solely on the distribution level. In particular, it does not state that any such sequence defines an additive subordinator in the same probability space. Moreover, it does not show how a de Finetti representation of an existing sequence can be obtained by extending the probability space slightly. My main theorem of this article, developed during my doctoral studies, significantly extends this by working on the probability space level. In particular, I derive that every exchangeable generalized Marshall–Olkin sequence defines an additive subordinator. Furthermore, I outline how the probability space needs to be extended to obtain a de Finetti representation. The second part of the proof of Theorem 1 is conceptually similar to the corresponding proof in my master thesis. However, I rewrote the proof significantly to be simpler, clearer, and more concise. For my doctoral thesis, to explicitly prevent double counting, only those parts of this article that go beyond my Master's thesis are to be accredited.

Permission to include

H. Sloot. “The deFinetti representation of generalised Marshall–Olkin sequences”. In: *Dependence Modeling* 8.1 (2020), pp. 107–118. DOI: 10.1515/demo-2020-0006

Reprinted according to the *CC BY 4.0 license*;[2] see human-readable summary included below.

This page is available in the following languages:

Creative Commons License Deed

Attribution 4.0 International (CC BY 4.0)

This is a human-readable summary of (and not a substitute for) the license.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material

for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

[2] https://creativecommons.org/licenses/by/4.0/.

Research Article **Open Access**

Henrik Sloot*

The deFinetti representation of generalised Marshall–Olkin sequences

https://doi.org/10.1515/demo-2020-0006
Received January 28, 2020; accepted May 22, 2020

Abstract: We show that each infinite exchangeable sequence $\tau_1, \tau_2, \ldots$ of random variables of the generalised Marshall–Olkin kind can be uniquely linked to an additive subordinator via its deFinetti representation. This is useful for simulation, model estimation, and model building.

Keywords: Shock model, Stochastic processes, Survival analysis, Marshall–Olkin, de Finetti

MSC: 60G09, 60G51, 62H05

1 Introduction

The Marshall–Olkin distribution was introduced by eponymous authors in [21] as the multivariate exponential distribution satisfying a strong multivariate lack-of-memory property. A random vector $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_d)'$ has a *Marshall–Olkin distribution* if non-negative parameters $\lambda_I, \emptyset \neq I \subseteq \{1, \ldots, d\}$, exist such that $\boldsymbol{\tau}$ has the survival-function

$$\bar{F}(\boldsymbol{t}) = \exp\left\{ - \sum_{\emptyset \neq I \subseteq \{1,\ldots,d\}} \lambda_I \max_{i \in I} t_i \right\}, \quad \forall \boldsymbol{t} = (t_1, \ldots, t_d)' \geq 0, \tag{1}$$

and the parameters $\lambda_I, \emptyset \neq I \subseteq \{1, \ldots, d\}$, fulfil the condition

$$\sum_{I \ni i} \lambda_I > 0, \quad \forall i \in \{1, \ldots, d\}. \tag{2}$$

A simple calculation shows that the sums in Eq. (2) correspond to the rates of the exponentially distributed univariate margins τ_i, $i \in \{1, \ldots, d\}$, respectively. Hence, the condition in Eq. (2) ensures that $\tau_i < \infty$ a.s. for all $i \in \{1, \ldots, d\}$.

In [21], the authors proposed the *exogenous shock model* as a natural stochastic model for the Marshall–Olkin distribution. This model is based on independent, exponentially distributed random times corresponding to the failure of multiple components of a system at once. In particular, for $\lambda_I \geq 0$, $\emptyset \neq I \subseteq \{1, \ldots, d\}$, fulfilling the condition in Eq. (2), let $E_I \sim \mathrm{Exp}(\lambda_I)$ be independent exponentially distributed random variables with rates $\lambda_I, \emptyset \neq I \subseteq \{1, \ldots, d\}$, respectively, where we use the convention that an exponentially distributed random variable with rate zero is almost surely infinite. Define $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_d)'$ by

$$\tau_i := \min\left\{ E_I \,:\, I \ni i \right\}, \quad i \in \{1, \ldots, d\}. \tag{3}$$

Then $\boldsymbol{\tau}$ has a Marshall–Olkin distribution with parameters $\lambda_I, \emptyset \neq I \subseteq \{1, \ldots, d\}$.

We are interested in exchangeable random vectors and sequences of (generalised) Marshall–Olkin kind. These subclasses have been intensively studied in the last decade for the *classical* Marshall–Olkin distribution.

***Corresponding Author: Henrik Sloot:** Technical University of Munich, Parkring 11, 85748 Garching-Hochbrück, E-mail: henrik.sloot@tum.de

- *The exchangeable subclass* is studied in [11, Chp. 3]. The author has proven that exchangeability corresponds to the property
$$\lambda_I = \lambda_J, \quad \forall \emptyset \neq I, J \subseteq \{1, \dots, d\} \ : \ |I| = |J|.$$
Furthermore, he has shown that the survival function in Eq. (1) of an exchangeable Marshall–Olkin distribution can be reparametrised as follows:
$$\bar{F}(\boldsymbol{t}) = \exp\left\{-\sum_{i=1}^{d} a_{i-1} t_{[i]}\right\}, \quad \forall \boldsymbol{t} = (t_1, \dots, t_d)' \geq 0, \tag{4}$$
where $t_{[1]} \geq \cdots \geq t_{[d]}$ is $\boldsymbol{t}$ in descending order. The sequence $a_0, a_1, \dots a_{d-1}$ is defined by
$$a_{i-1} = \sum_{j=0}^{d-i} \binom{d-i}{j} \lambda_{j+1}, \quad i \in \{1, \dots, d\},$$
where $\lambda_i = \lambda_I$ for $i = |I|$. Finally, he provides a characterisation theorem that states that a function $\bar{F}$ of the form of Eq. (4) is a survival function if, and only if, the sequence $a_0, a_1, \dots, a_{d-1}$ is *d-monotone*. A sequence $a_0, a_1, \dots, a_{d-1}$ is called *d-monotone* if $(-1)^{i-1}\Delta^{i-1} a_{d-i} \geq 0$ for $i = 1, \dots, d$. In this case, the author shows that $\lambda_i = (-1)^{i-1}\Delta^{i-1} a_{d-i}$, $i \in \{1, \dots, d\}$.
- The *extendible* subclass is studied in [11, Chp. 4]. A d-variate Marshall–Olkin distributed random vector $\boldsymbol{\tau} = (\tau_1, \dots, \tau_d)'$ is called *extendible (in the class of Marshall–Olkin distributions)* if an exchangeable sequence of random variables $\{\tilde{\tau}_i\}_{i \in \mathbb{N}}$ exists such that each d-variate subsequence is equal in law to $\boldsymbol{\tau}$ and each finite subsequence has a Marshall–Olkin distribution. The author of the aforementioned reference found a unique link between extendible Marshall–Olkin distributions and Lévy subordinators via a deFinetti representation. In particular, he has shown that an infinite exchangeable Marshall–Olkin sequence $\{\tau_i\}_{i \in \mathbb{N}}$ is conditionally iid and can be written as
$$\tau_i = \inf\{t > 0 \ : \ \Lambda_t \geq E_i\}, \quad i \in \mathbb{N}, \tag{5}$$
where E_i are iid unit exponential random variables independent of a Lévy subordinator $\{\Lambda_t\}_{t \geq 0}$ on $[0, \infty]$. This model is also called the *Lévy frailty model*.

A natural generalisation of the classical Marshall–Olkin distribution is achieved if we allow non-constant hazard rates in the exogenous shock model construction in Eq. (3), see [10]. This means that we replace $\lambda_I \cdot \max_{i \in I} t_i$ with a cumulative hazard rate function $H_I(\max_{i \in I} t_i)$ and the exponential shocks $E_I \sim \text{Exp}(\lambda_I)$ in Eq. (3) with $Z_I \sim 1 - \exp\{-H_I\}$, $\emptyset \neq I \subseteq \{1, \dots, d\}$, respectively. A *cumulative hazard rate function* is a non-negative, non-decreasing, and continuous function on the non-negative half-line that starts in zero. Previous works exist on special cases of this generalisation, e.g. [9], which discusses the bivariate case, and [22], which assumes that $H_I(t) \equiv \lambda_I H(t)$.

- The exchangeable generalised Marshall–Olkin distribution and the exchangeable exogenous shock model are studied in [25]. Similar to the classical Marshall–Olkin case, the author has proven that exchangeability corresponds to the property
$$H_I(t) = H_J(t), \quad \forall t > 0, \ \forall \emptyset \neq I, J \subseteq \{1, \dots, d\} \ : \ |I| = |J|.$$
Furthermore, he has shown that a reparametrisation is possible, similar to the classical Marshall–Olkin case, by replacing $a_{i-1} \cdot t_{[i]}$ by $A_{i-1}(t_{[i]})$, $i \in \{1, \dots, d\}$, in Eq. (4). He also provides an analytical characterisation, which is discussed in Section 3.
- The extendible subclass of generalised Marshall–Olkin distributions is studied in [14] and [25, Sec. 3]. In [14, Prop. 3.1], it is shown that if the subordinator Λ in Eq. (5) is assumed to be an additive subordinator in $[0, \infty]$, then each finite margin of $\{\tau_i\}_{i \in \mathbb{N}}$ has an extendible generalised Marshall–Olkin distribution. We call this stochastic model the *additive-frailty model*.

Contribution:
This article provides the following novel result, which was posed as an open problem for further research in [25, p. 147 sq.]: every exchangeable sequence $\tau_1, \tau_2, \ldots$ with finite margins of generalised Marshall–Olkin type has an implicit representation as an additive frailty model. In particular, an additive subordinator Λ and an iid sequence of unit exponential random variables $E_1, E_2, \ldots$, independent of Λ, exist such that Eq. (5) is fulfilled almost surely. Recall that the converse of this statement was proven in [14, Prop. 3.1]. Consequently, we complete this result and establish a novel one-to-one connection between sequences of generalised Marshall–Olkin type and additive subordinators.

The article is structured as follows: we introduce the mathematical background and notation in Section 2, we summarise existing results on exchangeable generalised Marshall–Olkin distributions in Section 3, and we present the main result in Section 4. In Section 5, we conclude the article. The main proof requires some technical results involving exchangeable sequences and Bernstein functions. For the interested reader, we summarise the required background in Appendices A and B.

2 Mathematical background and notation

In this section, we give a short overview of the required mathematical background and the used notation.

We assume basic knowledge of the theory on multivariate distribution functions and probability theory. Furthermore, we assume that the reader is familiar with the Lévy–Khintchine characterisation of additive subordinators. *Additive processes* are real-valued, stochastic processes, which are defined on the non-negative half-line, start at zero, have independent increments, and have càdlàg path. An *additive subordinator* is a non-decreasing additive process which tends almost surely to infinity. Excellent books on additive processes and Lévy processes in particular are [2, 24]. We deviate slightly from the standard theory by allowing the additive subordinator to jump to an absorbing point associated with ∞ at a random time, which is independent from the subordinator. The corresponding (cumulative) hazard rate is called *(cumulative) killing hazard rate* and is equal to the zero function if almost surely no killing occurs. The Lévy–Khintchine characterisation states that each additive subordinator is uniquely determined in law by its family of Laplace exponents. These Laplace exponents are from the family of *Bernstein functions*, hereafter denoted by $\mathcal{BF}$. A function $\psi : (0, \infty) \to (0, \infty)$ is a Bernstein function if it is infinitely often differentiable and has the following property

$$(-1)^{n-1}\psi^{(n)}(x) \geq 0, \quad \forall x > 0,\ n \in \mathbb{N}.$$

One can show, see e.g. [3, Prop. 6.12] and [26, Thm. 3.2], that a function $\psi : (0, \infty) \to (0, \infty)$ is a Bernstein function if, and only if,

$$(-1)^{n-1}\Delta^n\psi(x) \geq 0, \quad \forall x > 0,\ n \in \mathbb{N}.$$

Here, Δ is the forward iterated difference operator. A Bernstein function ψ is assumed to be extended to the domain $[0, \infty)$ by the convention $\psi(0) = 0$. Excellent books on Bernstein functions are [3, 26].

We denote random variables with capital or Greek letters, e.g. X or τ, and (random) vectors with bold letters, e.g. $\boldsymbol{X}$, $\boldsymbol{\tau}$, or $\boldsymbol{t}$. We write $X \sim F$ if X has the distribution function F. We assume that operators are applied component-wise to vectors. That means $\boldsymbol{\tau} > \boldsymbol{t}$ is equivalent to $\tau_i > t_i$ for all $i \in \{1, \ldots, d\}$. Finally, we denote the descending order of a vector $\boldsymbol{t} \in [0, \infty)^d$ by $t_{[1]} \geq \cdots \geq t_{[d]}$.

We denote the class of continuous, real functions by $\mathcal{C}^{(0)}$, we write $\Delta f \geq 0$ if the function f is non-decreasing everywhere, and we use the notation $f(x-) := \lim_{y \nearrow x} f(y)$ as well as $f(x+) := \lim_{y \searrow x} f(y)$. Finally, for a real number x, we denote the smallest integer i with $i \geq x$ by $\lceil x \rceil$.

3 Exchangeable generalised Marshall–Olkin distributions

In this section, we give a short introduction into exchangeable generalised Marshall–Olkin distributions. For a more detailed treatment of the exchangeable subclass, see [14, 25].

We generalise the *classical* Marshall–Olkin distribution by allowing arbitrary continuous, cumulative hazard rate functions in the exogenous shock model. This is equivalent to having continuous, non-negative, and unbounded shock-times. For this, we define the class of continuous, cumulative hazard rate functions $\mathcal{H}$ and its unbounded subclass $\mathcal{H}_0$ by

$$\mathcal{H} := \left\{ H : [0,\infty) \to [0,\infty) \ : \ H \in \mathcal{C}^{(0)},\ \Delta H \geq 0,\ H(0) = 0 \right\}$$

and

$$\mathcal{H}_0 := \left\{ H \in \mathcal{H} \ : \ H(\infty-) = \infty \right\}.$$

We say that a random vector $\boldsymbol{\tau} \in [0,\infty)^d$ has a *generalised Marshall–Olkin distribution* if functions $H_I \in \mathcal{H}$, $\emptyset \neq I \subseteq \{1,\dots,d\}$, exist such that $\boldsymbol{\tau}$ has survival function

$$\bar{F}(\boldsymbol{t}) = \exp\left\{ - \sum_{\emptyset \neq I \subseteq \{1,\dots,d\}} H_I\left(\max_{i \in I} t_i \right) \right\}, \quad \forall \boldsymbol{t} \geq 0,$$

and the hazard rate functions fulfil the condition

$$\sum_{I \ni i} H_I \in \mathcal{H}_0, \quad \forall i \in \{1,\dots,d\}.$$

This condition, which generalises the condition in Eq. (2), is equivalent to the margins being almost surely finite, since $\sum_{I \ni i} H_I$ are the marginal cumulative hazard rates. With a simple calculation, we can establish a generalised version of the exogenous shock model in Eq. (3) for generalised Marshall–Olkin distributions by replacing $\lambda_I \cdot \max_{i \in I} t_i$ with $H_I(\max_{i \in I} t_i)$ and E_I with $Z_I \sim 1 - \exp\{-H_I\}$, $\emptyset \neq I \subseteq \{1,\dots,d\}$, see [14, Proof of Thm. 1.1 (iv)⇒ (i)].

Below, we present a characterisation of exchangeable generalised Marshall–Olkin distributions. We know from [14, Prop. 2.1] that, similar to the classical Marshall–Olkin case, exchangeability is equivalent to the property $H_I = H_J$ for all $\emptyset \neq I, J \subseteq \{1,\dots,d\}$ with $|I| = |J|$. Furthermore, the following characterisation result has been proven in [14]:

Lemma 1 ([14, Thm. 1.1])**.** *Let $\bar{F} : [0,\infty)^d \to [0,1]$ be a function such that functions $A_0,\dots,A_{d-1} \in \mathcal{H}$ with $A_0 \in \mathcal{H}_0$ and $A_i(0) = 0$ with*

$$\bar{F}(\boldsymbol{t}) = \exp\left\{ - \sum_{i=1}^{d} A_{i-1}\left(t_{[i]}\right) \right\}, \quad \forall \boldsymbol{t} \geq 0,$$

exist, where $t_{[1]} \geq \cdots \geq t_{[d]}$ is $\boldsymbol{t}$ in descending order. Then the following statements are equivalent:

1. *$\bar{F}$ is the survival function of a random vector on $[0,\infty)^d$.*
2. *It holds that $H_i : [0,\infty) \to [0,\infty)$, $t \mapsto (-1)^{i-1}\Delta^{i-1}A_{d-i}(t) \in \mathcal{H}$ for all $i \in \{1,\dots,d\}$, where the difference operator is understood to be applied to the (finite) sequence $A_0(t),\dots,A_{d-1}(t)$ for fixed $t \geq 0$.*

Finally, we can construct a random vector $\boldsymbol{\tau}$ with survival function $\bar{F}$ via an exogenous shock model with $H_I := H_i$ if $|I| = i$, $\emptyset \neq I \subseteq \{1,\dots,d\}$.

Proof of Lemma 1. This is a direct corollary of [14, Thm. 1.1]. However, since we changed the notation, we will give a short explanation: if we take the standardisation of the margins into account, the aforementioned result affirms that the first statement of this lemma is equivalent to $H_i \in \mathcal{H}$ for all $i \in \{1,\dots,d\}$, where

$$H_i(t) := \sum_{j=0}^{i-1} (-1)^j \binom{i-1}{j} A_{d-i+j}(t), \quad \forall t \geq 0.$$

Now, we obtain the claim as a corollary from [14, Thm. 1.1] by using [11, Lem. 2.5.2] which implies that

$$(-1)^{i-1}\Delta^{i-1}A_{d-i}(t) = \sum_{j=0}^{i-1} (-1)^j \binom{i-1}{j} A_{d-i+j}(t), \quad \forall t \geq 0.$$

□

4 The deFinetti representation of GMO sequences

In this section, we characterise the deFinetti representation of exchangeable generalised Marshall–Olkin sequences. We begin with an overview of general deFinetti representations.

We know from *deFinetti's theorem*, see [1, Thm. 3.1], that an almost surely unique random distribution function F exists for each exchangeable sequence $\tau_1, \tau_2, \ldots$ such that almost surely

$$\mathbb{P}\left(\tau_1 \le x_1, \ldots, \tau_d \le x_d \middle| F\right) = \prod_{i=1}^{d} F_{x_i}, \quad \forall x_1, \ldots, x_d \in \mathbb{R},\ d \in \mathbb{N}. \tag{6}$$

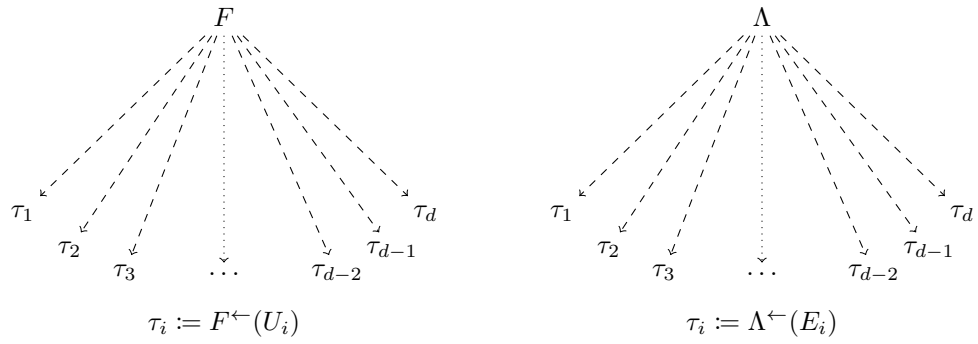
For a non-decreasing function h, we define its *generalised (right) inverse* $h^{\leftarrow}$ by $h^{\leftarrow}(y) := \inf\{x : h(x) \ge y\}$ with $\inf \emptyset = 0$, see [7] for a detailed discussion of generalised inverses. If the random distribution function F has almost surely no jumps, we have that almost surely

$$\tau_i = F^{\leftarrow}(U_i) = \inf\left\{x \in \mathbb{R} \ :\ F_x \ge U_i\right\}, \quad i \in \mathbb{N}, \tag{7}$$

for an iid uniform sequence $U_1, U_2, \ldots$, independent of F, which is defined by $U_i := F(\tau_i)$. If $\operatorname{supp}(F) \subseteq [0, \infty]$, we can rewrite Eq. (7) as

$$\tau_i = \Lambda^{\leftarrow}(E_i) = \inf\left\{t > 0 \ :\ \Lambda_t \ge E_i\right\}, \quad i \in \mathbb{N}, \tag{8}$$

for a (càdlàg) subordinator Λ and a sequence $E_1, E_2, \ldots$ of iid unit exponential random variables, independent of Λ. For this, we define $\Lambda = -\log(1 - F)$ and $E_i = -\log(1 - U_i)$, $i \in \{1, \ldots, d\}$. Note that we define a *subordinator* as a $[0, \infty]$-valued, non-decreasing, càdlàg process on $[0, \infty)$ that starts at zero and tends to infinity for $t \to \infty$. If the random distribution function F may possibly have jumps, then Eqs. (7) and (8) still hold if there is an additional iid uniform sequence $W_1, W_2, \ldots$, which is independent of $\tau_1, \tau_2, \ldots$, defined on the probability space. The sequence $W_1, W_2, \ldots$ is required to modify $F(\tau_i)$ to a uniform random variable by a random interpolation at its (random) atoms, see [23].



(a) With random distribution function **(b)** With random hazard rate

Figure 1: A visualisation of both deFinetti representations: (1) Draw F (resp. Λ) (2) For each component i, draw U_i (resp. E_i) and transform with generalised inverse of F (resp. Λ). We have $F = 1 - \exp\{-\Lambda\}$ and $U_i = 1 - \exp\{-E_i\}$.

Before moving on to the main result of this article, we want to outline three applications of the deFinetti representation:

1. We can use the deFinetti representation to sample from certain distributions efficiently in high-dimensions as illustrated in Fig. 1. See, e.g., [11, 16, 19] for applications of this technique.
2. We can build low-parametric, dimensionless families of multivariate distributions from parametrised subordinators, see, e.g., [4, 15, 17] for examples. We call these families *dimensionless*, since a random vector from such a model can be defined as the margin of an infinite sequence. Consequently, these families are not inherently linked to a specific dimension.

3. We can use the deFinetti representation for exchangeable sequences to build hierarchical models for non-exchangeable sequences. We refer the interested reader to [12, 18, 20].

Below, we state the main result of this article and investigate the subordinator, which is implied by the deFinetti representation of generalised Marshall–Olkin sequences. We already know from [11] that Marshall–Olkin sequences are uniquely linked to Lévy subordinators via Eq. (8). That remains true if we generalise the Marshall–Olkin definition as in Section 3 and generalise the Lévy subordinator to an additive subordinator.

Theorem 1 (Main result). *Let $\tau_1, \tau_2, \dots$ be an infinite exchangeable sequence of generalised Marshall–Olkin kind, i.e. a sequence of functions $A_0, A_1, \dots \in \mathcal{H}$ with $A_0 \in \mathcal{H}_0$ and $A_i(0) = 0$ exists such that for $d \geq 2$*

$$\mathbb{P}(\tau_1 > t_1, \dots, \tau_d > t_d) = \exp\left\{-\sum_{i=1}^{d} A_{i-1}\big(t_{[i]}\big)\right\}, \quad \forall \boldsymbol{t} \geq 0. \tag{9}$$

Furthermore, assume that an iid uniform sequence $W_1, W_2, \dots$ which is independent of $\tau_1, \tau_2, \dots$ is defined on the probability space. Then, an additive subordinator Λ and iid unit exponentially distributed random variables $E_1, E_2, \dots$, independent of Λ, exist such that almost surely

$$\tau_i = \inf\left\{t > 0 \,:\, \Lambda_t \geq E_i\right\}, \quad \forall i \in \mathbb{N}. \tag{10}$$

Conversely, if Λ is an additive subordinator with Laplace exponents $\{\psi_t\}_{t\geq 0}$, $E_1, E_2, \dots$ are iid unit exponentially distributed random variables, independent of Λ, and $\tau_1, \tau_2, \dots$ are constructed according to Eq. (10)*, then for all $d \geq 2$ the random vector $\boldsymbol{\tau}_d = (\tau_1, \dots, \tau_d)$ has an exchangeable generalised Marshall–Olkin distribution with*

$$A_i(t) = \psi_t(i+1) - \psi_t(i), \quad \forall t \geq 0,\ i \in \mathbb{N}_0. \tag{11}$$

Proof. Firstly, note that the backward direction is a corollary of [14, Prop. 3.1] by considering marginal transformations.

For the forward direction, which is the main contribution of this article, we use deFinetti's theorem, see [1, Thm. 3.1], to obtain the existence of a random distribution function F such that the sequence is conditionally iid given F and Eq. (6) holds. We define

$$U_i \coloneqq F_{\tau_i} + W_i\big(F_{\tau_i} - F_{\tau_i-}\big), \quad i \in \mathbb{N}.$$

We use [23, Sec. 2] to obtain that, conditioned on F, $U_1, U_2, \dots$ are uniform and fulfil almost surely Eq. (6). In particular, we have that almost surely

$$\mathbb{P}\big(U_1 \leq u_1, \dots, U_d \leq u_d \big| F\big) = \prod_{i=1}^{d} u_i, \quad \boldsymbol{u} \in [0,1]^d.$$

In summary, the sequence $U_1, U_2, \dots$ is iid uniform and independent of F. We use the transformations $\Lambda = -\log(1-F)$ and $E_i = -\log(1-U_i)$, $i \in \mathbb{N}$ and obtain a subordinator Λ and an iid unit exponential sequence $E_1, E_2, \dots$, independent of Λ, such that Eq. (10) holds almost surely.

Now, we have to prove that Λ is an additive subordinator. By a simple uniqueness-in-distribution argument and [25, p. 41], we determine that this is equivalent to the existence of a family of Bernstein functions $\{\psi_t\}_{t\geq 0} \subseteq \mathcal{BF}$, fulfilling the conditions

$$\psi_0 \equiv 0, \tag{12a}$$

$$\psi_s - \psi_t \in \mathcal{BF} \quad \forall s > t \geq 0, \tag{12b}$$

$$t \mapsto \psi_t(x) \in \mathcal{C}^{(0)}, \quad \forall x \geq 0, \tag{12c}$$

such that Eq. (11) holds. Below, we show that a family of Bernstein functions with these properties exists. With Lemma 1, we have

$$H_i^{(d)} \coloneqq (-1)^{i-1}\Delta^{i-1}A_{d-i} \in \mathcal{H}, \quad i \in \{1, \dots, d\}.$$

Fix $s > t \geq 0$. Then, we have for arbitrary $d \in \mathbb{N}$

$$\begin{aligned} & H_i^{(d)}(s) - H_i^{(d)}(t) \geq 0, \quad \forall i \in \{1, \dots, d\} \\ \Leftrightarrow \quad & (-1)^{i-1} \Delta^{i-1} \big(A_{d-i}(s) - A_{d-i}(t)\big) \geq 0, \quad \forall i \in \{1, \dots, d\}. \end{aligned}$$

This implies that the sequence $A_0(s) - A_0(t), A_1(s) - A_1(t), \dots$ is completely monotone, see [11, Lem. 3.3.2]. A completely monotone sequence can be uniquely represented by the series of first-order iterated differences of a Bernstein function on $\mathbb{N}_0$, see [11, Sec. 4.1], [8, Cor. 4.2], and [3, Prop. 6.12]. Therefore, we obtain the existence of a unique Bernstein function $\psi_{s,t}$ with $\psi_{s,t}(0) = 0$ such that

$$A_i(s) - A_i(t) = \psi_{s,t}(i+1) - \psi_{s,t}(i) \quad \forall i \in \mathbb{N}.$$

This implies

$$\begin{aligned} \psi_{s,t}(i+1) - \psi_{s,t}(i) &= A_i(s) - A_i(t) \\ &= \big[A_i(s) - A_i(0)\big] - \big[A_i(t) - A_i(0)\big] \\ &= \big[\psi_{s,0}(i+1) - \psi_{s,0}(i)\big] - \big[\psi_{t,0}(i+1) - \psi_{t,0}(i)\big], \quad \forall i \in \mathbb{N}. \end{aligned}$$

Thus, if we set $\psi_u = \psi_{u,0}$ for $u \in \{t, s\}$, we obtain $\psi_0 \equiv 0$ and

$$\psi_t(i) + \psi_{s,t}(i) = \psi_s(i), \quad \forall i \in \mathbb{N}.$$

We use the fact that Bernstein functions are determined by their values on $\mathbb{N}_0$, see [3, Prop. 6.12] and [26, Thm. 3.2], and we get $\psi_s - \psi_t = \psi_{s,t} \in \mathcal{BF}$. Finally, we use that a Bernstein function is non-negative and monotone increasing to obtain the following formula for $s > t \geq 0$ and $x \geq 0$ that

$$0 \leq \psi_s(x) - \psi_t(x) = \psi_{s,t}(x) \leq \psi_{s,t}(\lceil x \rceil) = \sum_{j=1}^{\lceil x \rceil} A_{j-1}(s) - A_{j-1}(t).$$

Hence, the continuity of $A_0, A_1, \dots$ implies $\lim_{t_k \to t} \psi_{t_k}(x) = \psi_t(x)$ for all $t, x \geq 0$. □

Recovery of the subordinator

Theorem 1 motivates the following questions: firstly, what are non-trivial examples of how the forward direction of this theorem can be used and secondly, how can we use the theorem to learn more about the implied subordinator. A non-trivial example is an exchangeable, but not comonotone or independent, generalised Marshall–Olkin sequence, which is not directly generated by a deFinetti model. Given such a sequence, the theorem only guarantees the existence of a deFinetti representation, but does not explicitly state the law of the subordinator or how it can be explicitly recovered. In the following, we use an example adapted from [13, Expl. 6.3] to demonstrate how the subordinator can be identified and recovered.

We consider an exogenous shock model in which each component can fail due to independent individual shocks or a common global shock. For this, let $H, H^G \in \mathcal{H}$ with $H + H^G \in \mathcal{H}_0$ be cumulative hazard rate functions and define $A = H$ and $A_0 = H + H^G$. Furthermore, let $Z_G \sim 1 - \exp\{-H^G\}$ and let $Z_1, Z_2, \dots$ be an iid sequence with distribution function $1 - \exp\{-H\}$ that is independent of Z_G. We define the random sequence $\tau_1, \tau_2, \dots$ by

$$\tau_i := \min\Big\{Z_i, Z_G\Big\}, \quad i \in \mathbb{N}.$$

Recovery of the subordinator law

In the first step, we use the generalised version of the exogenous shock model representation from Eq. (3) and the novel result from Theorem 1 to determine that the subordinator, implied by the deFinetti representation, is

an additive subordinator with cumulative killing hazard rate H^G and deterministic part H. Since the sequence $\tau_1, \tau_2, \dots$ is exchangeable and of generalised Marshall–Olkin kind, we know that the random vector $\boldsymbol{\tau} = (\tau_1, \dots, \tau_d)$ has an exchangeable generalised Marshall–Olkin distribution for each $d \in \mathbb{N}$. We use Lemma 1 and determine that the corresponding survival function is

$$\bar{F}(\boldsymbol{t}) = \exp\left\{-A_0\left(t_{[1]}\right) - \sum_{i=2}^{d} A\left(t_{[i]}\right)\right\}, \quad \boldsymbol{t} = (\tau_1, \dots, \tau_d) \geq 0.$$

Then, we conclude with Theorem 1 that an additive subordinator Λ with the characterising family of Bernstein functions $\{\psi_t\}_{t\geq 0}$ exists such that

$$A_i(t) = \psi_t(i+1) - \psi_t(i), \quad i \in \mathbb{N}_0 \quad \forall t \geq 0.$$

This implies for $t \geq 0$ that

$$\begin{aligned} A_0(t) &= \psi_t(1), \\ A(t) &= \psi_t(i+1) - \psi_t(i), \quad i \in \mathbb{N}, \\ \psi_t(i) &= \begin{cases} A_0(t) + (i-1)A(t) & i \in \mathbb{N}, \\ 0 & i = 0. \end{cases} \end{aligned}$$

As Bernstein functions are uniquely defined by their values on $\mathbb{N}_0$, we verify that

$$\psi_t(x) = \underbrace{(A_0(t) - A(t))}_{=H^G(t)} 1_{\{x>0\}} + x \underbrace{A(t)}_{=H(t)}, \quad t, x \geq 0.$$

This family of Bernstein functions can be identified with an additive subordinator with (inhomogeneous) cumulative killing hazard rate $H^G(t)$ and deterministic part $H(t)$. In particular, a random variable $Z \sim 1 - \exp\{-H^G\}$ exists such that

$$\Lambda_t = H(t) + \infty \cdot 1_{\{Z\leq t\}} = \begin{cases} H(t) & t < Z, \\ \infty & t \geq Z. \end{cases}$$

Note that so far, we only know that some random variable Z exists such that this equation holds. A natural conjecture is that $Z = Z^G$, which is proven in the following.

Explicit recovery of the subordinator

In the second step, to derive the subordinator explicitly, we use that

$$F_t(\omega) = \mathbb{E}\left[1_{\{\tau_1\leq t\}} \middle| \mathcal{T}\right](\omega), \quad \forall \omega \in \Omega \setminus N$$

for the tail-σ-algebra $\mathcal{T}$ of the sequence $\tau_1, \tau_2, \dots$ and a $\mathbb{P}$-nullset N, see [1, Lem. 2.15 and 2.19]. Furthermore, we use that $Z_j > t$ for infinitely many j and therefore

$$\left\{Z^G > t\right\} = \bigcap_{i\geq 1} \bigcup_{j\geq i} \left\{\min\{Z_j, Z_G\} > t\right\}, \quad t \geq 0.$$

Consequently, Z_G is measurable with respect to $\mathcal{T}$. Moreover, we have for $\omega \in \Omega\setminus N$

$$\begin{aligned} \Lambda_t(\omega) &= -\log\left(1 - F_t(\omega)\right) = -\log\left(\mathbb{E}\left[1_{\{\tau_1>t\}} \middle| \mathcal{T}\right](\omega)\right) \\ &= -\log\left(\mathbb{E}\left[1_{\{Z_1>t\}}\right] 1_{\{Z^G(\omega)>t\}}\right) \\ &= -\log\left(\mathbb{E}\left[1_{\{Z_1>t\}}\right]\right) + \infty 1_{\{Z^G(\omega)\leq t\}} \end{aligned}$$

$$= \begin{cases} H(t) & t < Z_G(\omega), \\ \infty & t \geq Z_G(\omega), \end{cases}$$

where we use the convention that $0 \cdot \infty = 0$. Finally, given an iid uniform sequence $W_1, W_2, \ldots$, independent of $\tau_1, \tau_2, \ldots$, we can construct the sequence $E_1, E_2, \ldots$ by

$$U_i := \begin{cases} \left(1 - \mathrm{e}^{-H(Z_i)}\right) & Z_i < Z_G, \\ 1 - (1 - W_i)\,\mathrm{e}^{-H(Z_G)} & Z_i \geq Z_G. \end{cases} \qquad i \in \mathbb{N}$$

and

$$E_i := -\log(1 - U_i) = \begin{cases} H(Z_i) & Z_i < Z_G \\ H(Z_G) - \log(1 - W_i) & Z_i \geq Z_G. \end{cases} \qquad i \in \mathbb{N}.$$

Now, Theorem 1 implies that the sequence $E_1, E_2, \ldots$ is iid unit exponential, independent of F, and we conclude that almost surely

$$\tau_i = \inf\left\{t > 0 \,:\, \Lambda_t \geq E_i\right\}, \quad i \in \mathbb{N}.$$

5 Conclusion

We have shown that exchangeable sequences $\tau_1, \tau_2, \ldots$ of a generalised Marshall–Olkin kind are uniquely linked to additive subordinators via a deFinetti representation. In particular, in a suitably extended probability space, we have almost surely that

$$\tau_i = \inf\{t > 0 \,:\, \Lambda_t \geq E_i\}, \quad i \in \mathbb{N},$$

where Λ is an additive subordinator and the sequence $E_1, E_2, \ldots$ is iid unit exponential and independent of Λ.

Acknowledgements: Many thanks to Lexuri Fernández, Florian Brück, Matthias Scherer, and the two anonymous reviewers for their feedback on this article.

A Exchangeable sequences and DeFinetti's theorem

In this section, we summarise the background on exchangeable sequences and deFinetti representations. An extensive reference on the deFinetti representation of exchangeable sequences and exchangeability in general, which contains all results that are presented in this section, is [1].

We call a sequence $\tau_1, \tau_2, \ldots$ *exchangeable* if

$$(\tau_1, \ldots, \tau_d) \overset{d}{=} (\tau_{\pi(1)}, \ldots, \tau_{\pi(d)}),$$

for each $d \in \mathbb{N}$ and permutation π on $\{1, \ldots, d\}$. A well-known result, first established by Bruno deFinetti in [6], states that sequences $\tau_1, \tau_2, \ldots$ are exchangeable if, and only if, they are conditionally iid. While this statement is clear and simple, there are some technical details hidden in the expression *conditionally iid*. In our case, since generalised Marshall–Olkin distributions have singular components and additive subordinators have jumps, these details become very important. This is explained in more detail with an example at the end of this section. For this reason, we outline below how an exchangeable sequence can be represented by a random distribution function and an iid uniform sequence.

DeFinetti Representation (See [1, Thm. 3.1]). *A sequence $\tau_1, \tau_2, \ldots$ is exchangeable if, and only if, a random measure α exists such that the product measure α^∞ is a regular conditional distribution of $\tau_1, \tau_2, \ldots$ given $\sigma(\alpha)$.*

In the following, we show how the *directing measure* α can be calculated from the sequence $\tau_1, \tau_2, \ldots$. For this, assume that the sequence $\tau_1, \tau_2, \ldots$ is defined in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{T}$ be its tail-σ-algebra. On existence, α is a.s. unique, $\mathcal{T}$ measurable, and a regular conditional distribution for τ_1 given $\mathcal{T}$, see [1, Lem. 2.15 and 2.19]. Thus, we have

$$\alpha(\omega, A) = \mathbb{P}\big(\tau_1 \in A | \mathcal{T}\big)(\omega), \quad \forall \omega \in \Omega \setminus N,\ A \in \mathcal{B},$$

where N is a $\mathbb{P}$-nullset. In the following, we assume w.l.o.g. that $\alpha(\omega, A) = 0$ for all $\omega \in N$, $A \in \mathcal{B}$. Finally, since $\alpha(\omega, \cdot)$ is a (random) probability measure on $\mathbb{R}$, we may identify $\alpha(\omega, \cdot)$ with a random distribution function $F(\omega)$ via

$$F_t(\omega) := \alpha(\omega, (-\infty, t]), \quad \forall t \in \mathbb{R},\ \omega \in \Omega.$$

If another sequence of iid uniform random variables $W_1, W_2, \ldots$, which is independent of $\tau_1, \tau_2, \ldots$ is defined on the probability space, we can refine deFinetti's theorem:

Corollary 1 (Cf. [1, Thm. 3.1])**.** *Let $W_1, W_2, \ldots$ be an iid uniform sequence and let $\tau_1, \tau_2, \ldots$ be independent thereof. The sequence $\tau_1, \tau_2, \ldots$ is exchangeable if, and only if, a random distribution function F and an iid uniform sequence $U_1, U_2, \ldots$, independent of F, exist such that*

$$\tau_i = \inf\left\{t \in \mathbb{R} \ :\ F_t \geq U_i\right\} \text{ a.s.}, \quad \forall i \in \mathbb{N}. \tag{13}$$

Corollary 2 (Cf. [1, Thm. 3.1])**.** *Let $W_1, W_2, \ldots$ be an iid uniform sequence and let $\tau_1, \tau_2, \ldots \geq 0$ be independent thereof. The sequence $\tau_1, \tau_2, \ldots$ is exchangeable if, and only if, a random subordinator Λ and an iid unit exponential sequence $E_1, E_2, \ldots$, independent of Λ, exist such that*

$$\tau_i = \inf\left\{t \geq 0 \ :\ \Lambda_t \geq E_i\right\} \text{ a.s.}, \quad \forall i \in \mathbb{N}.$$

Proof of Corollaries 1 and 2. Firstly, the claim from Corollary 2 follows directly from Corollary 1 with the transformations $\Lambda = -\log(1 - F)$ and $E_i = -\log(1 - U_i)$, $i \in \mathbb{N}$. Secondly, we use [23, Prop. 2.1] to ascertain that $U_1, U_2, \ldots$ are iid uniform conditioned on $\mathcal{T}$ and that Eq. (13) holds, where we define

$$U_i := F_{\tau_i -} + W_i \left(F_{\tau_i} - F_{\tau_i -}\right), \quad \forall i \in \mathbb{N}.$$

Finally, with the definition of the regular conditional distribution, we establish that $U_1, U_2, \ldots$ is an iid uniform sequence that is independent of $\mathcal{T}$, hence also independent of F. □

We conclude this section with an example that explains the need for additional randomness, in form of an iid uniform sequence $W_1, W_2, \ldots$, in the two preceding theorems. This example also highlights that not every *conditionally independent* sequence has a representation as in Eq. (13) when only the original probability space is considered. For this, let $(\Omega, \mathcal{F}, \mathbb{P})$ be the Lebesgue probability space on the interval $[0, 1]$ and define

$$U_i(\omega) := \omega, \quad \omega \in [0, 1],\ i \in \mathbb{N}.$$

Clearly, the sequence $U_1, U_2, \ldots$ is exchangeable and U_1 is measurable with respect to the sequences tail-σ-algebra $\mathcal{T}$. Hence, we can calculate the random distribution function F, corresponding to the sequences directing measure α, for all ω excluding a Lebesgue-nullset and $u \in [0, 1]$ by

$$F_u(\omega) = \mathbb{E}\left[1_{\{U_1 \in [0,u]\}} \middle| \mathcal{T}\right](\omega) = 1_{\{U_1(\omega) \in [0,u]\}}.$$

Since $\sigma(F) = \sigma(U_1) = \mathcal{F}$, there is no additional iid sequence independent of F defined on this probability space. If we now consider the enclosing probability product space, on which U_1 as well as an iid uniform sequence $W_1, W_2, \ldots$, independent of U_1, are defined, we have

$$\begin{aligned} U_i = U_1 &= \inf\left\{u \in [0, 1] \ :\ 1_{\{U_1 \in [0,u]\}} \geq W_i\right\} \\ &= \inf\left\{u \in [0, 1] \ :\ F_u \geq W_i\right\}, \quad i \in \mathbb{N}. \end{aligned}$$

B Bernstein functions and completely monotone sequences

The proof of the main theorem relies heavily on the connection between additive and Lévy subordinators, so-called *Bernstein functions*, and *completely monotone sequences*. As the topic cannot be treated in detail without using deep results of functional analysis and measure theory, we will limit ourselves to presenting the main results. Extensive references on this topic are [3, 26]. Another excellent reference is [11, Chp. 3 and 4].

A *Bernstein function* is a function $\psi : (0, \infty) \to [0, \infty)$ that has derivatives of arbitrary order $\psi^{(i)}$, $i \in \mathbb{N}$, and fulfils

$$(-1)^{i-1}\psi^{(i)}(x) \geq 0, \quad \forall x > 0,\ i \in \mathbb{N}.$$

We denote the set of all Bernstein functions by $\mathcal{BF}$ and use the convention that a Bernstein function may be extended to $[0, \infty)$ by setting $\psi(0) := 0$. It is well-known, see, e.g. [26, Thm. 3.2], that a function $\psi : (0, \infty) \to [0, \infty)$ is a Bernstein function if, and only if, real numbers $a, b \geq 0$ and a Lévy-measure ν on $(0, \infty)$ exist such that

$$\psi(x) = a + bx + \int_{(0,\infty)} \left(1 - \mathrm{e}^{-xu}\right) \nu(\mathrm{d}u), \quad \forall x > 0,$$

where we call a measure ν on $(0, \infty)$ a *Lévy measure* if $\int_{(0,\infty)}(1 \wedge x)\, \nu(\mathrm{d}x) < \infty$. In that case (a, b, ν) is uniquely determined by ψ and is called the *Lévy triplet*.

Bernstein functions ψ with $\psi(0) = 0$ can be uniquely linked to so-called completely monotone sequences. For a (countably infinite) sequence $a_0, a_1, \ldots$, let Δ be the discrete difference operator defined by $\Delta a_i := a_{i+1} - a_i$ and define recursively $\Delta^n a_i := \Delta(\Delta^{n-1} a_i)$. We call the sequence $a_0, a_1, \ldots$ *completely monotone* if

$$(-1)^i \Delta^i a_k \geq 0, \quad \forall i, k \in \mathbb{N}_0.$$

We call a finite sequence $a_0, \ldots, a_{d-1}$ *d-monotone* if

$$(-1)^i \Delta^i a_k \geq 0, \quad \forall i, k \in \mathbb{N}_0 \ :\ i + k < d. \tag{14}$$

In particular, a sequence $a_0, a_1, \ldots$ is completely monotone if, and only if, the sequences $a_0, a_1, \ldots, a_{d-1}$ are d-monotone for all $d \in \mathbb{N}$. Furthermore, a sequence $a_0, a_1, \ldots, a_{d-1}$ is d-monotone if, and only if, Eq. (14) is fulfilled for $i, k \in \mathbb{N}_0$ with $i + k = d - 1$, see [11, Lem. 3.3.2]. Moreover, a sequence $a_0, a_1, \ldots$ is completely monotone if, and only if, a Bernstein function ψ exists with $a_i = \psi(i + 1) - \psi(i)$ for all $i \in \mathbb{N}_0$, see [11, Sec. 4.1] and cf. [8, Cor. 4.2] or [3, Prop. 6.12]. Note, that this implies that Bernstein function are uniquely determined by their values on the natural numbers.

A Bernstein function, and subsequently a completely monotone sequence, can be uniquely linked to the law of a Lévy subordinator, see [5, Thm. 1.2]. In particular, let ψ be a Bernstein function, then a Lévy subordinator Λ, uniquely determined in law, exists with Laplace exponent $x \mapsto t\psi(x)$ for all $t \geq 0$, i.e.

$$\mathbb{E}\left[\mathrm{e}^{-x\Lambda_t}\right] = \mathrm{e}^{-t\psi(x)}, \quad \forall t, x \geq 0. \tag{15}$$

Conversely, if Λ is a Lévy subordinator, then a Bernstein function ψ exists such that Eq. (15) holds.

This can be generalised, see, e.g. [25, p. 41]: let $\{\psi_t\}_{t\geq 0}$ be a family of Bernstein functions fulfilling

$$\psi_0 \equiv 0, \tag{12a rev.}$$

$$\psi_s - \psi_t \in \mathcal{BF} \quad \forall s > t \geq 0, \tag{12b rev.}$$

$$t \mapsto \psi_t(x) \in \mathcal{C}^{(0)}, \quad \forall x \geq 0. \tag{12c rev.}$$

Then, an additive subordinator, uniquely determined in law, exists with Laplace exponent $x \mapsto \psi_t(x)$ for all $t \geq 0$ and

$$\mathbb{E}\left[\mathrm{e}^{-x(\Lambda_s - \Lambda_t)}\right] = \mathrm{e}^{-(\psi_s - \psi_t)(x)}, \quad \forall s \geq t, x \geq 0. \tag{16}$$

Conversely, if Λ is an additive subordinator, then a family of Bernstein functions $\{\psi_t\}_{t\geq 0}$ exists fulfilling Eqs. (12a) to (12c) such that Eq. (16) holds.

References

[1] Aldous, D. J. (1985). *Exchangeability and Related Topics*. Springer, Berlin.

[2] Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus*. Second edition. Cambridge University Press.

[3] Berg, C., J. P. R. Christensen, and P. Ressel (1984). *Harmonic Analysis on Semigroups*, Springer, New York.

[4] Bernhart, G., J.-F. Mai, and M. Scherer (2015). On the construction of low-parametric families of min-stable multivariate exponential distributions in large dimensions. *Depend. Model. 3*, 29–46.

[5] Bertoin, J. (1999). *Subordinators: Examples and Applications*. Springer, Berlin.

[6] de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré 7*, 1–68.

[7] Embrechts, P. and M. Hofert (2013). A note on generalized inverses. *Math. Methods Oper. Res. 77*(3), 423–432.

[8] Gnedin, A. and J. Pitman (2008). Moments of convex distribution functions and completely alternating sequences. In D. Nolan and T. Speed (Eds.), *Probability and statistics: essays in honor of David A. Freedman.*, pp. 30–41, Inst. Math. Statist., Beachwood OH.

[9] Li, X. and F. Pellerey (2011). Generalized Marshall–Olkin distributions and related bivariate aging properties. *J. Multivariate Anal. 102*(10), 1399–1409.

[10] Lin, J. and X. Li (2014). Multivariate generalized Marshall–Olkin distributions and copulas. *Methodol. Comput. Appl. Probab. 16*(1), 53–78.

[11] Mai, J.-F. (2010). *Extendibility of Marshall–Olkin Distributions via Lévy Subordinators and an Application to Portfolio Credit Risk*. PhD thesis, Technische Universität München, Germany.

[12] Mai, J.-F. (2014). *Multivariate Exponential Distributions with Latent Factor Structure and Related Topics*. Habilitation thesis, Technische Universität München, Germany.

[13] Mai, J.-F. (2019). The infinite extendibility problem for exchangeable real-valued random vectors. Available at: https://arxiv.org/abs/1907.04054.

[14] Mai, J.-F., S. Schenk, and M. Scherer (2016). Exchangeable exogenous shock models. *Bernoulli 22*(2), 1278–1299.

[15] Mai, J.-F., S. Schenk, and M. Scherer (2017). Two novel characterizations of self-decomposability on the half-line. *J. Theoret. Probab. 30*(1), 365–383.

[16] Mai, J.-F. and M. Scherer (2009a). Efficiently sampling exchangeable Cuadras–Augé copulas in high dimensions. *Inform. Sci. 179*(17), 2872–2877.

[17] Mai, J.-F. and M. Scherer (2009b). Lévy-frailty copulas. *J. Multivariate Anal. 100*(7), 1567–1585.

[18] Mai, J.-F. and M. Scherer (2013). Sampling exchangeable and hierarchical Marshall-Olkin distributions. *Comm. Statist. Theory Methods 42*(4), 619–632.

[19] Mai, J.-F. and M. Scherer (2014). Characterization of extendible distributions with exponential minima via processes that are infinitely divisible with respect to time. *Extremes 17*(1), 77–95.

[20] Mai, J.-F. and M. Scherer (2017). *Simulating Copulas: Stochastic Models, Sampling Algorithms and Applications*. Second edition. World Scientific.

[21] Marshall, A. W. and I. Olkin (1967). A multivariate exponential distribution. *J. Amer. Statist. Assoc. 62*(317), 30–44.

[22] Muliere, P. and M. Scarsini (1987). Characterization of a Marshall–Olkin type class of distributions. *Ann. Inst. Statist. Math. 39*(2), 429–441.

[23] Rüschendorf, L. (2009). On the distributional transform, Sklar's theorem, and the empirical copula process. *J. Statist. Plann. Inference 139*(11), 3921–3927.

[24] Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

[25] Schenk, S. (2016). *Exchangeable Exogenous Shock Models*. PhD thesis, Technische Universität München, Germany.

[26] Schilling, R. L., R. Song, and Z. Vondracek (2012). *Bernstein Functions. Theory and Applications*. Second edition. De Gruyter, Berlin.

C Further articles as co-author

C.1 Consistent Iterated Simulation of Multivariate Defaults: Markov Indicators, Lack of Memory, Extreme-Value Copulas, and the Marshall–Olkin Distribution

Consistent Iterated Simulation of Multivariate Defaults: Markov Indicators, Lack of Memory, Extreme-Value Copulas, and the Marshall–Olkin Distribution

Damiano Brigo, Jan-Frederik Mai, Matthias Scherer, and Henrik Sloot

---

The review article [5] critically examines the practice of iteratively simulating multivariate credit events over a discretized time grid. Primarily, we outline the effect of falsely assuming a multivariate lack of memory property, carrying inter-period to terminal default time dependencies. Secondarily, we propose using distributions fulfilling a suitable lack of memory property. In particular, we promote using hierarchical Marshall–Olkin factor models to address a group of modeling issues for credit portfolios.

The introduction summarizes practitioners' requirements for a joint model for financial risk factors and credit events: Asset and risk management of path-dependent products often involves the simulation of financial risk factors, for example, equity returns, over a discretized time grid. In addition, measuring counterparty credit risk or simulating credit products must include credit events and default probabilities in these simulations. We illustrate why many practitioners choose increment-driven credit models over directly simulating default times despite the complexities of their terminal default time probabilities. Reasons are, for example, their consistency with existing increment-driven models for other risk factors and the natural consideration of different risk horizons.

A common simplification assumes iid inter-period credit events, discarding already defaulted entities. The idea for this article originated from the observation of D. Brigo that some practitioners used this iid assumption while falsely presuming equality, or at least similarity, of the terminal and inter-period default time dependencies.

Section 2 outlines why assuming the default-indicator process being embeddable in a continuous-time, (homogeneous) Markov process is reasonable if one desires a feasible and consistent model. Following, we recall the existing *lack of memory (LOM)* properties for random vectors: The *full* lack of memory property, equivalent to independence; the *Marshall–Olkin (MO)* lack of memory property, characterizing the multivariate exponential distribution of the same name and equivalent to Markovian survival indicator processes; the *min-stable multivariate exponential (MSMVE)* lack of memory property, equivalent to an *extreme-value survival copula* coupled with exponential margins; and the *exponential minima (EM)* property, which assumes all minima of components are exponential. All of these properties require *margin stability* in the sense that they have to be fulfilled not only for the random vector itself but also for all of its margins. Subsequently, we coin *weak* versions of these properties, which do not require this margin stability.

Section 3 investigates *survival-of-all* events. These events are interesting as they are relevant for the valuation of *first-to-default* swaps. Additionally, we can reduce the requirements to obtain consistent inter-period and terminal distributions to the following discretized version of the *weak exponential minima* property:

$$\mathbb{P}(\boldsymbol{\tau} > n\Delta) = \mathbb{P}(\boldsymbol{\tau} > \Delta)^n. \tag{C.1}$$

Inspired by previous work, we label distributions as *common Δ-period self-chaining* if they fulfill Eq. (C.1) for all $n \in \mathbb{N}$ and as *self-chaining* if they are *common Δ-period self-chaining* for all $\Delta > 0$.

We link this property to the previously established lack of memory property by noting that the latter property is equivalent to the WEM property and provide two examples of WEM distributions. Consequently, we identify that EM distributions, including MSMVE and MO distributions, allow consistent iterated simulation of *survival of all* events. Additionally, we recall that the Clayton, Frank, and Gaussian copula are in the *maximum domain of attraction* of the independence copula. Consequently, using them as survival copulas in the simplified approach above leads to asymptotically vanishing dependence.

In section 4, we finally investigate arbitrary *mixed default-survival* events. We identify that *multivariate phase-type distributions*, characterized by having Markovian default indicator processes, fulfill our desired consistency requirement. As examples, we discuss the *looping default* model, an extension of the bivariate Freund distribution, and the *Marshall–Olkin* distribution. However, we argue that a feasible model should

allow dropping or adding entities without losing the consistency property or introducing model ambiguity. For this, we promote using previous hierarchical Marshall–Olkin factor models, rooted in the Lévy-frailty representation of extendible MO distributions, and provide ample background information.

To demonstrate issues of vanishing terminal dependence for some instances of the simplified approach above with selected survival copulas, we included two case studies for *survival-of-all* and *mixed-default-survival* events in Sections 3 and 4. The first case study considers survival copulas of type normal, Frank, Clayton, Gumbel, and t. The second case study also considers Marshall–Olkin, Freund, and independence distributions. The results of the case studies are in line with the previously presented theoretical results: only the Gumbel copula (*survival of all*) resp. Marshall–Olkin and independent (*mixed default-survival*) produce consistent inter-period and terminal probabilities. We created these case studies to warn practitioners against choosing an inter-period dependence, particularly the normal copula, resulting in asymptotically vanishing terminal dependence.

Statement of individual contribution

An initial draft of the article was written by D. Brigo, J.-F. Mai, and M. Scherer in a collaborative effort for the publication in the proceedings of the conference “Innovations in Insurance, Risk- and Asset Management” at the Technical University of Munich, April 5–7, 2017. I was invited to contribute after pointing out possible improvements in the preceding conference talk, associated with the working paper [88], on this topic; the talk coined the properties *common period lack of memory* and *self-chaining*, but wrongly equated random vectors being *self-chaining with exponential margins* and their survival copula having the *extreme-value property*. First, I significantly rewrote the paper, except for the introduction, without changing its general structure and synopsis. In particular, I aligned it with existing notions of lack of memory properties, such as *min-stable multivariate exponential* and *exponential minima* and weak versions thereof, which do not require that all margins inherit the property. Second, I improved the second case study in Sec. 4.3 by using exact probabilities instead of Monte-Carlo estimates and considering additional copulas, specifically, the *Clayton*, *Frank*, *independent*, *t*, and *Freund* copulas. Most of the paper cannot be attributed unanimously to a single author; however, the following parts were contributed solely by myself: Examples 3.1 and 3.2, Remarks 3.1 and 4.3, and a significant portion of Section 2. Additionally, I was responsible for performing the case studies, creating the visualizations, and writing the final draft; my co-authors contributed further by reviewing and commenting on the final version before submission and during the peer-review process.

Addendum

In a numerical study, Sec. 4 of the article demonstrates that the terminal probabilities of a *mixed default-survival* event coincide with those calculated from the inter-period distribution for the considered Marshall–Olkin (MO) distribution. However, the article misses to investigate whether this is a general property of all MO distributions for arbitrary mixed default-survival events. That this is true is shown in the following.[1]

Theorem C.1. *Consider a d-variate MO distributed random vector $\boldsymbol{\tau}$, and let the survival indicator chain $\tilde{Z}^{(\Delta)}$ for the discretization $\Delta > 0$ and final time horizon $T = n\Delta$ be the Markov chain defined by multiplicatively accumulated survival indicators until the time Δ of iid copies from $\boldsymbol{\tau}$ as described in Sec. 4 of the article. Then, $\boldsymbol{\tau}$ is also the terminal distribution of this survival indicator process, particularly,*

$$\Big\{\tilde{Z}^{(\Delta)}(k) \;:\; k \in \{0,\dots,T/\Delta\}\Big\} \overset{d}{=} \Big\{(1_{\{\tau_i > k\Delta\}} \;:\; i \in [d]) \;:\; k \in \{0,\dots,T/\Delta\}\Big\}.$$

Proof. Consider the *Arnold model* proposed in [20], represented as in the contributed core article [2]. In particular, for the *shock-arrival intensities* $\{\lambda_I : \emptyset \neq I \subseteq [d]\}$ of $\boldsymbol{\tau}$, define $\lambda := \sum_{\emptyset \neq I \subseteq [d]} \lambda_I$ and $p_I := \lambda_I/\lambda$, $\emptyset \neq I \subseteq [d]$, and let the survival-indicator process Z be defined via

$$Z_t = \prod_{j=1}^{N_t} Y_j, \quad t \geq 0,$$

[1] The theorem uses the notation of the contributed article and the references of this thesis.

where N is a Poisson process with intensity λ, and $Y_1, Y_2, \ldots$ is an iid sequence of multivariate binary random variables with $\mathbb{P}(Y_1 = (1_{\{i \notin I\}} : i \in [d])) = p_I$, $\emptyset \neq I \subseteq [d]$.

Now, the claim follows from

$$Z_t = \prod_{k=1}^{t/\Delta} \underbrace{\prod_{j=N_{(k-1)\Delta}+1}^{N_{k\Delta}} Y_j}_{\stackrel{d}{=} Z_\Delta} \stackrel{d}{=} \prod_{k=1}^{t/\Delta} Z_\Delta^{(k)}, \quad \forall t \in \{0, \Delta, 2\Delta, \ldots, T\},$$

where $Z^{(k)}$ are iid copies of Z; which uses that increments of N with the same time-step are iid. □

Permission to include

D. Brigo et al. “Consistent iterated simulation of multivariate defaults: Markov indicators, lack of memory, extreme-value copulas, and the Marshall–Olkin distribution”. In: *Innovations in Insurance, Risk- and Asset Management*. Ed. by K. Glau et al. Conference at the Technical University of Munich, 5–7 April 2017. Singapore: World Scientific, 2018, pp. 47–93. DOI: 10.1142/9789813272569_0003

Reprinted according to the *CC BY-NC-ND 4.0 license*;[2] see human-readable summary included below.

This page is available in the following languages:

Creative Commons License Deed

Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)

This is a human-readable summary of (and not a substitute for) the license.

You are free to:

Share — copy and redistribute the material in any medium or format

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial — You may not use the material for commercial purposes.

NoDerivatives — If you remix, transform, or build upon the material, you may not distribute the modified material.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

[2] https://creativecommons.org/licenses/by-nc-nd/4.0/.

Chapter 3

Consistent Iterated Simulation of Multivariate Defaults: Markov Indicators, Lack of Memory, Extreme-Value Copulas, and the Marshall–Olkin Distribution

Damiano Brigo
Department of Mathematics, Imperial College London,
London, SW7 2AZ, UK
damiano.brigo@imperial.ac.uk

Jan-Frederik Mai
Xaia Investment,
Sonnenstr. 19, 80331 München, Germany
jan-frederik.mai@xaia.com

Matthias Scherer
Technische Universität München, Lehrstuhl für Finanzmathematik
Parkring 11, 85748 Garching, Germany
scherer@tum.de

Henrik Sloot
Technische Universität München, Lehrstuhl für Finanzmathematik
Parkring 11, 85748 Garching, Germany
henrik.sloot@tum.de

A current market-practice to incorporate multivariate defaults in global risk-factor simulations is the iteration of (multiplicative) i.i.d. survival indicator increments along a given time-grid, where the indicator distribution is based on a copula ansatz. The underlying assumption is that the behavior of the resulting iterated default distribution is similar to the one-shot distribution. It is shown that in most cases this assumption is not fulfilled and furthermore numerical analysis is presented that shows sizable differences in probabilities assigned to both "survival-of-all" and "mixed default/survival" events. Moreover, the classes of distributions for which probabilities from the "terminal one-shot" and "terminal iterated" distribution coincide are derived for problems considering "survival-of-all" events as well as "mixed default/survival" events. For the former problem, distributions must fulfill a lack-of-memory type property, which is, e.g., fulfilled by min-stable multivariate exponential distributions. These correspond in a copula-framework to exponential margins coupled via extreme-value copulas. For the latter problem, while looping default inspired multivariate Freund distributions and more generally multivariate phase-type

distributions could be a solution, under practically relevant and reasonable additional assumptions on portfolio rebalancing and nested distributions, the unique solution is the Marshall–Olkin class.

Keywords: stepwise default simulation, default dependence, extreme-value copulas, Marshall–Olkin distribution, nested margining, Freund distribution, looping default models, multivariate phase-type distribution.

1. Introduction

The increasingly global nature of financial products and risks calls for adequately complex stochastic models and simulation procedures. These are required for valuation purposes as well as for risk analysis and often involve thousands of risk factors that can be different in nature. Investment banks and financial service companies are devoting a sizable effort to design software and hardware architectures that support such global simulations effectively, see, e.g. [1]. The path-dependent nature of many risks and the necessity to analyze risks at different time horizons lead to an iterated simulation of all risk factors across time steps. The consistent statistical representation of default-times of multiple entities and their inter-dependence-structure is the main motivation for this paper. For the simulation of default-times, up to a final horizon, two possible approaches are considered:

(i) Simulate the default-times, at the beginning, once and for all in each given scenario. The resulting values are stored and the other risk factors are simulated iteratively up to the final time horizon.
(ii) Alternatively, one simulates in each given scenario for every time-period a "default/no default" indicator of all non-defaulted entities conditional on the default history — i.e. the survival of non-defaulted entities up to the beginning of this period and the default-times of already defaulted entities.

We anticipate that we will be concerned with the consistency of the two approaches above under a number of additional specifications. The basic question is:

When is an iterated default simulation, often done by sampling a given type of multivariate distribution, equivalent to a one-shot simulation under essentially the same distribution?

Although this appears to be a simple question, it is in fact rather nuanced. For this question to fully make sense we need to be a little more precise on

our definitions and on our problem specification, and it is indeed one of the main purposes of this paper to fully clarify this question, its implications, and some possible answers. It is worth putting this pre-question in the open now, and we would like to mention that the first named author has witnessed cases in the industry where the two procedures were assumed to be equivalent when they were not, and this both in the valuation/hedging space and in the risk measurement space. While the author is not allowed to provide details on such cases for confidentiality issues, we will see some numerical examples clarifying this discrepancy in the course of the paper.

Going back to our introduction, the dependence between default-times and other risk factors has to be introduced on the whole risk factor evolution in approach (i) and on the period steps in approach (ii), respectively. In this formulation both approaches are mathematically equivalent — however, this equivalence is based on conditional probabilities, which can be arbitrarily complex.[1] Consider, for example, the case of wrong way risk for credit valuation adjustments for credit default swap (CDS) trades under collateralization in [2], where the first approach is used: even with just three default times involved, the CDS and the two trading parties, the formulas become very involved and cumbersome. Thus, generally, one either has a model for the default-times in approach (i) with complex conditional probabilities, or one has a model for the indicator increment process in approach (ii) with unknown "terminal iterated" dependence. The mathematical underpinning — if any — for company-wide, global simulation of defaults is often, or can be translated into, a copula-based ansatz. Such a model originates from the statistical literature and renders approach (i) more natural from the company-default perspective. However, when dealing with large portfolios, the literature on financial risk management mostly prefers models relying on a repeated evolution of risk factors on common time grids. Approach (ii) is more consistent with this way of thinking and therefore more desirable both from a theoretical and practical point of view, for the following reasons:

- **Software consistency with "Brownian-driven" asset classes**: Consider a bank that runs a global simulation on a large portfolio, including complex products and defaults, in order to obtain a risk measure.

[1]Contrary to the univariate case, where sampling from conditional probability distributions can be handled using the distributional transform, even if we can calculate the probabilities, conditional multivariate probability distributions can be very difficult to sample from.

One example would be computing the value-at-risk or the expected shortfall of CVA, a task that is numerically very intensive, see, e.g. [3]. In this context, there is need to evolve risk factors according to controlled time steps that are common to all factors, to have all required variables at each step of the simulation. While this is relatively natural for asset models that are driven by Brownian-type processes and even extensions with jumps, it becomes harder when trying to include defaults of underlying entities or counterparties. The reason for this is that default-times, typically represented through intensity models, should be simulated just once, being static random variables as opposed to stochastic processes. Once simulated, there would be nothing left to iterate. However, the consistency of the global simulation and the desire to have all variables simulated at every step is prompting the design of iterated survival or default flags across the time steps that are already used in the simulation of more traditional assets.

- **Basel III requirement for risk horizons:** A further motivation for iterating the global simulation across standard time steps is coming from the Basel III framework when trying to address liquidity risk. The Bank of International Settlements (BIS) suggests the following solution, see [4].
  "The Committee has agreed that the differentiation of market liquidity across the trading book will be based on the concept of liquidity horizons. It proposes that banks' trading book exposures be assigned to a small number of liquidity horizon categories: [10 days, 1 month, 3 months, 6 months, 1 year]. The shortest liquidity horizon (most liquid exposures) is in line with the current 10-day VaR treatment in the trading book. The longest liquidity horizon (least liquid exposures) matches the banking book horizon at one year. The Committee believes that such a framework will deliver a more graduated treatment of risks across the balance sheet. Among other benefits, this should also serve to reduce arbitrage opportunities between the banking and trading books."
  It is clear then that a bank will need to simulate the risk factors of the portfolio across a grid including the standardized holding periods above. In this sense it will be practical to simulate all variables, including defaults and survivals, in common time steps. Software architecture and the possibility to effectively decompose the simulation across steps, prompt to the possibility to iterate the default simulation rather than trying to simulate random default-times just once.
- **General need for dependence modeling in the context of the current counterparty credit risk debate:** As an example, the current

debate on valuation adjustments (as the partly overlapping credit CVA, debit DVA, and funding FVA adjustments, see, e.g., [3]), is forcing financial institutions to run global simulations over very large portfolios. By nature, CVA is an option on a very large portfolio containing the most disparate risk factors. A key quantity in valuing this option is the dependence between the default of a counterparty and the value of the underlying portfolio that is traded with that counterparty. When such dependence is adverse for the agent making the calculation we have wrong way risk (WWR), a risk that is at the center of the agenda of the Bank of International Settlements in reforming current regulation. Modeling the dynamics of dependence is not only essential for the current emergencies of the industry, such as CVA/DVA/FVA and risk measures on these quantities, but it is also necessary for the management of pure credit products, such as, e.g., Collateralized Debt or Loan Obligations (CDO, CLO).

Before shifting the focus solely to default-times, it is important to consider not only the distribution of default-times but also the dependence on other risk factors:

(a) In reality, default risk is correlated with other risk-factors. These can be risk-factors belonging to other asset classes, e.g. equity, or even macro-economic risk factors. These dependencies, however, are usually not considered in model building for the following reasons: It might be easy to reject the independence-assumption between a default-time and some other risk-factor with qualitative arguments or statistical tests, but the determination of a good model for this dependence (or directly for the joint distribution) is usually far from trivial. Even if one can formulate a satisfying model for other risk-factors and default-times — or the survival-indicator increments — the additional complexity can lead to computational problems (as explained in the following). Furthermore, the design of such a global model, including dependence between risk-factor classes, would require different departments of the financial institution to work together. For most institutions this is infeasible as business is often separated into different sections, of which each models their relevant risk factors to their own appropriate level of complexity.
(b) The computation of transition probabilities, or sampling from these transition-distributions, for the risk-factor evolution will be very

difficult and non-trivial in most cases. In particular, if there are no closed-form expressions, one usually has to rely on numerical-integration techniques — if available — which becomes time-consuming and is difficult to implement.

(c) Dependency information requires additional storage — especially if the dependency is conditional on the full histories of risk-factors, which is even challenging in low, but especially in high dimensions.

Focussing on the (discrete) survival-indicator process, there are more problems which have to be considered:

(d) Assume that there are d entities and N simulation steps up to the final time-horizon. In the worst case of full default-evolution-dependence this leads to $\sum_{k=0}^{N-1}(k+2)^d$ transition-probabilities. In the case of simple time-dependence, we have $N \cdot 3^d$ transition-probabilities. In the case of complete time-homogeneity, one "only" has 3^d transition-probabilities. For a large number of entities d or/and a large number of simulation steps N the issue of over-parameterization becomes apparent.

(e) Let T be the final time-horizon. Then the number of time-steps, and subsequently the number of parameters, depends on the step-size Δ, i.e. $N = T/\Delta$. This can lead to problems if different step-sizes have to be simulated (e.g. days, weeks, months, ...) as all probabilities should be consistent.

An additional problem is that the definition of all transition-probabilities have to be re-assessed in case the composition of the defaultable portfolio changes.

In summary, approach (i) appears more natural from the perspective of default modeling itself, however, in a global risk factor model, approach (ii) might be more desirable and is mostly used in the financial industry. Summing up, this involves the following questions:

(1) What are convenient conditions on the multivariate distribution of the default-times such that approach (i) and approach (ii) are consistent in the sense that if one knows the distribution of default-times for approach (i), one has a manageable "default/no default" indicator process for approach (ii) yielding the same results, and vice versa.

(2) What can go wrong, if one uses some indicator evolution which is not consistent in the sense of (1) — e.g. based on a Gaussian coupling of exponential random variables?

The consistency in question (1) can be weakened if the problem only concerns "survival-of-all" events instead of "mixed default/survival" events. The class of consistent distributions in the sense of question (1) might be very large — as the requirement of understanding the distribution as a model in approach (i) and approach (ii) can be fulfilled for many distributions with enough time at hand. However, most of these distributions are not feasible in practice, as we do not only need a model which is fully understood, but also feasible for simulation in terms of memory usage and sampling strategy. Therefore, a convenient *assumption*, which resolves — or at least diminishes — problems (a)–(e) from above, is a (continuous-time) time-homogeneous Markovian survival-indicator process. This is equivalent to conditional probabilities being determined by the current set of defaulted entities, but not on their specific default-times. The idea of using Markovian survival-indicator processes (even possibly time-inhomogeneous and only Markovian conditional on a set of intensity processes) is not new and has been discussed in [5] and [6]. These papers focus on the issue of pricing portfolio-credit derivatives. In the following we give a short overview on the "survival-of-all" and "mixed default/survival" problems.

1.1. *Problem one: "Survival-of-all" events*

In this special case the underlying problem only concerns the default/survival-of-all entities up to certain points in time. An example for such a problem is the valuation of a first-to-default swap on a basket of entities. Subsequently, one can demand a weaker version of consistency and feasibility — namely that the "survival-of-all" event and the corresponding indicator process are consistent and feasible. The class of consistent and feasible distributions for this problem was first studied in [7] and is related to a multivariate generalization of the univariate lack-of-memory property. In particular, a subclass fulfilling this property are min-stable multivariate exponential distributions. These are multivariate distributions with exponential margins and an extreme-value copula. Fundamental examples of this subclass, such as the Marshall–Olkin and the Gumbel–Hougaard distribution, are presented in this paper.

1.2. *Problem two: "Mixed default/survival" events*

Problems which depend on "mixed default/survival" events — and thus do not fall in the same category as problem one — require the original strict version of consistency. This leads (under previously outlined feasibility conditions) to time-homogeneous Markovian survival-indicators. This

general class is already known under the name *multivariate phase-type distributions*. This article analyzes further desirable theoretical and practical conditions on the resulting simulation process and as a result focuses on the subclasses of Marshall–Olkin distributions as well as a multivariate extension of the bivariate Freund distribution. In particular, the practically important requirement of having the Markov property also for sub-vectors of indicators leads to a new characterization of the Marshall–Olkin law that has been first discussed in [8] and is recalled here in the context of the present paper. Our general aim is to increase awareness of the fact that the stepwise simulation of default indicators (approach (ii) above) is a hard task in general, and in particular that the practical implementation is not feasible without huge efforts (both theoretical and computational), and that sizable errors and undesired effects may occur by iterating under the wrong conditions.

1.3. *Structure of the paper*

In Sec. 2 the survival-indicator process is introduced. It is shown that Markovianity of this process can be identified on a distributional level with a lack-of-memory type property. Subsequently, multiple lack-of-memory properties are presented and associated with certain classes of multivariate probability distributions. In particular, the min-stable multivariate exponential property (MSMVE) is introduced and is related to its characterization via extreme-value copulas and exponential margins.

Section 3 addresses the "survival-of-all" problem. Therefore, the concepts of self-chaining distributions and copulas, which were introduced in [7], are revisited and advanced. In particular, it is shown that the MSMVE characterization in terms of extreme-value copulas with exponential margins solves the problem. Then it is outlined that the widely used Gaussian-coupled exponential distributions do not fulfill that property. Moreover, choosing such a distribution for the step-innovations leads asymptotically to independence of the default-times, completely destroying dependence in the limit if the step size in time tends to zero.

In Sec. 4 the "mixed default/survival" problem is discussed, for looping default models, Freund distributions, and multivariate phase-type distributions. A special focus lies on the Marshall–Olkin class, leveraging its new characterization in terms of Markov property of vectors and subvectors of indicators, as in [8], and different simulation strategies as well as a convenient construction through Lévy-frailty models.

The final section concludes the article.

2. Default-time distributions and survival-indicator processes

Assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space on which all random objects of this section are defined. Throughout this article, let $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_d)'$ be a (non-negative) random vector of default-times[2] for d entities with joint- and marginal survival function(s) $\bar{F}$ and $\bar{F}_i, i \in [d] := \{1, \ldots, d\}$, respectively[3] and $\boldsymbol{Z} = \boldsymbol{Z}(t)$ be the corresponding survival indicator process which is defined by

$$Z_i(t) := 1_{\{\tau_i > t\}},\ i \in [d], t \geq 0.$$

In light of the introduction — and particularly as our questions of interest rely on iterating the survival-indicator process over periods with fixed length Δ — it may seem more appropriate (and also simpler) to work with the discretized version of $\boldsymbol{Z}$, hereby denoted by $\boldsymbol{Z}^{(\Delta)}$ and defined by

$$Z_i^{(\Delta)}(j) := Z_i(j\Delta),\ j \in \{0, \ldots, N\},\ i \in [d].$$

As outlined in the introduction, there are various arguments why it is convenient to assume that the underlying continuous-time process $\boldsymbol{Z}$ is also time-homogeneous Markovian. In the following another technical and a model building argument for this assumption are presented:

(a) *Technical argument:* The period-length, $\Delta > 0$, is usually an externally given quantity — e.g. set by the regulator as liquidity horizon or it is implicitly given from the existing IT-infrastructure. Hence, a model which can only be used consistently and feasible for very specific Δ is not desirable, as any (externally driven) change in Δ might destroy the models usability.

(b) *Model building argument:* From a model building perspective it is reasonable to assume that $\boldsymbol{Z}^{(\Delta)}$ has a representation with an underlying continuous-time process $\boldsymbol{Z}$. A deviation from the Markovian assumption above implies that the process $\boldsymbol{Z}$ either violates the time-homogeneity or the Markovian assumption entirely. However, if one

[2]For consistency, these "event"-times are referred to as default-times throughout this article, however, other notions such as fatality-, inter-arrival-, or inter-failure-times are equally applicable.

[3]For $\boldsymbol{\tau}$ and $\boldsymbol{s}, t \geq 0$, the multivariate survival function is defined by $\bar{F}(\boldsymbol{s}) := \mathbb{P}(\boldsymbol{\tau} > \boldsymbol{s})$ and the ith marginal survival function by $\bar{F}_i(t) := \mathbb{P}(\tau_i > t)$.

assumes that the time-homogeneous Markovian property of $\boldsymbol{Z}^{(\Delta)}$ is a tolerable deviation from reality — one should avoid choosing a model which violates those very properties on the continuous-time scale.

In summary, one can conclude that assuming an implied continuous-time, time-homogeneous Markovian survival-indicator process $\boldsymbol{Z}$ is a reasonable assumption, if one wants a feasible and consistent approach. In particular, this assumption is desirable from a technical aspect and also from a model building view if the underlying entities do note make the time-homogeneity assumption in itself unusable. Therefore, it is assumed throughout this article that, as a feasibility condition, $\boldsymbol{Z}$ is a continuous-time, time-homogeneous Markovian survival-indicator process.

2.1. *Markovian survival indicator-processes*

Let $\mathcal{I} = \{0,1\}^d$ and define the auxiliary function h to establish a bijection between the power set of $[d]$, denoted by $\mathcal{P}([d])$, and $\mathcal{I}$ by

$$h : \mathcal{P}([d]) \to \mathcal{I}, I \mapsto (1_{\{1\in I\}}, \ldots, 1_{\{d\in I\}})'.$$

A *survival-indicator process* is a stochastic process $\boldsymbol{Z} = \boldsymbol{Z}(t)$ on $\mathcal{I}$ fulfilling for all $s,t \geq 0$ and $J \subsetneq I \subseteq [d]$

$$\mathbb{P}(\boldsymbol{Z}(t+s) = h(I) \mid \boldsymbol{Z}(t) = h(J)) = 0.$$

This process is *Markovian* if for all $I, J \subseteq [d]$, $A \in \sigma(\boldsymbol{Z}(v) : v \leq t)$, and $s,t \geq 0$

$$\begin{aligned}\mathbb{P}(\boldsymbol{Z}(t+s) = h(I) \mid \boldsymbol{Z}(t) = h(J), A)\\ = \mathbb{P}(\boldsymbol{Z}(s+t) = h(I) \mid \boldsymbol{Z}(t) = h(J)).\end{aligned}$$

It is furthermore called *time-homogeneous* if additionally for all $s,t,v \geq 0$

$$\begin{aligned}\mathbb{P}(\boldsymbol{Z}(t+s+v) = h(I) \mid \boldsymbol{Z}(t+v) = h(J))\\ = \mathbb{P}(\boldsymbol{Z}(t+s) = h(I) \mid \boldsymbol{Z}(t) = h(J)).\end{aligned}$$

A time-homogeneous Markovian process satisfies

$$\mathbb{P}(\boldsymbol{Z}(t+s) = h(I) \mid \boldsymbol{Z}(t) = h(J)) = (\vec{e}_{\tilde{h}(J)})' \exp\{Qs\}\vec{e}_{\tilde{h}(I)},$$

where $\tilde{h} : \mathcal{P}([d]) \to \{0,1\}^{2^d}$ is an arbitrary bijection between the power set of $[d]$ and the set $\{1,\ldots,2^d\}$, which fulfills $\tilde{h}(I) < \tilde{h}(J) \Leftrightarrow |I| > |J|$ for all $I, J \subseteq [d]$,[4] $\vec{e}_k, k \in [2^d]$, is the canonical basis of $\mathbb{R}^{2^d}$, and $Q \in \mathbb{R}^{2^d \times 2^d}$ is an

[4]This property guarantees, that the resulting intensity matrix Q is an upper-triagonal matrix.

intensity matrix.[5] As it is assumed that $\tilde{h}$ is chosen such that for two sets with different cardinality, the one with more elements has the lower index, the matrix Q is upper trigonal with non-negative off-diagonal values and rows summing up to zero, i.e.

$$Q = \begin{pmatrix} q_{1,1} & & \star \\ & \ddots & \\ 0 & & q_{d,d} \end{pmatrix}.$$

Remark 2.1 (Intensities of a Markovian Process). *Let $Q \in \mathbb{R}^{n\times n}$ be a (not necessarily upper trigonal) intensity matrix for n states S — w.l.o.g. assume $S = [n]$. Then, one can construct a continuous-time, time-homogeneous Markovian process Z as follows (see [9]):*

(i) Let X_0 be the (possibly random) initial state, i.e. define $Z(0) := X_0$.
(ii) For $k \in \mathbb{N}_0$ define the kth jump time of Z by T_k (for $k = 0$ let $T_0 := 0$). Furthermore, assume that $Z(T_k) = i \in S$.
  (a) Let $E_{k+1} \sim \text{Exp}(-q_{ii})$ be an exponential random variable with rate $-q_{ii}$ which is, conditional on $Z(T_k)$, independent of $\sigma(\{E_l, T_l, l \leq k\})$.
  (b) Define $T_{k+1} := T_k + E_{k+1}$ and define $Z(t) = i\ \forall t \in (T_k, T_{k+1})$.
  (c) Let X_{k+1} be a discrete random variable on $S\backslash\{i\}$ with probabilities proportional to the ith row, i.e. $\mathbb{P}(X_{k+1} = j) = -q_{ij}/q_{ii}$. Moreover, assume that X_{k+1} is independent of $\sigma(\{E_l, T_l, l \leq k\})$ as well as independent of T_{k+1}.
  (d) Let $Z(T_{k+1}) = X_{k+1}$.
(iii) Repeat (ii) either infinitely often or until an absorbing state is reached, i.e. a state i with $q_{ii} = 0$. Note that for practical application the algorithm stops if $T_{k+1} > T$ for some terminal time-horizon $T > 0$.

It is useful to know that a time-homogeneous Markovian survival-indicator process is uniquely defined if for every non-zero transition, i.e. $h(J) \to h(I), I \subseteq J$, the transition probability for an arbitrary positive transition-time is known. This will be shown in the sequel. Let $\boldsymbol{\tau}$ be a default-vector with corresponding time-homogeneous Markovian survival-process $\boldsymbol{Z}$ and intensity-matrix Q. Furthermore, let $1 \leq K \leq d$,

[5]For a thorough introduction to continuous-time Markovian processes and a reference for this result, see [9], Ch. 8 and 9.

$I = \{i_1, \ldots, i_K\} \subseteq [d]$, $\boldsymbol{t}_I \geq 0$, $\pi \in \mathcal{S}_d$ be a permutation[6] with $\pi([K]) = I$ and $t_{\pi(1)} \geq \ldots \geq t_{\pi(K)}$, and define $A_{\pi,K}$ as the finite set

$$A_{\pi,K} := \{(I_1, \ldots, I_K) : \pi([k]) \subseteq I_k, I_k \subseteq I_{k+1}\ \forall k = 1, \ldots, K\},$$

where $t_{\pi(K+1)} = 0$ and $I_{K+1} = [d]$. Then

$$\mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{t}_I) = \sum_{(I_1,\ldots,I_K)\in A_{\pi,K}} \prod_{k=1}^{K} (\vec{e}_{\tilde{h}(I_{k+1})})' \exp\left\{(t_{\pi(k)} - t_{\pi(k+1)})Q\right\} \vec{e}_{\tilde{h}(I_k)}.$$

The assumption that the survival-indicator process is time-homogeneous Markovian has an important implication: Let $\boldsymbol{s} = (s_1, \ldots, s_d)' \geq 0$ be a deterministic vector of non-negative times and let $\pi \in \mathcal{S}_d$ be a permutation such that $s_{\pi(1)} \geq \ldots \geq s_{\pi(d)}$. Then for $t \geq 0$, $\boldsymbol{v} = \boldsymbol{s} + t$, and $v_{\pi(d+1)} = 0$ as well as $I_{d+1} = [d]$

$$\begin{aligned}
\mathbb{P}(\boldsymbol{\tau} > \boldsymbol{s} + t) &= \sum_{(I_1,\ldots,I_K)\in A_{\pi,d}} \prod_{k=1}^{d} (\vec{e}_{\tilde{h}(I_{k+1})})' \exp\left\{(v_{\pi(k)} - v_{\pi(k+1)})Q\right\} \vec{e}_{\tilde{h}(I_k)} \\
&= (\vec{e}_{\tilde{h}([d])})' \exp\{tQ\} \vec{e}_{\tilde{h}([d])} \\
&\quad \times \sum_{(I_1,\ldots,I_K)\in A_{\pi,d}} \prod_{k=1}^{d} (\vec{e}_{\tilde{h}(I_{k+1})})' \exp\left\{(s_{\pi(k)} - s_{\pi(k+1)})Q\right\} \vec{e}_{\tilde{h}(I_k)} \\
&= \mathbb{P}(\boldsymbol{\tau} > \boldsymbol{s})\mathbb{P}(\boldsymbol{\tau} > t).
\end{aligned}$$

This is equivalent to

$$\mathbb{P}(\boldsymbol{\tau} > \boldsymbol{s} + t \mid \boldsymbol{\tau} > t) = \mathbb{P}(\boldsymbol{\tau} > \boldsymbol{s}). \tag{1}$$

Analogously, one can derive for some $\emptyset \neq I \subseteq J \subseteq [d]$, and $t, v \geq 0$, that

$$\begin{aligned}
&\mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I + t + v \mid \boldsymbol{\tau}_J > t + v, \tau_{[d]\setminus J} \leq t + v) \\
&\quad = \mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I + t \mid \boldsymbol{\tau}_J > t, \tau_{[d]\setminus J} \leq t).
\end{aligned}$$

2.2. *Lack-of-memory properties*

It is not a coincidence that Eq. (1) collapses in the univariate case to the well-known *univariate lack-of-memory property* — also known as Cauchy's

[6]A permutation on $[d]$ is a bijection from $[d]$ to $[d]$; the set of all permutations on $[d]$ is denoted by $\mathcal{S}_d$.

functional equation — as in that case the time-homogeneity of the survival-indicator process implies exactly that *the probability of a survival-time bigger than $s+t$ conditional on a survival-time bigger than s is stationary with respect to t*, i.e.

$$\mathbb{P}(\tau > s + t \mid \tau > t) = \mathbb{P}(\tau > s). \tag{2}$$

It is a well-known fact that the class of non-negative distributions fulfilling Eq. (2) and having at least one continuity point[7] are exponential distributions — see, e.g., [10], p. 190. This property implies a very convenient simulation scheme if one is interested in the exponentially distributed survival-time of some entity:

$$1_{\{\tau > j\Delta\}} \overset{d}{=} \prod_{k=1}^{j} 1_{\{\tau^{(k)} > \Delta\}},$$

where $\tau^{(k)} \sim \tau$ are i.i.d. copies of τ and $\overset{d}{=}$ denotes equality in distribution.

The univariate lack-of-memory property, Eq. (2), can be extended to a multivariate property in multiple ways. In the following, a few of these are presented. Therefore, let $\boldsymbol{\tau}$ be a vector of non-negative random default-times and assume that the following conditions hold for all $\emptyset \neq I \subseteq [d]$ and $s_I, t_I, c_I, s, t \geq 0$.

- *Multivariate independent exponential lack-of-memory (MIELOM):*

$$\mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I + \boldsymbol{t}_I \mid \boldsymbol{\tau}_I > \boldsymbol{t}_I) = \mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I). \tag{3}$$

- *Multivariate Marshall–Olkin lack-of-memory (MMOLOM):*

$$\mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I + t \mid \boldsymbol{\tau}_I > t) = \mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I). \tag{4}$$

- *Min-stable multivariate exponential lack-of-memory (MSMVE):*

$$\mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{c}_I(s + t) \mid \boldsymbol{\tau}_I > \boldsymbol{c}_I t) = \mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{c}_I s). \tag{5}$$

- *Exponential-minima lack-of-memory (EM):*

$$\mathbb{P}(\boldsymbol{\tau}_I > s + t \mid \boldsymbol{\tau}_I > t) = \mathbb{P}(\boldsymbol{\tau}_I > s). \tag{6}$$

[7]This condition can be weakened in this context.

It was shown in [11] that (MIELOM) is equivalent to $\boldsymbol{\tau}$ having independent exponential components and (MMOLOM) is equivalent to $\boldsymbol{\tau}$ having a *Marshall–Olkin distribution*, i.e. there exist $\lambda_I \geq 0$, $\emptyset \neq I \subseteq [d]$, with $\sum_{I:i\in I} \lambda_I > 0$ for all $i \in [d]$, such that for all $\boldsymbol{t} \geq 0$

$$\mathbb{P}(\boldsymbol{\tau} > \boldsymbol{t}) = \exp\left\{ - \sum_{I:\emptyset\neq I\subseteq[d]} \lambda_I \max_{i\in I} t_i \right\}. \tag{7}$$

Furthermore, the authors provided the following stochastic model: Let $E^I, \emptyset \neq I \subseteq [d]$, be exponential random variables with rates $\lambda_I, \emptyset \neq I \subseteq [d]$, as above. Then the random vector $\boldsymbol{\tau}$ has the survival function in Eq. (7), where $\boldsymbol{\tau}$ is defined by

$$\tau_i := \min\{E^I : i \in I\},\ i \in [d]. \tag{8}$$

Marshall–Olkin distributions and continuous-time, time-homogeneous Markovian survival-indicator processes are deeply connected. In [8] it was shown that $\boldsymbol{\tau}$ has a Marshall–Olkin distribution if and only if for every non-empty subset I the marginal survival-indicator process $\boldsymbol{Z}_I(t) := (1_{\{\tau_i > t\}}, i \in I)'$ is time-homogeneous Markovian. The following theorem shows that every continuous-time, time-homogeneous Markovian survival-indicator process can be constructed using a finite sequence of Marshall–Olkin distributed random vectors.

Theorem 2.1. *Let Q be an intensity matrix of a time-homogeneous Markovian survival-indicator process. Consider the process $\boldsymbol{Z}$, which is constructed as follows:*

(i) Define $\boldsymbol{Z}(0) = h([d]) = (1, \ldots, 1)'$ (All entities are alive at time 0).

(ii) Assume that $\boldsymbol{Z}$ jumped $k \in \mathbb{N}_0$ times and define the time of the kth jump by T_k (for $k = 0$ let $T_0 := 0$). Furthermore, assume that $h^{-1}(\boldsymbol{Z}(T_k)) = I \subseteq [d]$.

(a) For $\emptyset \neq J \subseteq I$, let $E^J_{k+1} \sim \mathrm{Exp}(q_{h(I),h(I\setminus J)})$ be independent exponential random variables with rates $q_{h(I),h(I\setminus J)}$, which are, conditional on $\boldsymbol{Z}(T_k)$, also independent of all previously used random variables.

(b) Define
$T_{k+1} := T_k + \min_{\emptyset\neq J\subseteq I} E^J_{k+1}$ and $D_{k+1} := \mathrm{argmin}_{\emptyset\neq J\subseteq I} E^J_{k+1}$. Furthermore, define $\boldsymbol{Z}(t) := h(I)\ \forall t \in (T_k, T_{k+1})$ and $\boldsymbol{Z}(T_{k+1}) := h(I\setminus D_{k+1})$.

The resulting process $\boldsymbol{Z}$ *is time-homogeneous Markovian with intensity matrix* Q*. Note how the minimum operation in (b) is related to the Marshall–Olkin fatal shock model.*

Proof. The statement follows directly from Thm. A.1. □

It is a well-known fact, see e.g. [12], p. 174, that the class of MSMVE distributions is characterized by having exponential margins and a survival copula of *extreme-value kind*, i.e. a copula $\hat{C}$ that satisfies

$$\hat{C}(\boldsymbol{u}^t) = \hat{C}(\boldsymbol{u})^t, \ \forall \boldsymbol{u} \in [0,1]^d, t \geq 0. \tag{9}$$

Furthermore, it holds that (see, e.g., [13])

$$\text{MIELOM} \subsetneq \text{MMOLOM} \subsetneq \text{MSMVE} \subsetneq \text{EM}.$$

For the purpose of this article, we also define weaker versions of these properties, where the respective property only has to be fulfilled for $I = [d]$, and these are then referred to as weak versions of the respective properties, e.g., *weak exponential minima property (WEM).*

3. Problem one: Iterating "survival-of-all"

This section addresses problem one, for which only "survival-of-all" events are relevant. Let the vector of default-times be denoted by $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_d)'$. A "survival-of-all" event (similarly for a "first-to-default" event) has the form

$$\left\{ \min_{i \in [d]} \tau_i > s \right\}, \text{ for some } s > 0.$$

In practical applications, one has the options of either directly modeling the joint minimum of all default-times, or modeling the vector of all default-times and considering its minimum. Note that these approaches are sometimes called top-down- and bottom-up approach, respectively, not to be confused with the related but different top-down and bottom-up approaches for collateralized debt or loan obligations, see for example [14]. The top-down approach has the appealing advantage that everything becomes simpler and more advanced models, e.g. with stochastic intensity, become feasible. On the contrary, the bottom-up approach has the advantage that the default-times themselves are more "natural," compared to their joint minimum, as a model. This means in particular that in bottom-up models:

- There is usually good knowledge on the single default-times τ_i through historic data or CDS-quotes.
- On the contrary, there is comparably little understanding of the "first-to-default"-time that, barring heroic assumptions on pool homogeneity, granularity, and dependence, is usually accessed through brute force simulation methods.
- The dependence of other risk factors, e.g. equity, to the default-times is usually less complex than their dependence to the "first-to-default" time.
- A dependence-structure between default-times can be found, e.g., by mixtures of expert-judgment and model calibration to portfolio credit derivative data (e.g. CDO's), even though at the moment these markets are much less liquid than before the 2007-2008 crisis.

For the rest of this section the second option of modeling the default-times vector, namely the bottom up option, is considered.

The assumption of a continuous-time, time-homogeneous Markovian survival-indicator process has been motivated with the need to understand the increment- as well as the "terminal iterated"-distribution and to limit the data which has to be stored for simulation. For this very problem we can weaken these requirements by simply asking that the survival-indicator process has a time-homogeneous probability to stay in the "no default"-state. In other words, for this particular problem, the distribution of default-times is feasible if it fulfills the *weak exponential minima (WEM)* property:

$$\mathbb{P}(\boldsymbol{\tau} > s + t \mid \boldsymbol{\tau} > t) = \mathbb{P}(\boldsymbol{\tau} > s). \tag{10}$$

Another formulation of this class, fulfilling Eq. (10), is the following:

$$\begin{aligned} &\text{"terminal one-shot survival probability up to } t_1 + \ldots + t_N\text{"} \\ &\quad = \mathbb{P}(\boldsymbol{\tau} > t_1 + \ldots + t_N) = \mathbb{P}(\boldsymbol{\tau}^{(1)} > t_1) \cdot \ldots \cdot \mathbb{P}(\boldsymbol{\tau}^{(N)} > t_N) \\ &\quad = \text{"terminal iterated survival probability with steps } t_1, \ldots, t_N\text{,"} \end{aligned}$$

where $\boldsymbol{\tau}^{(k)}, k \in [d]$, are i.i.d. copies of $\boldsymbol{\tau}$. The class of distributions fulfilling the WEM-property is potentially large, as the following examples show, and to the best knowledge of the authors it is not characterized in any other way.

Example 3.1. Let $\boldsymbol{\tau}$ have a bivariate survival function corresponding to an independence survival-copula and the marginal survival functions $\bar{F}_1(t) = (t+1)\exp\{-t\}$ and $\bar{F}_2(t) = (1+t)^{-1}$, respectively. The functions $\bar{F}_i, i \in [2]$, are both proper survival functions as they are decreasing, continuous, and tend to zero and one for $t \to 0$ and $t \to \infty$, respectively. Then the joint

minimum, $\min_{i\in[2]} \tau_i$, is exponential, and in particular $\boldsymbol{\tau}$ fulfills the WEM-property, but neither τ_1 nor τ_2 are exponential,

$$\mathbb{P}\left(\min_{i\in[2]} \tau_i > t\right) = (t+1)\exp\{-t\} \cdot (1+t)^{-1} = \exp\{-t\}.$$

Example 3.2. Let $\boldsymbol{\eta}$ be a $(d-1)$-dimensional non-negative random vector, E an exponential random variable with rate $\lambda > 0$, and Π a random variable on the set of permutations on $[d]$. Define $\tilde{\boldsymbol{\tau}} := (E, E + \boldsymbol{\eta}')'$ and $\boldsymbol{\tau}$ by

$$\boldsymbol{\tau} := (\tilde{\tau}_{\Pi(1)}, \ldots, \tilde{\tau}_{\Pi(d)})'.$$

Then $\boldsymbol{\tau}$ has the WEM-property, as by construction $\min_{i\in[d]} \tau_i = E$.

The rest of this section has two purposes:

- The assumption of a time-homogeneous Markovian first-default survival indicator has strong links to multivariate lack-of-memory properties. It is shown that, in particular, all MSMVE distributions fulfill this property. As a well-known representative of this class, the Gumbel–Hougaard copula and the corresponding Gumbel–Hougaard exponential distribution[8] are introduced as an example.
- Showing that the popular approach of (independent in time) Gaussian-coupled exponential increments does not fulfill the WEM-property. Furthermore, it is shown that this approach kills dependence asymptotically for $N \to \infty$ — meaning the "terminal iterated" dependence is approximately that of independent-coupled exponential random variables.

3.1. *Lack-of-memory properties revisited*

Let Δ be the period step-size, T the final horizon, and N the number of periods up to T, i.e. $T = N\Delta$.

In [7], in the context of the problem of "survival-of-all", the authors tried to bridge the gap between the question

[8]The Gumbel–Hougaard distribution is the multivariate extension defined later in Eq. (12). This was originally introduced in [15] for the bivariate case. It is not to be confused with the two other bivariate exponential distributions introduced in that very paper that are also named after Emil J. Gumbel. One of those, with the survival function $\exp\{-\lambda_1 t_1 - \lambda_2 t_2 - \theta t_1 t_2\}$, is characterized by a lack-of-memory property called *bivariate remaining life constancy*, see, e.g., [16], [17], which has the interpretation that, conditional on the survival of the respective other component up to an arbitrary time, both variables are exponential, cf. [18].

Which distributions have equal "terminal one-shot" and "terminal iterated" survival probabilities for common step-size Δ?

and properties of survival copulas corresponding to multivariate exponential distributions. This leads to the definition of so called *self-chaining* copulas — or *self-chaining distributions*.

In the following, this approach will be (broadly) outlined, advanced and generalized, exploring the full lack-of-memory implications and characterization for the extreme-value copula with exponential margin solution obtained initially in [7]. We will confirm also the special solutions found in [7], namely the Gumbel–Hougaard copula and the Marshall–Olkin copula, further specifying the properties of these solutions, although we will not address the bivariate Pickands functions solution here. For further details on Pickands functions see, for example, [19] or [20].

Definition 3.1. The distribution of $\boldsymbol{\tau}$ has the *weak common Δ-period exponential minima* (WCPEM(Δ))-property if for every two natural numbers $j, k \in \mathbb{N}$

$$\mathbb{P}(\boldsymbol{\tau} > (j+k)\Delta \mid \boldsymbol{\tau} > j\Delta) = \mathbb{P}(\boldsymbol{\tau} > k\Delta).$$

It has the *common Δ-period exponential minima (CPEM (Δ))*-property if for all non-empty $I \subseteq [d]$ the vector $\boldsymbol{\tau}_I$ has the (WCPEM(Δ))-property.

It can be easily shown that this property can be rewritten as follows:

Definition 3.2. A random vector $\boldsymbol{\tau}$ is *Δ-periodic self-chaining* if for all $j \in \mathbb{N}$

$$\mathbb{P}(\boldsymbol{\tau} > j\Delta) = \mathbb{P}(\boldsymbol{\tau} > \Delta)^j.$$

For a Δ-periodic self-chaining distribution, the corresponding survival-copula $\hat{C}$ is called *$\mathbb{N}$-self-chaining in the point* $(\bar{F}_1(\Delta), \ldots, \bar{F}_d(\Delta))'$.

From Def. 3.1 it is visible that a distribution fulfilling the (W)CPEM(Δ)-property for all $\Delta > 0$ fulfills the (W)EM-property and vice versa. Therefore, in light of Def. 3.2, the following definition follows.

Definition 3.3. A random vector $\boldsymbol{\tau}$ is *self-chaining* if for all $t > 0$

$$\mathbb{P}(\boldsymbol{\tau} > t) = \mathbb{P}(\boldsymbol{\tau} > 1)^t.$$

For a self-chaining distribution, the corresponding survival-copula $\hat{C}$ is called *$\mathbb{R}$-self-chaining* (or *self-chaining*) in the point $(\bar{F}_1(1), \ldots, \bar{F}_d(1))'$.

Let $\boldsymbol{\tau}$ have exponential margins and define $\boldsymbol{u} := (\bar{F}_1(1), \ldots, \bar{F}_d(1))'$. Then $\boldsymbol{\tau}$ is self-chaining if and only if the survival-copula $\hat{C}$ fulfills (for the specific $\boldsymbol{u}$)

$$\hat{C}(\boldsymbol{u}^t) = \hat{C}(\boldsymbol{u})^t, \ \forall t > 0. \tag{11}$$

Equation (11) is well-known from extreme-value theory, as the class of copulas fulfilling Eq. (11) for all $\boldsymbol{u} \in [0,1]^d$, cf. Eq. (9), is that of *extreme-value copulas (EVCs)* and furthermore, that the class of min-stable multivariate exponential distributions, cf. Eq. (5), is characterized by a coupling of EVC's and exponential margins, see [12], p. 174.

A self-chaining survival-copula in the point $\boldsymbol{u} \in [0,1]^d$ can only be coupled with exponential margins with rates $\lambda_i = -\ln u_i, i \in [d]$, to a self-chaining distribution, while an extreme-value copula can be coupled with any exponential margin to a self-chaining distribution. In general, it should be noted that almost all lack-of-memory properties get lost if the underlying survival-copula is re-coupled with different marginal distributions — even if one stays in the exponential class.

An example for a (survival-)copula which is self-chaining in arbitrary points $\boldsymbol{u} \in [0,1]^d$ is the *Gumbel–Hougaard copula*, see [15],[16],[21],[22], which is implicitly defined by the following multivariate exponential distribution ($\boldsymbol{\lambda} > 0, \theta \geq 1$)

$$\mathbb{P}(\boldsymbol{\tau} > \boldsymbol{s}) = \exp\left\{ -\left(\sum_{i=1}^{d} (\lambda_i s_i)^\theta \right)^{\frac{1}{\theta}} \right\}, \ \boldsymbol{s} \geq 0. \tag{12}$$

In [22], it was proven that the class of Gumbel–Hougaard copulas are the only copulas which are both extreme-value- and Archimedean copulas, see also [7] for an alternative proof.

An example for a distribution with exponential minima, which is not min-stable multivariate exponential, with a recipe from [13] for the bivariate case.

(1) Let $E_I^{(k)}$ be independent exponential random variables with rates $\lambda_I^{(k)}, k \in [2], \emptyset \neq I \subseteq [2]$.
(2) Let $\tilde{\boldsymbol{\tau}}^{(k)} = (\tilde{\tau}_1^{(k)}, \tilde{\tau}_2^{(k)})', k \in [2]$, be defined by

$$\tilde{\tau}_i^{(k)} := \min\{E_{\{i\}}^{(k)}, E_{[2]}^{(k)}\}, \ i, k \in [2],$$

i.e. both $\tilde{\boldsymbol{\tau}}^{(1)}$ and $\tilde{\boldsymbol{\tau}}^{(2)}$ are Marshall–Olkin distributed.

(3) Let $\boldsymbol{\tau}$ for $p \in (0,1)$ and $a_i^{(k)}, i, k \in [2]$, be defined by

$$\tau_i = X a_i^{(1)} \tilde{\tau}_i^{(1)} + (1-X) a_i^{(2)} \tilde{\tau}_i^{(2)},\ i \in [2],$$

where X is a Bernoulli variable with "success probability" p.

Choose $\lambda_{\{1\}}^{(1)} = 1/2$, $\lambda_{\{2\}}^{(1)} = 1$, $\lambda_{[2]}^{(1)} = 2$, $\lambda_{\{1\}}^{(2)} = 2/3$, $\lambda_{\{2\}}^{(2)} = 1/2$, $\lambda_{[2]}^{(2)} = 1$ as well as $a_1^{(1)} = 1/2$, $a_2^{(1)} = 1$, $a_1^{(2)} = 1/3$, and $a_2^{(2)} = 1/2$; then the attained distribution has EM but is not MSMVE. The attained distribution is a mixture of MO-coupled, i.e. having a copula from a Marshall–Olkin survival copula, exponential random variables. The key for the EM-property to hold is to make sure that the mixed MO-coupled exponential distributions have equal diagonal-functions for all margins. This concept can be extended to arbitrary dimensions for the creation of distributions with EM.

In more basic terms, this discussion highlights a tension between the full Marshall–Olkin law and the Marshall–Olkin copula with possibly different exponential margins. The initial results in [7] include the solution given by the Marshall–Olkin copula with possibly re-scaled exponential margins, leading to a multivariate distribution that is different from a fully consistent Marshall–Olkin law. In more intuitive terms, we can say that re-scaling the margins with new exponentials breaks the natural consistency between margins and dependence that is a key property of the Marshall–Olkin law. In general, arbitrarily decoupling the margins and the dependence structure may result in paradoxical results when analyzing wrong way risk in CDS trades, see, for example, the low dimensional examples in [23], [3], [2], and [24].

For the construction of high-dimensional models it might be convenient to know that there is another recent approach for the generation of (extendible) EM-distributed random vectors via first hitting times of matrix-mixtures of subordinators which are weakly infinitely divisible with respect to time over random exponential barriers, see [25], [26].

3.2. *Change in dependence when iterating non-self chaining copulas*

In the following, a standard approach which is widely used in the financial industry is critically analyzed: The discretely iterated Gaussian-coupled exponential margins survival-indicator process. Let, as before, $T > 0$, $N \in \mathbb{N}$, and $\Delta := T/N$ and define for $j \in \mathbb{N}$

$$\boldsymbol{Z}^{(\Delta)}(j+1) \mid \{\boldsymbol{Z}^{(\Delta)}(j) = \mathbf{1}\} := 1_{\{\boldsymbol{\zeta}_{j+1} > \Delta\}},$$

for independent and identically distributed $\boldsymbol{\zeta}_{j+1} \sim C_\Phi(\rho) \oplus (\bar{F}_1, \ldots, \bar{F}_d)$, where $C_\Phi(\rho)$ is the Gaussian copula with equi-correlation $\rho > 0$ and $\bar{F}_i, i \in [d]$, are exponential survival functions.

Assume first that $\boldsymbol{\zeta}_j, j \in [N]$, are constructed with an arbitrary copula coupled with exponential margins; then the "terminal iterated" probability for the "survival-of-all" event is

$$\mathbb{P}\left(\boldsymbol{Z}^{(T/N)}(N) = \mathbf{1}\right) = \left(\mathbb{P}\left(\boldsymbol{\zeta} > \frac{T}{N}\right)\right)^N. \tag{13}$$

From multivariate extreme-value theory it is known that for $N \to \infty$ the expression in Eq. (13) either converges to a min-stable multivariate exponential distribution[9] or does not converge at all, see [12].

Definition 3.4. Let $\hat{C}$ be an extreme-value copula. Every copula $\hat{C}_F$ with

$$\lim_{n\to\infty} \hat{C}_F(\boldsymbol{u}^{1/n})^n = \hat{C}(\boldsymbol{u}),\ \forall \boldsymbol{u} \in [0,1]^d,$$

is said to be in the *domain of attraction* of $\hat{C}$.

Theorem 3.1. *Let $d = 2$, then the* Clayton copula, Frank copula, *and the Gaussian copula for $\rho < 1$ are in the domain of attraction of the independence copula.*

Proof. See [12],[27]–[29]. □

This implies in particular for $d = 2$ and large N that the distribution of $\boldsymbol{\tau}$ is approximately that of independent exponential random variables. Hence, and this is a word of warning, for large N the Gaussian-coupling kills the correlation of the "terminal iterated" law.

Remark 3.1. The asymptotic "terminal iterated" dependence can be inferred if the survival-copula of the iterated law lies in the domain of attraction of some extreme-value copula, e.g. in Thm. 3.1, it was shown that the bivariate non-comonotonic Gaussian-, Clayton-, and Frank copulas are in the domain of attraction of the independence copula, see [12], p. 141 and also [29] for an early account on asymptotic independence of the Gaussian copula. The bivariate exchangeable t-copula lies in the domain of attraction of the *t-EV copula*, which is for finite degrees of freedom not the independence copula and depends on the degrees of freedom as well as the correlation parameter, see [30]. Furthermore, if $\hat{C}_F$ lies in the domain of

[9] A vector of independent exponentially distributed random variables is also MSMVE.

attraction of $\hat{C}$, then their upper-tail-dependence coefficient coincides — in particular, if a copula $\hat{C}_F$ incorporates asymptotic independence and lies in the domain of attraction of an extreme-value copula $\hat{C}$, then $\hat{C}$ is the independence copula, see e.g. [30], pp. 587–588. Moreover, if $\hat{C}_F$ is a d-dimensional copula which lies in the domain of attraction of $\hat{C}$ and incorporates pairwise asymptotic independence, then $\hat{C}$ is the independence copula, see, e.g., [30], p. 591. This implies in particular that also the d-dimensional exchangeable Gaussian-copula with $\rho < 1$ lies in the domain of attraction of the independence copula.

In the following example, this effect is analyzed numerically for bivariate Gaussian-coupled exponential distributions with rates $\lambda_{\mathrm{IG}} = 1\%$ and $\lambda_{\mathrm{SG}} = 4.5\%$, corresponding to an investment grade (IG) or speculative grade (SG) entity. The "terminal one-shot" and "terminal iterated" probability for the "survival-of-all" event is denoted by

$$p_T := \mathbb{P}(\zeta > T) \text{ or } p_\Delta^N := \mathbb{P}(\zeta > \Delta)^N = \mathbb{P}(\tau > T).$$

In Tables 1 and 2, the result of this analysis for two different settings with different final time-horizons as well as different numbers of iterations can be observed. The results illustrate the statement from Thm. 3.1, i.e. that Gaussian-coupled exponential distributions with $\rho < 1$ do not have the WEM-property. Moreover, the relative error is sizable and becomes larger for higher marginal rates and higher correlation, which is especially undesirable.

Table 1. Comparison of "terminal one-shot" and "terminal iterated" survival probabilities for $T = 5y$ and $N = 1000$.

λ_1	λ_2	ρ	p_T	p_Δ^N	% Diff.
0.010	0.010	0.25	0.9084	0.9049	0.38%
0.010	0.010	0.50	0.9142	0.9057	0.95%
0.010	0.010	0.75	0.9238	0.9103	1.48%
0.010	0.045	0.25	0.7679	0.7598	1.07%
0.010	0.045	0.50	0.7785	0.7614	2.24%
0.010	0.045	0.75	0.7908	0.7698	2.73%
0.045	0.045	0.25	0.6592	0.6382	3.29%
0.045	0.045	0.50	0.6851	0.6421	6.7%
0.045	0.045	0.75	0.7187	0.6605	8.81%

In Fig. 1, the relative error is visualized for four additional survival-copulas, i.e. the t-, Clayton-, Frank-, and Gumbel-copula, and multiple

Table 2. Comparison of "terminal one-shot" and "terminal iterated" survival probabilities for $T = 30y$ and $N = 1000$.

λ_1	λ_2	ρ	p_T	p_Δ^N	% Diff.
0.010	0.010	0.25	0.5765	0.5496	4.91%
0.010	0.010	0.50	0.6084	0.5545	9.71%
0.010	0.010	0.75	0.6483	0.5766	12.43%
0.010	0.045	0.25	0.2169	0.1929	12.47%
0.010	0.045	0.50	0.2389	0.1974	21.01%
0.010	0.045	0.75	0.2553	0.2142	19.2%
0.045	0.045	0.25	0.0949	0.0682	39.17%
0.045	0.045	0.50	0.1268	0.0728	74.09%
0.045	0.045	0.75	0.1667	0.0899	85.38%

Kendall's τ, denoted by τ_K, where the underlying copula parameters are calibrated such that a certain τ_K is achieved. One can see that the error is strongly dependent on the chosen rank correlation. Furthermore, the Gaussian coupling seems to have the largest errors for $\tau_K \leq 75\%$, while the error for the t-coupling is rather small in comparison. An explanation for the latter observation could be that the bivariate t-copula converges for a low degree of freedoms comparably fast, see [27], and the t-EV copula still incorporates information on ν and τ_K.

In conclusion, these calculations show that a coupling with the Gaussian-, Frank-, or Clayton copula can lead to sizable differences in the terminal probabilities. This is not a surprising result, as it was already shown theoretically that the terminal probabilities can only match if the iterated distribution has the WEM-property (e.g. an MSMVE-distribution) and that the iteration of Gaussian-copulas leads asymptotically to independence; however, this analysis underscores the severity of the mismatch.

4. Problem two: "Mixed default/survival" events

So far, the problem of finding conditions under which the "survival-of-all" simulation can be iterated (feasible) in a way that makes it consistent to a single step simulation was addressed. However, while the "survival-of-all" may be of interest in situations where one wishes to exclude even a single default, or for the valuation of a first-to-default CDS, it is more interesting to look at the general problem of iterating in presence of "mixed-default/survival"-states. This problem, "problem two," is the topic of the present section and conditions for the feasible and consistent simulation of "mixed-default/survival"-indicators up to a terminal time are analyzed.

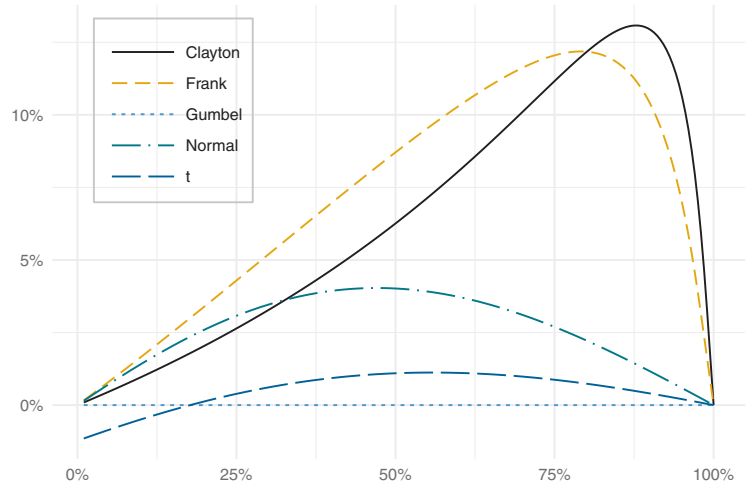


Fig. 1. Relative deviation of p_T and p_Δ^N in % vs. Kendalls's τ for $T = 5y$, $\lambda_i = 4.5\%, i = 1, 2$, $N = 10$, and 3 degrees of freedom for the t-distribution, see [28].

Finally, examples such as the Marshall–Olkin distribution and a multivariate extension of the Freund distribution are presented.

4.1. *The looping default model and the Freund distribution*

One of the most intuitive models for contagion effects in portfolio-credit risk is the so-called "looping default"-model, the terminology being introduced in one of the first works on counterparty credit risk pricing, see [31]. In the bivariate case, the model can easily be explained: Let C_1 and C_2 be two companies with respective default intensities for $t \geq 0$

$$\tilde{\lambda}_1(t) = \lambda_1 + 1_{\{\tau_2 \leq t\}}(\eta_1 - \lambda_1),$$
$$\tilde{\lambda}_2(t) = \lambda_2 + 1_{\{\tau_1 \leq t\}}(\eta_2 - \lambda_2),$$

where $\lambda_1, \lambda_2, \eta_1, \eta_2 > 0$. Loosely speaking, this means that the default/survival-probabilities of company C_1 depend on the default/survival of company C_2 and vice versa. This explains the notion of a "looping-default" model, as the influence of companies on each-others default/survival-probabilities can be depicted as a loop. This model formulation can easily be generalized to non-linear or stochastic hazard functions. Constructing a well-defined probability space, however, supporting such a multivariate distribution is non-trivial. Therefore, it was initially assumed

that the set of companies can be divided into two classes $\mathcal{A}$ and $\mathcal{B}$, such that the default of a company from set $\mathcal{A}$ can influence the default of a company from $\mathcal{B}$, but not vice versa. As a consequence the model can be formulated recursively in the spirit of a classical intensity-based model, see [31]. The problem of constructing the distribution in the general model (with hazard-rate functions which are deterministic functions of time and default history) on a well-defined probability space has been investigated in subsequent articles and finally was resolved in [32], where the "looping-default" model is defined using the so-called "total hazard construction," which originates from the statistical literature, see [33] and [34]. The total hazard construction defines a d-dimensional random vector $\boldsymbol{\tau}$ of default-times as a function of d independent unit exponential random variables $E_1, \ldots, E_d$, such that the corresponding default intensities satisfy certain relations that are specified a priori. This construction algorithm is, however, rather complicated to implement in practice, and in particular has no natural coherence with stepwise simulation — rendering it inconvenient for our purpose. As a first example of the total hazard construction, [32] reconsiders the "looping default" of [31] in a two-dimensional setup. In [6] and [5], it was shown that the "looping default" model falls into the class of default models whose survival indicator process is a Markov chain, which provides an alternative stochastic construction being naturally consistent with stepwise simulation. Interestingly, in the bivariate case the probability law of $\boldsymbol{\tau} = (\tau_1, \tau_2)'$ is well-known in the statistical literature as well.

Remark 4.1 (Looping default model/Freund distribution). *The bivariate distribution which is derived in [32] coincides precisely with the so-called bivariate Freund distribution, which is an "old friend" from reliability theory, see [35]. In other words, the looping default has incidentally been known for many years in the statistical literature by the name "Freund distribution." The fact that both distributions coincide can be observed by comparing the bivariate densities derived in [32] and [35], respectively. The details are provided below.*

In the sequel, a new construction for the Freund distribution based on continuous-time, time-homogeneous Markovian processes is presented. This construction provides an alternative access to this probability law, which is in particular based on a stepwise-simulation ansatz. Moreover, it can be easily generalized to dimensions $d > 2$ and to extensions with simultaneous defaults.

Consider two companies' default-times $\boldsymbol{\tau} = (\tau_1, \tau_2)'$. We construct the associated survival indicator process $\boldsymbol{Z}(t) := (1_{\{\tau_1 > t\}}, 1_{\{\tau_2 > t\}})'$ as a continuous-time, time-homogeneous Markov chain. This process is fully described by its intensity matrix Q. Let the four states $(1,1), (0,1), (1,0)$, and $(0,0)$ be indexed by the numbers $1, 2, 3$, and 4 and define the intensity matrix $Q \in \mathbb{R}^{4\times 4}$ by

$$Q = \begin{pmatrix} -(\lambda_1 + \lambda_2) & \lambda_1 & \lambda_2 & 0 \\ 0 & -\eta_2 & 0 & \eta_2 \\ 0 & 0 & -\eta_1 & \eta_1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

where the "initial intensities" $\lambda_i > 0, i \in [2]$, and the "intensities conditional on second-party default" $\eta_i > 0, i \in [2]$, are positive real numbers. It is easy to verify that in case the condition $\eta_i \neq \lambda_1 + \lambda_2, i \in [2]$, is fulfilled the matrix Q is diagonalizable,[10] i.e. we can find a matrix M such that

$$M^{-1} Q M = \operatorname{diag}(-(\lambda_1 + \lambda_2), -\eta_2, -\eta_1, 0),$$

where the transformation-matrix M has the eigenvectors of Q as column vectors, i.e.

$$M = \begin{pmatrix} 1 & \frac{\lambda_1}{\lambda_1+\lambda_2-\eta_2} & \frac{\lambda_2}{\lambda_1+\lambda_2-\eta_1} & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

This intensity matrix Q can be interpreted as follows (cf. Thm. 2.1): Being in a certain state corresponds to a certain row of the matrix — e.g. the process starts in state $(1,1)$ corresponding to row 1. For each other state $(0,1), (1,0)$, and $(0,0)$ there are independent latent exponential random variables with rates $Q_{(1,1),(0,1)}, Q_{(1,1),(1,0)}$, and $Q_{(1,1),(0,0)}$. The process $\boldsymbol{Z}$ reacts only on the smallest of these random variables and moves to the corresponding target state. A rate of zero corresponds to the corresponding random variable being "degenerate," i.e. almost surely equal to infinity. Therefore, the chain cannot go directly from no default $(1,1)$ to joint default $(0,0)$. Finally, as Q has vanishing row sums, the ith diagonal entry corresponds to the negative rate of the minimum of all latent exponential random variables for transition out of i. The same logic applies to the other rows of Q. In particular, after the default of one company, the hazard rate

[10]The case $\eta_i = \lambda_1 + \lambda_2$ for some $i \in [2]$ is still a valid model. However, as the matrix Q is not diagonalizable, the analytical calculation of probabilities becomes more involved.

of the remaining company changes from λ_i to η_i, and the bottom row of Q is zero because the state of two defaults is an absorbing state. Using diagonalization, one can show that for $t > 0$ the entries of the transition matrix

$$P[t] := e^{t\,Q} = M^{-1} \exp\{tMQM^{-1}\}M$$

are given by

$$\begin{aligned}
P_{(1,1),(1,1)}[t] &= e^{-(\lambda_1+\lambda_2)\,t},\\
P_{(1,1),(0,1)}[t] &= \frac{\lambda_1}{\lambda_1+\lambda_2-\eta_2}\left(e^{-\eta_2\,t} - e^{-(\lambda_1+\lambda_2)\,t}\right),\\
P_{(1,1),(1,0)}[t] &= \frac{\lambda_2}{\lambda_1+\lambda_2-\eta_1}\left(e^{-\eta_1\,t} - e^{-(\lambda_1+\lambda_2)\,t}\right),\\
P_{(1,1),(0,0)}[t] &= -\frac{\lambda_1}{\lambda_1+\lambda_2-\eta_2}\,e^{-\eta_2\,t} - \frac{\lambda_2}{\lambda_1+\lambda_2-\eta_1}\,e^{-\eta_1\,t}\\
&\quad + 1 + \left(\frac{\lambda_1}{\lambda_1+\lambda_2-\eta_2} + \frac{\lambda_2}{\lambda_1+\lambda_2-\eta_1} - 1\right) e^{-(\lambda_1+\lambda_2)\,t},\\
P_{(0,1),(0,1)}[t] &= e^{-\eta_2\,t}, \quad P_{(0,1),(0,0)}(t) = 1 - e^{-\eta_2\,t},\\
P_{(1,0),(1,0)}[t] &= e^{-\eta_1\,t}, \quad P_{(1,0),(0,0)}(t) = 1 - e^{-\eta_1\,t},
\end{aligned}$$

and all other entries of P being zero. In particular, we calculate

$$\begin{aligned}
&\mathbb{P}(\tau_1 > t_1, \tau_2 > t_2)\\
&\quad = \begin{cases} P_{(1,1),(1,1)}(t_1)\left(P_{(1,1),(1,1)}(t_2-t_1) + P_{(1,1),(0,1)}(t_2-t_1)\right), & t_2 \geq t_1\\ P_{(1,1),(1,1)}(t_2)\left(P_{(1,1),(1,1)}(t_1-t_2) + P_{(1,1),(1,0)}(t_1-t_2)\right), & t_1 > t_2 \end{cases}\\
&\quad = \begin{cases} \frac{\lambda_2-\eta_2}{\lambda_1+\lambda_2-\eta_2}\,e^{-(\lambda_1+\lambda_2)\,t_2} + \frac{\lambda_1}{\lambda_1+\lambda_2-\eta_2}e^{-\eta_2\,t_2-(\lambda_1+\lambda_2-\eta_2)\,t_1}, & t_2 \geq t_1\\ \frac{\lambda_1-\eta_1}{\lambda_1+\lambda_2-\eta_1}\,e^{-(\lambda_1+\lambda_2)\,t_1} + \frac{\lambda_2}{\lambda_1+\lambda_2-\eta_1}e^{-\eta_1\,t_1-(\lambda_1+\lambda_2-\eta_1)\,t_2}, & t_1 > t_2. \end{cases}
\end{aligned}$$

The latter distribution is precisely the Freund distribution, which can be seen by comparing it to Eq. (47.26) in [16], p. 356. Note additionally, that the so-called ACBVE$(\tilde{\eta}_1, \tilde{\eta}_2, \tilde{\eta}_{12})$-distribution, defined in [36], arises as the three-parametric subfamily of the Freund distribution, obtained from the parameters

$$\lambda_1 = \tilde{\eta}_1 + \frac{\tilde{\eta}_{12}\tilde{\eta}_1}{\tilde{\eta}_1+\tilde{\eta}_2},\ \lambda_2 = \tilde{\eta}_2 + \frac{\tilde{\eta}_{12}\tilde{\eta}_2}{\tilde{\eta}_1+\tilde{\eta}_2},\ \eta_1 = \tilde{\eta}_1 + \tilde{\eta}_{12},\ \eta_2 = \tilde{\eta}_2 + \tilde{\eta}_{12}.$$

Multivariate extensions of the described Markov chain construction, leading to the Freund distribution, are now clearly straightforward. One can simply define the intensity matrix Q as follows: For each set $I \subseteq [d]$ one has to define exponential rates η_J for all subsets $J \subseteq I$ with $|J| = |I|-1$, i.e. corresponding to exactly one additional default scenario, and write them in

the respective entry $Q_{h(I),h(J)}$. All other off-diagonal entries of Q are set to zero, and then the diagonal elements are computed as the negative of the sum over all previously defined row entries. Similarly, one can generalize the model to allow for multiple defaults and also assign positive exponential rates to subsets $J \subseteq I$ with $|J| = |I| - k$, $k \geq 1$.

For stepwise simulation along the Δ-grid, one only requires the matrix $P[\Delta] = \exp\{\Delta\, Q\}$, which can be computed easily if Q is diagonalizable or otherwise numerically (e.g. `expm` in MATLAB or `Matrix::expm` in R).

Remark 4.2. The class of distributions attained in continuous-time, time-homogeneous Markovian survival-indicator processes coincides with the class of *multivariate phase-type distributions* which were introduced in [37], see also [38]. Multivariate phase-type distributed random vectors $\boldsymbol{\tau}$ are defined implicitly through a continuous-time, time-homogeneous Markovian process Z and absorbing sets $A_i, i \in [d]$, such that $\bigcap_{i\in d} A_i$ is absorbing and

$$\tau_i := \inf\{t > 0 : Z(t) \in A_i\},\ i \in [d].$$

In particular, it follows that all resulting marginal distributions of $\boldsymbol{\tau}$ are univariate phase-type distributions.

4.2. *Marshall–Olkin distributions*

Throughout this section, we denote by $\boldsymbol{Z}_I$ the I-margin of the survival-indicator process $\boldsymbol{Z}$ which only consists of the components indexed by $I \subseteq [d]$. This section starts with summarizing the findings and results of [8], in which it is emphasized that for practical applications even the assumption of a continuous-time, time-homogeneous Markovian survival-indicator process has serious drawbacks if the corresponding default-times vector $\boldsymbol{\tau}$ does not have a Marshall–Olkin distribution. The findings are:

(a) In general, even if $\boldsymbol{Z}$ is time-homogeneous Markovian the survival-indicator $\boldsymbol{Z}_I$, corresponding to a subportfolio $\emptyset \neq I \subsetneq [d]$, might not fulfill this property. As a result, even if a certain study involves only the default-times $\boldsymbol{\tau}_I$ one has to simulate the full survival-indicator process $\boldsymbol{Z}$. This is undesirable for two reasons: Firstly, simulations only considering subportfolios cannot be performed more efficiently than via the full portfolio simulation. Second, every restructuring of the credit portfolio requires a careful adjustment and possibly a reevaluation of the whole default model (see (b) for a detailed account).

(b) If the underlying credit portfolio is subject to restructuring, the Markovian survival-indicator model is, in general, problematic. This is best

explained in the case where an additional entity $d+1$ is added to the credit portfolio. Then, each state L splits into two separate states L and $\tilde{L} := L \cup \{d+1\}$, and following this logic each "transition-rate" in the intensity matrix has the interpretation

$$P_{h(I),h(J)}[\Delta] = \mathbb{P}\left(Z((k+1)\Delta) \in h(\{J, \tilde{J}\}) \mid Z(k\Delta) \in h(\{I, \tilde{I}\})\right),$$

with an extended version of h. Hence, to be consistent with the model before restructuring, generally all transition probabilities have to be carefully translated into a new model. Therefore, models which have a "dimension-less" specification are very popular in the industry — an example for such a model, which particularly does not correspond to a Markovian survival-indicator, is the Gaussian one-factor model.

(c) A general drawback of *all* Markovian survival-indicator models is that one-dimensional marginals are heavily dependent on the specification of Q. Moreover, given an intensity-matrix Q, the construction of finite state space Markovian processes, cf. Rmk. 2.1 or Thm. 2.1, gives a particular interpretation of the joint behavior, which is lost after applying arbitrary marginal transformation. Finally, if there exists a positive rate $q_{h(I),h(J)}$ for two sets with $|J| \leq |I| - 2$, the default-time distribution has a singular component, i.e. joint defaults are possible. As a result, marginal transformation is even more difficult and can introduce undesired effects if performed without care, see e.g. [39], Sec. 5.

A Markovian characterization of the Marshall–Olkin law

The problem described in (a) can easily be resolved by requiring that also all marginal survival-indicator processes $\boldsymbol{Z}_I$ have to be time-homogeneous Markovian. The main result of [8] is the following theorem.

Theorem 4.1. ***(Markovian characterization of MO).*** *The $|I|$-dimensional survival indicator processes $\boldsymbol{Z}_I$ are time-homogeneous Markovian for all subsets $\emptyset \neq I \subseteq [d]$ if and only if $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_d)'$ has a Marshall–Olkin distribution.*

Simulation and Application

There are multiple stochastic models that produce Marshall–Olkin distributed random vectors, which can be used for model specification and simulation. We will consider three models. The seminal interpretation is

an exogenous shock model representation with $2^d - 1$ independent exponential shock arrival-times, one for each subset of components, cf. Eq. (8), see also [11]. An alternative model, in the following denoted as the *Arnold model*, was introduced in [40] and is based on compound sums of exponential random variables. The model can be summarized as follows: Let $\{E_i\}_{i\in\mathbb{N}}$ be an i.i.d. family of exponential random variables with a rate $c = \sum_{\emptyset\neq I\subseteq[d]} \lambda_I$ and $\{X_i\}_{i\in\mathbb{N}}$ a discrete Markov-chain on $\{I : \emptyset \neq I \subseteq [d]\}$, which has a probability of λ_I/c for a transition from an arbitrary state each into I. Then, the random vector $\boldsymbol{\tau}$ is defined by

$$\tau_i := \inf\{t > 0 : i \in X_{N(t)}\},$$

where $N(t) := \sum_{i=1}^{\infty} 1_{\{E_1+\ldots+E_i\leq t\}}$. The latter is closely linked to the classical model for the underlying Markovian survival-indicator as introduced in the previous sections, which is the third model.

Remark 4.3 (Comparison of MO-models). *All three models require a full model specification, i.e.* $2^d - 1$ *parameters, one for every non-empty set of components. The original model has the advantage of being very simple and easy to implement, however, for large dimensions* d *one has to sample* $2^d - 1$ *exponential shocks — therefore the simulation of* n *independent samples has a runtime of the order* $\mathcal{O}(n2^d)$*, see [41]. The Arnold-model is a little more difficult to implement efficiently, see [41], Alg. 3.3 and Alg. 3.4 for details, however the sampling of* n *independent samples has an expected runtime of the order* $\mathcal{O}(2^d + nd^3)$*. The classical Markov simulation is very similar to the Arnold model, with two important differences, which make this approach either more or less desirable. The Arnold model has the property that the distributions of waiting times to the next "event" as well as the random set-variable of "killed" components corresponding to that event are i.i.d. However, if all set-components have already defaulted nothing happens. In the classical Markovian setup the exponential-rates of the waiting times as well as the (random) new state depend on the current state. As a result the initial setup and storage for transition probabilities of the Arnold model is less costly. The price to pay is that not every "event" corresponds to an action. In summary, which of these models is most appropriate depends on the dimension* d*, the number of simulations* n*, and the computational capabilities.*

A possible way to reduce the number of model parameters as well as the computational effort for simulation (with all models) is to assume that all, but a few selected shock-rates equal zero: In [42] the shock model is defined

using only idiosyncratic shocks, a global shock, and a few additional shocks which are chosen on some classification, e.g. industry segment, country, etc., see also [43] for a similar approach.

Considering default modeling, the dynamic properties of the aggregated default counting process and the related loss process have been studied in [44] and [45] under pool homogeneity assumptions and time-inhomogeneous *cluster default-intensities*[11] in dimensions up to $d = 125$. These authors build on the framework of [43], one of the few frameworks allowing for an explicit joint bottom-up and top-down approach, where a Marshall–Olkin bottom up setting corresponds to a generalized Poisson process top-down setup. The GPL model in [44] is one of the first pre-crisis arbitrage-free aggregate loss model to be consistently calibrated to the whole panel of different CDO tranches and maturities for the iTraxx (or CDX) portfolio, including a discussion on tranchelets. For a summary of related models and a calibration study ranging from 2005 to 2009 iTraxx tranches data see [14]. For an example of the calibration of a (time-inhomogeneous) Markovian model to market data, see [46] and [47].

Marshall–Olkin one-factor models

While survival-indicator processes defined on a latent Marshall–Olkin distribution solve the problem described in (a), it is still a model with a large number of parameters, which is in general inefficient to sample. Furthermore, the problem described in (b) is not resolved, as a Marshall–Olkin distributed vector $\boldsymbol{\tau}$ attained with the classical shock model representation as a model tied to a specific dimension d, and certain objects indexed by $\{1, \ldots, d\}$. Assume, that a $d+1$ dimensional Marshall–Olkin distribution $\tilde{\boldsymbol{\tau}}$ exists with $\tilde{\boldsymbol{\tau}}_{[d]} \overset{d}{=} \boldsymbol{\tau}$. Then, for $i \in [d]$, it holds that (cf. Eq. (8))

$$\begin{aligned}\tilde{\tau}_i &= \min\{\tilde{E}^I : i \in I\} \\ &= \min\{\min\{\tilde{E}^I, \tilde{E}^{I\cup\{d+1\}}\} : i \in I \subseteq [d]\},\end{aligned}$$

where $\tilde{E}^I, \emptyset \neq I \subseteq [d+1]$, are the independent exponential random shocks from the shock model representation of $\tilde{\boldsymbol{\tau}}$. In particular, it follows for the rates of $\boldsymbol{\tau}$ that

$$\lambda_I = \tilde{\lambda}_I + \tilde{\lambda}_{I\cup\{d+1\}},\ \emptyset \neq I \subseteq [d],$$

[11] In this model, all defaults are triggered by independent, time-inhomogeneous Poisson processes for subsets (clusters) of entities.

which shows that there are infinitely many possibilities to embed a Marshall–Olkin distribution into a higher dimensional Marshall–Olkin distribution. Summarizing, one can conclude that, in general, for large d the Marshall–Olkin distribution has too many parameters and has no direct intuition for the extension into higher dimensions.

The simplest way to circumvent this issue is to assume that there exists an exchangeable sequence $\tilde{\tau}_i,\ i \in \mathbb{N}$, such that for every finite $\emptyset \neq I \subseteq \mathbb{N}$ the random vector $\tilde{\boldsymbol{\tau}}_I := (\tilde{\tau}_i)_{i\in I}$ has a Marshall–Olkin distribution. Random vectors $\boldsymbol{\tau}$ which have such a construction are said to have an *extendible Marshall–Olkin distribution*. A thorough treatment of these distributions can be found in [20], which also shows that an extendible Marshall–Olkin distribution can be characterized and constructed by a Lévy-subordinator Λ.

Theorem 4.2 (Lévy-frailty construction). *Let $\{\tau_i\}_{i\in\mathbb{N}}$ be an exchangeable sequence on some probability space, such that each finite margin has a Marshall–Olkin distribution. Denote by $\mathcal{H} = \bigcap_{n\geq 1} \sigma(\tau_n, \tau_{n+1}, \ldots)$ the tail-σ-field of $\{\tau_i\}_{i\in\mathbb{N}}$.*

(a) The stochastic process $\Lambda(t) := -\log \mathbb{P}(\tau_1 > t \mid \mathcal{H}), t \geq 0$, is a (possibly killed) Lévy subordinator.

(b) There exists a sequence of i.i.d. unit exponential random variables $\{E_i\}_{i\in\mathbb{N}}$, independent of Λ, such that almost surely

$$\tau_i = \inf\{t > 0 : \Lambda(t) > E_i\},\ i \in \mathbb{N}.$$

(c) Denote by $x \mapsto \psi(x)$ the associated Bernstein function,[12] *i.e. $\exp\{-t\psi(x)\} = \mathbb{E}[\exp\{-x\Lambda(t)\}]$, then*

$$\mathbb{P}(\boldsymbol{\tau} > \boldsymbol{t}) = \prod_{i=1}^{d} \mathrm{e}^{-t_{\pi(i)}(\psi(i)-\psi(i-1))}$$

for each $d \geq 1$ and $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_d)'$, $\boldsymbol{t} \in \mathbb{R}^d_+$ and a permutation π on $[d]$ with $t_{\pi(1)} \geq \ldots \geq t_{\pi(d)}$.

Proof. By De Finetti's Theorem, conditional on $\mathcal{H}$ the sequence $\{\tau_i\}_{i\in\mathbb{N}}$ is i.i.d., with distribution function $1 - \exp\{-\Lambda(t)\}$ for $\Lambda(t) := -\log \mathbb{P}(\tau_1 > t \mid \mathcal{H})$, see [50]. The claim on the variables $\{E_i\}_{i\in\mathbb{N}}$ can be established

[12] A Bernstein function ψ is characterized by a Lévy-triplet (a, b, ν) for $a, b \geq 0$ and a Lévy-measure ν on $(0, \infty)$ fulfilling the integrability condition $\int_{(0,\infty)} 1 \wedge v\nu(dv) < \infty$, where $\psi(x) = a1_{(0,\infty)}(x) + bx + \int_{(0,\infty)}(1 - \mathrm{e}^{-xv})\nu(dv), x \geq 0$, see [48], [49].

with a modified distribution function, see [51], Prop. 2.1. Furthermore, the law of $\{\Lambda(t)\}_{t\geq 0}$ is almost surely uniquely determined by $\mathcal{H}$, and by [41], Chapter 3.3, it is a (possibly killed) Lévy subordinator with the claimed properties. □

The alternative stochastic model of extendible Marshall–Olkin distributions via the so-called Lévy-frailty construction in Thm. 4.2 has the advantage of being a *De Finetti model* for extendible sequences, which renders the approach independent of the dimension d. This solves not only the problem described in (b), but also provides an alternative simulation strategy, see [8] for a detailed account. The alternative simulation strategy has the advantage that its runtime scales linearly with increasing dimension, which makes it particularly interesting for large d. The approach comes with the drawback that a simulation bias is introduced as we can only sample the random walk corresponding to some embedding of Λ on a discrete time-grid. This bias, however, can be controlled through the step size of the discrete time-grid.

In the following we present five examples of Lévy-subordinators which can be used to define parametric one-factor Marshall–Olkin distributions.

Example 4.1 (Linear drift). *Let $\Lambda(t) = bt, t \geq 0$ for some $b > 0$, then $\boldsymbol{\tau}$ corresponds to d independent exponentially distributed random variables with common rate b. A simple extension can be attained assuming a "global shock" $E \sim \mathrm{Exp}(a)$, $a > 0$, which "kills" all entities. This corresponds to a (killed) Lévy-subordinator $\Lambda(t) = bt + \infty \cdot 1_{\{E\leq t\}}, t \geq 0$ with the convention $0 \cdot \infty = 0$. The corresponding Bernstein-function is $\psi(x) = a1_{(0,\infty)}(x) + bx$.*

This model is, e.g., implicitly used in [52]. A "global shock" can analogously be introduced in every Lévy-frailty model by assuming that Λ is "killed" — that is, sent to the absorbing state ∞ — at a rate $a > 0$, i.e. there exists an independent exponential random variable E with rate a and we assume that $\Lambda(t) = \infty$ for $t > E$. The corresponding new Bernstein-function can be attained by adding the term $a1_{(0,\infty)}(x)$ to the old one.

Example 4.2 (Compound Poisson subordinator). *Let $\Lambda(t) = bt + \sum_{k=1}^{N(t)} J_k$ for independent N and $\{J_k\}_{k\in\mathbb{N}}$, where the former is a classical Poisson-process with rate $\lambda > 0$ and the latter an i.i.d. family of random variables on $(0, \infty)$. The corresponding Bernstein-function is $\psi(x) = bx + \lambda(1 - \mathcal{L}(x; J_1))$, where $\mathcal{L}(x; J_1)$ is the Laplace-transformation corresponding to J_1.*

For a compound Poisson subordinator, defined as above, the number of jumps in the time-intervals $(0, t_1], (t_1, t_2], \ldots$ are independent and $\mathrm{Poi}(\lambda(t_k - t_{k-1}))$ distributed on $\mathbb{N}_0$, respectively, and the jth jump-size is J_j.

Example 4.3 (Gamma subordinator). *Let Λ have a Bernstein function of the form $\psi(x) = \alpha \ln(1 + x/\beta)$ for $\alpha, \beta > 0$. The corresponding increments $\Lambda(s) - \Lambda(t)$ are Gamma-distributed and can easily be sampled, see e.g. [41], Alg. 6.5 and Alg. 6.6, pp. 242–243.*

Example 4.4 (Inverse-Gaussian subordinator). *Let Λ have a Bernstein function of the form $\psi(x) = \beta(\sqrt{2x + \eta^2} - \eta)$ for $\beta, \eta > 0$. The corresponding increments $\Lambda(s) - \Lambda(t)$ are Inverse-Gaussian distributed and can easily be sampled, see e.g. [41], Alg. 6.10, p. 245.*

Example 4.5 (Stable subordinator). *Let Λ have a Bernstein function of the form $\psi(x) = x^\alpha$ for some $1 \geq \alpha > 0$. Then the increments $\Lambda(s) - \Lambda(t)$ belong to the class of stable distributions and can be sampled, see e.g. [41], Alg. 6.11, p. 246.*

Marshall–Olkin multi-factor models

The Lévy-frailty model has the serious drawback of being a one-factor model. This implies not only homogeneity with respect to marginal distributions, but also an exchangeable dependence structure. However, we can exploit that independent Lévy subordinators form a cone and we can consider the extended Lévy-frailty model, where $\boldsymbol{\tau}$ is defined by

$$\tau_i := \inf\{t > 0 : \Lambda_i(t) > E_i\},\ i \in [d], \tag{14}$$

where $\Lambda_i, i \in [d]$, are Lévy subordinators from the cone spanned from independent Lévy subordinators $\Upsilon_1, \ldots, \Upsilon_n$ and $E_1, \ldots, E_d$ are i.i.d. unit exponentials, which are independent thereof. In the following, a result of [26] regarding this model is presented. Assume that $\boldsymbol{\Upsilon}$ is an n-dimensional vector of independent Lévy subordinators corresponding to Bernstein functions $\hat{\psi}_1, \ldots, \hat{\psi}_n$ and $\Theta = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_d) \in \mathbb{R}_+^{n \times d}$ is a matrix with non-negative entries. Define the process $\boldsymbol{\Lambda}$ by $\Lambda_i := \boldsymbol{\theta}_i' \boldsymbol{\Upsilon}, i \in [d]$.

Theorem 4.3. *Let $\boldsymbol{t} \geq 0$ and $\pi \in \mathcal{S}_d$ be a permutation with $t_{\pi(1)} \geq \ldots \geq t_{\pi(d)}$ and let $\boldsymbol{\tau}$ be defined as in Eq. (14). Then*

$$\mathbb{P}(\boldsymbol{\tau} > \boldsymbol{t}) = \exp\left\{ -\sum_{i=1}^{d} t_{\pi(i)} \sum_{k=1}^{n} \hat{\psi}_k \left(\sum_{j=1}^{i} \Theta_{k,\pi(j)} \right) - \hat{\psi}_k \left(\sum_{j=1}^{i-1} \Theta_{k,\pi(j)} \right) \right\}.$$

Furthermore, $\boldsymbol{\tau}$ has a Marshall–Olkin distribution.

Proof. See [26]. □

A slightly simplified extension with $n = 1$ has the interpretation of allowing inhomogeneous trigger rates in the original Lévy-frailty model, cf. [53]. Furthermore, a useful alternative representation of the vector in Thm. 4.3 can be attained as follows, cf. [41], Sec. 3.3.4: Let $\boldsymbol{\tau}^{(k)}$ be independent random vectors corresponding to Lévy-frailty models with inhomogeneous trigger rates $\boldsymbol{\theta}_k$ and trigger processes $\hat{\psi}_k$ for $k = 1, \ldots, n$. Then $\boldsymbol{\tau}$ has the survival function in Thm. 4.3, where $\boldsymbol{\tau}$ is defined by

$$\tau_i := \min\{\tau_i^{(k)} : k \in [n]\},\ i \in [d].$$

Remark 4.4 (Constructing the full Marshall–Olkin class). The multi-factor Lévy-frailty construction is general enough to comprise the full family of Marshall–Olkin distributions. To this end, we use $m = 2^d - 1$ independent killed subordinators $\Upsilon^{(I)}(t) := \infty \mathbf{1}_{\{E_I \le t\}}$ and $\Lambda^{(k)}(t) := \sum_{I:k\in I} \hat{\Lambda}^{(I)}(t)$, which is basically just a complicated way of writing the original Marshall–Olkin shock model, cf. Eq. (8). This construction is not unique in the class of Lévy-frailty models and provides an alternative proof of [54], Thm. 4.2.

Closely related, a hierarchical and h-extendible Marshall–Olkin law is constructed in [55] and [56]. The idea is to group the components according to some (economic) criterion (e.g., geographic region, industry segment, etc.). In the simplest case one has only one classification criterion, say for illustration purposes the industry segment, and each component is affected by a global and an industry specific factor. With respect to the factor model described in Thm. 4.3, assume that the components can be separated into J industry segments. Let $\Upsilon_1, \ldots, \Upsilon_J$ be independent Lévy subordinators, each corresponding to a specific segment. Furthermore, let Υ_0 be another independent Lévy subordinator corresponding to a global factor affecting all components. For component $i \in [d]$ which is in segment k, an individual trigger-processes Λ_i is defined using the weights $\boldsymbol{\theta}_i$ which are for $\alpha, \beta_k > 0$ defined by

$$\boldsymbol{\theta}_i = (\alpha, \underbrace{0, \ldots, 0}_{k-1 \text{ times}}, \beta_k, \underbrace{0, \ldots, 0}_{(J-k) \text{ times}})' \in \mathbb{R}_+^{J+1}$$

and by

$$\Lambda_i = \boldsymbol{\theta}_i' \boldsymbol{\Upsilon} = \sum_{k=0}^{n} \Theta_{k,i} \Upsilon_k.$$

This model is said to be h-extendible with two levels of hierarchy — meaning that there exists a σ-algebra $\mathcal{G}_0$ such that, conditional on this $\mathcal{G}_0$, the vector of default-times separates into independent groups and there exist group specific σ-algebras $\mathcal{G}_k$ such that the marginal group vectors of default-times are conditionally i.i.d., see [56]. For more levels of hierarchy, say one wants an additional regional classification, the model can be extended easily.

This model specification solves the problems (a), (b), and partially also (c), which were described at the beginning of this section:

(a) As shown in the previous paragraph, Marshall–Olkin distributions have the unique property that all marginal survival indicators are time-homogeneous Markovian. Therefore, simulation-studies on subportfolios can be performed efficiently using lower dimensional Markovian processes.
(b) The hierarchical construction gives an intuitive way to deal with portfolio restructuring. In case of a downsize, we can simply use the reduced model as each of the factors should be chosen in a way that they are (mostly) independent of the portfolio. If an additional component has to be modeled, one only has to specify factor-loadings corresponding to the "risk" regarding to each factor.
(c) Even though this model setup is not a copula ansatz, the factor approach offers a schematic picture of the inner- and outer-group dependence between components. In particular, it follows that the dependence, measured with the upper-tail dependence coefficient, between two components of the same group is higher than that of two components of different groups, see [55] for a similar result with temporal-, instead of spatial scaling of the underlying subordinators. However, the complete dependence structure, in form of the underlying copula, as well as the marginal distributions, are influenced by the specific weights. If only marginal distributions should be altered, this is possible by using a component specific factor. However, the choice of the marginal is restricted to the class of exponential distributions (as otherwise the Markov property is lost) and the minimal marginal rate is determined by the remaining weights.

In default modeling, the historical data is rarely substantial enough to perform goodness-of-fit tests for the chosen copula. Therefore, a good qualitative understanding of the schematic dependence is crucial. A slight modification of this model, which then partially solves (c), can be specified, if the loadings are assumed to be constant, e.g. $\alpha_i = \beta_i = 1$, and the group-components of the resulting vector are scaled with group specific scalar values to attain a group specific exponential-rate.

In Fig. 2, most of the distributional classes discussed in this paper are summarized in a schematic picture.

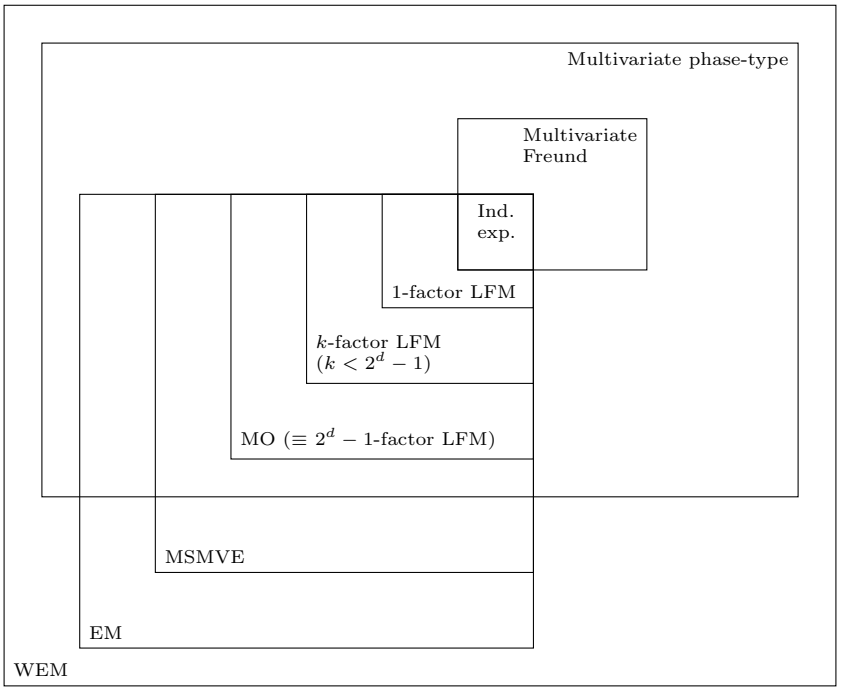


Fig. 2. Venn-diagram of (selected) multivariate exponential, Phase-type distributions, and distributions fulfilling the WEM-property. See Chap. 2.2 as well as [41], [37], [35], [13] for details.

4.3. *Case study: Iteration bias for selected multivariate distributions*

In Thm. 3.1 it was highlighted that iterating bivariate (non-comonotonic) Gaussian-, Clayton-, or Frank-coupled exponential margins "kills" dependence asymptotically. In the first numerical case study, cf. Sec. 3.2, it was demonstrated that probabilities for "survival-of-all" events can divert significantly if "terminal one-shot" are compared to "terminal iterated" laws. Only distributions fulfilling the weak exponential minima property have the property that "survival-of-all" events have the same probability under the "terminal one-shot" and "terminal iterated" law.

In Thm. 4.1 it was shown that the "terminal one-shot" and the "terminal iterated" law are equal if and only if it is a Marshall–Olkin distribution. The purpose of this section is to underscore this statement with a second numerical case study.

The model

Before numerical results are presented, it is specified mathematically what was referred to loosely as the "terminal one-shot" and "terminal iterated" law. It is assumed that the multivariate probability and survival distribution of "mixed default/survival" events are replaced by corresponding events using discretely iterated survival indicators, i.e. instead of

$$\mathbb{P}\left(\left(\bigcap_{i\in I}\{\tau_i > k_i\Delta\}\right)\cap\left(\bigcap_{i\notin I}\{\tau_i \leq k_i\Delta\}\right)\right)$$

we consider the probabilities

$$\mathbb{P}\left(\left(\bigcap_{i\in I}\{\tilde{Z}_i^{(\Delta)}(k_i)=1\}\right)\cap\left(\bigcap_{i\notin I}\{\tilde{Z}_i^{(\Delta)}(k_i)=0\}\right)\right),$$

where $\tilde{\boldsymbol{Z}}^{(\Delta)}$ is a (discrete-time) Markov-chain with i.i.d. multiplicative increments that are fully determined by

$$\tilde{\boldsymbol{Z}}^{(\Delta)}(1) \stackrel{d}{=} (1_{\{\tau_1>\Delta\}},\ldots,1_{\{\tau_d>\Delta\}})'.$$

This approach corresponds to the widespread industry-practice of defining a default distribution and iterating (multiplicative) i.i.d. increments of the corresponding survival-indicator for the step-size Δ through a discrete time grid up to the final horizon $T = N\Delta$.

The case study

It is assumed that $\Delta = 1$, $k_1 = 10$, and $k_2 \in \{5, 10\}$ for the event $\{\tau_1 > k_1\Delta, \tau_2 > k_2\Delta\}$ and the following distributions with common marginal rate $\lambda > 0$ are considered:

- *Marshall–Olkin:* A bivariate exchangeable Marshall–Olkin distribution with copula-parameter $\alpha_{MO} \in [0, 1]$, in the exchangeable Cuadras-Augé parameterization.
- *Gumbel:* A bivariate Gumbel distribution with parameter $\theta_{Gu} \in [1, \infty]$.
- *Clayton:* An exchangeable Clayton-coupled exponential distribution with parameter $\theta_{Cl} \geq -1$.
- *Frank:* An exchangeable Frank-coupled exponential distribution with parameter $\theta_{Fr} \in \mathbb{R}$.
- *Gaussian:* An exchangeable Gaussian-coupled exponential distribution with parameter $\rho_{Ga} \in [-1, 1]$.
- *t:* An exchangeable t-coupled exponential distribution for $\nu = 3$ degrees of freedom, parameter $\rho_t \in [-1, 1]$.

The marginal rates are assumed to be $\lambda_{SG} = 4.5\%$ (speculative grade) and the copula parameters are calibrated such that Kendall's τ equals 50%, see [30], pp. 260–261 for an overview on the Gumbel, Clayton, and Frank copula. Additionally, the following distributions are considered:

- *Freund:* An exchangeable Freund distribution with rates $\lambda_1 = \lambda_2 = \lambda_{SG}$ and $\eta_1 = \eta_2 = 3\lambda_{SG}$. The corresponding marginal distributions are not exponential and the resulting Kendall's τ is not set up to equal 50%.
- *Independent:* Two independent exponential random variables with common marginal rate $\lambda_{SG} > 0$. The independence copula is contained in all previously mentioned copulas families and is included as a reference point in this analysis.

In Tables 3 and 4 the results for both events can be observed. As expected, apart from Marshall–Olkin, Gumbel, Freund, and the independence copula, all copulas yield sizable differences for the "survival-of-all" event. For the "mixed default/survival" event only the Marshall–Olkin distribution and the independence copula yield equal "terminal one-shot" and "terminal iterated" probabilities. The effect is particularly strong for the Clayton- and Frank copula, where the "terminal iterated" probabilities are almost at the level of the independence copula.

Table 3. Comparison of "terminal one-shot" and "terminal iterated" survival probabilities for $k_1 = 10$, $k_2 = 10$, and $\Delta = 1y$ (survival-of-all case).

Copula	Exact law	Iterated law	%Diff
Marshall–Olkin	0.5488	0.5488	0%
Gumbel	0.5292	0.5292	0%
Clayton	0.5051	0.4220	19.71%
Frank	0.5299	0.4388	20.77%
Gaussian	0.5205	0.4788	8.72%
t	0.5219	0.5053	3.28%
Independent	0.4066	0.4066	0%
Freund	0.4066	0.4066	0%

Table 4. Comparison "terminal one-shot" and "terminal iterated" survival probabilities for $k_1 = 10$, $k_2 = 5$, and $\Delta = 1y$ (mixed default-survival case).

Copula	Exact law	Iterated law	%Diff
Marshall–Olkin	0.5916	0.5916	0%
Gumbel	0.6046	0.5809	4.09%
Clayton	0.5747	0.5187	10.79%
Frank	0.5965	0.5289	12.77%
Gaussian	0.5956	0.5525	7.8%
t	0.5956	0.5676	4.93%
Independent	0.5092	0.5092	0%
Freund	0.4885	0.5042	−3.13%

5. Conclusions

The problem of simulating the survival-indicator process on a discrete time-grid along with the remaining risk-factors has been investigated. It has been argued that, especially for high dimensions, good candidates for consistent and feasible joint simulations are continuous-time, time-homogeneous Markovian survival-indicators processes. In particular, the market practice of modeling the survival-indicator process as a discrete-time Markov chain with i.i.d. multiplicative increments, corresponding to a step distribution which is based on a copula-based ansatz, has been analyzed, criticized, and rectified. It has been shown theoretically and demonstrated with numerical examples that if we are concerned only with the "survival-of-all" event, then in order for "terminal one-shot" and "terminal iterated" probabilities to coincide, the multivariate default times distribution must fulfill the weak exponential minima property. In particular, this property is fulfilled for

exponential margins with a survival copula of extreme-value kind. If we are concerned with more general "mixed default/survival" events, this consistency is only achieved by Marshall–Olkin distributions. A special emphasis is on warning practitioners who are iterating Gaussian-coupled exponential distributions, which fulfill neither the weak exponential minima property nor do they belong to the class of Marshall–Olkin distributions. Indeed, since these distributions lie in the domain of attraction of the independence copula, iterating them completely "kills" dependence asymptotically, when the number of iterations increases.

Appendix A. Alternative construction of Markovian processes

An alternative construction of continuous-time, time-homogeneous Markovian processes on finite state spaces is presented. The construction is a variation of the classical construction, where (state specifically) jumps are constructed with exponential waiting times and independent new (random) states, cf. Rmk. 2.1.

Theorem A.1. *Let Q be an intensity matrix of a continuous-time, time-homogeneous Markovian process on a finite state space S (which is w.l.o.g. assumed to be $\{1, \dots, |S|\}$). Consider a process Z which is constructed as follows:*

(i) Let X_0 be the (possibly random) initial state, i.e. define $Z(0) := X_0$.

(ii) Assume that Z jumped $k \in \mathbb{N}_0$ times and define the time of the kth jump by T_k (for $k = 0$ we define $T_0 := 0$). Furthermore, assume that $Z(T_k) = i \in S$.

(a) For $j \in S \backslash \{i\}$ let $E^j_{k+1} \sim \mathrm{Exp}(q_{ij})$ be independent exponential random variables and define $E^i_{k+1} := \infty$. Assume additionally that $\boldsymbol{E}_{k+1}$, conditional on $Z(T_k)$, is independent of $\{\boldsymbol{E}_l : l \leq k\}$, $\boldsymbol{E}_{k+1} := (E^1_{k+1}, \dots, E^d_{k+1})'$.

(b) Define $T_{k+1} := T_k + \min_{j \in S} E^j_{k+1}$ and $Z(t) := i \ \forall t \in (T_k, T_{k+1})$.

(c) Define $Z(T_{k+1}) := \mathrm{argmin}_{j \in S} E^j_{k+1}$.

(iii) Repeat (ii) either infinitely often or until an absorbing state is reached.

Then the process Z is time-homogeneous Markovian with intensity-matrix Q.

Proof. For $k \geq 0$ and $i \in S$ define $\mathbb{P}_k(\cdot) = \mathbb{P}(\cdot \mid Z(T_k) = i)$. It suffices to show that for every $k \geq 0$ and $i \in S$ the following three conditions hold, as this implies the classical construction:

(I) $\mathbb{P}_k(\min_{j\in S} E^j_{k+1} > t) = \exp\{q_{ii}t\}\ \forall t > 0$.
(II) $\mathbb{P}_k(\operatorname{argmin}_{j\in S\setminus\{i\}} E^j_{k+1} = j^\star) = -q_{ij^\star}/q_{ii}\ \forall j^\star \in S\setminus\{i\}$.
(III) The random variables $\min_{j\in S} E^j_{k+1}$ and $\operatorname{argmin}_{j\in S\setminus\{i\}} E^j_{k+1}$ are independent conditional on $\{Z(T_k) = i\}$.

Condition (I) holds as the minimum of independent exponential random variables is again exponential with the rate corresponding the sum of all rates. In this particular case this implies, conditional on $\{Z(T_k) = i\}$, $\min_{j\in S} E^j_{k+1}$ is exponential with rate

$$\sum_{j\in S\setminus\{i\}} q_{ij} \overset{(\star)}{=} -q_{ii},$$

where $(\star)$ follows because Q is an intensity matrix.

The following calculation shows that condition (II) hold:

$$\begin{aligned}\mathbb{P}_k\left(\operatorname*{argmin}_{j\in S\setminus\{i\}} E^j_{k+1} = j^\star\right) &= \mathbb{P}_k\left(E^{j^\star}_{k+1} < \min_{j\in S\setminus\{i,j^\star\}} E^j_{k+1}\right)\\ &\overset{(\star)}{=} \mathbb{E}_k\left[\exp\left\{-E^{j^\star}_{k+1}\sum_{j\in S\setminus\{i,j^\star\}} q_{ij}\right\}\right]\\ &\overset{(\dagger)}{=} \frac{q_{ij^\star}}{q_{ij^\star} + \sum_{j\in S\setminus\{i,j^\star\}} q_{ij}} \overset{(\ddagger)}{=} -\frac{q_{ij^\star}}{q_{ii}},\end{aligned}$$

where $(\star)$ follows using the tower property conditioning on $E^{j^\star}_{k+1}$, $(\dagger)$ follows with the Laplace-transform of the exponential distribution, and $(\ddagger)$ follows using that Q has vanishing row sums.

Finally, the following calculate proves that condition (III) holds:

$$\begin{aligned}&\mathbb{P}_k\left(\min_{j\in S\setminus\{i\}} E^j_{k+1} > t, \operatorname*{argmin}_{j\in S\setminus\{i\}} E^j_{k+1} = j^\star\right)\\ &\quad\overset{(\star)}{=} \mathbb{E}_k\left[1_{\{E^{j^\star}_{k+1}>t\}}\exp\left\{-E^{j^\star}_{k+1}\sum_{j\in S\setminus\{i,j^\star\}} q_{ij}\right\}\right]\\ &\quad\overset{(\dagger)}{=} -\frac{q_{ij^\star}}{q_{ii}}\exp\{q_{ii}t\},\end{aligned}$$

where $(\star)$ follows using the tower property conditioning on $E^{j^\star}_{k+1}$ and $(\dagger)$ follows using that for an exponential random variable E with rate $\eta > 0$

one has for $t, x > 0$

$$\begin{aligned}\mathbb{E}\left[1_{\{E>t\}} \exp\{-xE\}\right] &= \int_t^\infty \eta \exp\{-(x+\eta)v\} dv \\ &= \frac{\eta}{\eta + x} \exp\{-(x+\eta)t\}.\end{aligned}$$

□

Acknowledgments

We thank the anonymous referees and the handling editor for useful remarks. Valuable help by Annette Wenninger to organize the references is appreciated.

References

1. C. Albanese, T. Bellaj, G. Gimonet and G. Pietronero, Coherent global market simulations and securitization measures for counterparty credit risk, *Quantitative Finance* **11**, 1–20 (2011).
2. D. Brigo, A. Capponi and A. Pallavicini, Arbitrage-free bilateral counterparty risk valuation under collateralization and application to credit default swaps, *Mathematical Finance* **24**, 125–146 (2014).
3. D. Brigo, A. Pallavicini and M. Morini, *Counterparty Credit Risk, Collateral and Funding: With Pricing Cases for all Asset Classes* (Wiley, 2013).
4. Bank of International Settlement, Fundamental review of the trading book, consulting document, annex 4 (May 2012), Available at BIS.org.
5. T. R. Bielecki, S. Crépey and A. Herbertsson, Markov chain models of portfolio credit risk, in *The Oxford Handbook of Credit Derivatives*, eds. A. Lipton and A. Rennie (Oxford University Press, 2011) pp. 327–382.
6. A. Herbertsson and H. Rootzen, Pricing kth-to-default swaps under default contagion: the matrix-analytic approach, *Journal of Computational Finance* **12**, 49–78 (2008).
7. D. Brigo and K. Chourdakis, Consistent single-and multi-step sampling of multivariate arrival times: a characterization of self-chaining copulas, Available at SSRN: https://ssrn.com/abstract=2047474 or http://dx.doi.org/10.2139/ssrn.2047474 (April, 2012).
8. D. Brigo, J.-F. Mai and M. Scherer, Markov multi-variate survival indicators for default simulation as a new characterization of the Marshall–Olkin law, *Statistics & Probability Letters* **114**, 60–66 (2016).

9. P. Brémaud, *Markov chains*, Texts in applied mathematics, Vol. 31 (Springer, 1999).
10. P. Billingsley, *Probability and measure*, Wiley Series in Probability and Statistics, 3 edn. (Wiley, 1995).
11. A. W. Marshall and I. Olkin, A multivariate exponential distribution, *Journal of the American Statistical Association* **62**, 30–44 (March 1967).
12. H. Joe, *Multivariate Models and Multivariate Dependence Concepts* (CRC Press, 1997).
13. J. D. Esary and A. W. Marshall, Multivariate distributions with exponential minimums, *The Annals of Statistics* **2**, 84–96 (1974).
14. D. Brigo, A. Pallavicini and R. Torresetti, *Credit Models and the Crisis: A Journey into CDOs, Copulas, Correlations and Dynamic Models* (Wiley, 2010).
15. E. J. Gumbel, Bivariate exponential distributions, *Journal of the American Statistical Association* **55**, 698–707 (December 1960).
16. S. Kotz, N. Balakrishnan and N. L. Johnson, *Continuous multivariate distributions, models and applications (volume 1)* (John Wiley & Sons, 2000).
17. N. Ebrahimi and H. Zahedi, Testing for bivariate gumbel against bivariate new better than used in expectation, *Communications in Statistics - Theory and Methods* **18**, 1357–1371 (1989).
18. K. M. Nair and N. U. Nair, On characterizing the bivariate exponential and geometric distributions, *Annals of the institute of Statistical Mathematics* **40**, 267–271 (1988).
19. G. Gudendorf and J. Segers, Extreme-value copulas, in *Copula Theory and Its Applications*, eds. P. Jaworski, F. Durante, W. K. Härdle and T. Rychlik, Lecture Notes in Statistics, Vol. 198 (Springer, Berlin, Heidelberg, 2010) pp. 127–145.
20. J.-F. Mai, Extendibility of Marshall–Olkin distributions via Lévy subordinators and an application to portfolio credit risk, dissertation, Technische Universität München 2010. Available at http://mediatum.ub.tum.de?id=969547.
21. J.-C. Lu and G. K. Bhattacharyya, Inference procedures for a bivariate exponential model of Gumbel based on life test of component and system, *Journal of Statistical Planning and Inference* **27**, 383–396 (March 1991).
22. C. Genest and L.-P. Rivest, A characterization of Gumbel's family of extreme value distributions, *Statistics & Probability Letters* **8**, 207 (1989).

23. D. Brigo and K. Chourdakis, Counterparty risk for credit default swaps: impact of spread volatility and default correlation, *International Journal of Theoretical and Applied Finance* **12**, 1007–1026 (2009).
24. M. Morini, *Understanding and Managing Model Risk: A Practical Guide for Quants, Traders and Validators* (John Wiley & Sons, 2011).
25. J.-F. Mai and M. Scherer, Characterization of extendible distributions with exponential minima via processes that are infinitely divisible with respect to time, *Extremes* **17**, 77–95 (2014).
26. J.-F. Mai, Multivariate exponential distributions with latent factor structure and related topics, habilitation, Technische Universität München 2014.
27. D. K. Dey and J. Yan (eds.), *Extreme value modeling and risk analysis: methods and applications* (Chapman and Hall/CRC, 2015).
28. C. Fang, *Terminal vs. iterated statistics of multivariate default times*, UROP project report, Imperial College of London (2016), Supervised by Brigo, D.
29. M. Sibuya, Bivariate extreme statistics, *Annals of the Institute of Statistical Mathematics* **11**, 195–210 (1960).
30. P. Embrechts, R. Frey and A. McNeil, *Quantitative risk management*, Princeton Series in Finance, Vol. 5, revised edn. (Princeton University Press, 2015).
31. R. A. Jarrow and F. Yu, Counterparty risk and the pricing of defaultable securities, *The Journal of Finance* **56**, 1765–1799 (2001).
32. F. Yu, Correlated defaults in intensity-based models, *Mathematical Finance* **17**, 155–173 (2007).
33. I. Norros, A compensator representation of multivariate life length distributions, with applications, *Scandinavian Journal of Statistics* **13**, 99–112 (1986).
34. M. Shaked and G. J. Shanthikumar, The multivariate hazard construction, *Stochastic Processes and Their Applications* **24**, 241–258 (1987).
35. J. E. Freund, A bivariate extension of the exponential distribution, *Journal of the American Statistical Association* **56**, 971–977 (December 1961).
36. H. W. Block and A. P. Basu, A continuous, bivariate exponential extension, *Journal of the American Statistical Association* **69**, 1031–1037 (1974).
37. D. Assaf, N. A. Langberg, T. H. Savits and M. Shaked, Multivariate phase-type distributions, *Operations Research* **32**, 688–702 (1984).

38. J. Cai and H. Li, Conditional tail expectations for multivariate phase-type distributions, *Journal of Applied Probability* **42**, 810–825 (September 2005).
39. J.-F. Mai and M. Scherer, What makes dependence modeling challenging? Pitfalls and ways to circumvent them, *Statistics & Risk Modeling* **11**, 287–306 (2013).
40. B. C. Arnold, A characterization of the exponential distribution by multivariate geometric compounding, *Sankhyā: The Indian Journal of Statistics, Series A* **37**, 164–173 (1975).
41. J.-F. Mai and M. Scherer, *Simulating copulas: stochastic models, sampling algorithms and applications*, Series in Quantitative Finance, Vol. 6, 2 edn. (World Scientific, 2017).
42. K. Giesecke, A simple exponential model for dependent defaults, *The Journal of Fixed Income* **13**, 74–83 (2003).
43. F. Lindskog and A. J. McNeil, Common Poisson shock models: applications to insurance and credit risk modelling, *Astin Bulletin* **33**, 209–238 (2003).
44. D. Brigo, A. Pallavicini and R. Torresetti, Calibration of CDO tranches with the dynamical Generalized-Poisson Loss model, *Risk Magazine* (June 2007).
45. D. Brigo, A. Pallavicini and R. Torresetti, Cluster-based extension of the generalized Poisson loss dynamics and consistency with single names, *International Journal of Theoretical and Applied Finance* **10**, 607–632 (2007), Also in: Liption, A., Rennie, A. (Eds.), Credit Correlation – Life after Copulas, World Scientific, 2007.
46. T. R. Bielecki, A. Cousin, S. Crépey and A. Herbertsson, A bottom-up dynamic model of portfolio credit risk; part i: Markov copula perspective, Available at SSRN: https://ssrn.com/abstract=1844574 or http://dx.doi.org/10.2139/ssrn.1844574 (March, 2013).
47. T. R. Bielecki, A. Cousin, S. Crépey and A. Herbertsson, A bottom-up dynamic model of portfolio credit risk: part ii: common-shock interpretation, calibration and hedging issues, Available at SSRN: https://ssrn.com/abstract=2245130 or http://dx.doi.org/10.2139/ssrn.2245130 (March, 2013).
48. R. L. Schilling, R. Song and Z. Vondracek, *Bernstein functions*, De Gruyter Studies in Mathematics, Vol. 37 (De Gruyter, Berlin, 2010).
49. C. Berg, J. P. R. Christensen and P. Ressel, *Harmonic analysis on semigroups*, Graduate Texts in Mathematics, Vol. 100 (Springer New York, 1984).

50. D. J. Aldous, Exchangeability and related topics, in *École d'Été de Probabilités de Saint-Flour XIII — 1983*, ed. P. L. Hennequin (Springer Berlin Heidelberg, Berlin, Heidelberg, 1985) pp. 1–198.
51. L. Rüschendorf, On the distributional transform, Sklar's theorem, and the empirical copula process, *Journal of Statistical Planning and Inference* **139**, 3921–3927 (2009), Special Issue: The 8th Tartu Conference on Multivariate Statistics & The 6th Conference on Multivariate Distributions with Fixed Marginals.
52. X. Burtschell, J. Gregory and J.-P. Laurent, A comparative analysis of CDO pricing models under the factor copula framework, in *The Definitive Guide to CDOs*, ed. G. Meissner (Risk Books, 2009) pp. 389–427.
53. J. Engel, M. Scherer and L. Spiegelberg, One-factor lévy-frailty copulas with inhomogeneous trigger rates, in *Soft Methods for Data Science*, eds. M. B. Ferraro, P. Giordani, B. Vantaggi, M. Gagolewski, M. Ángeles Gil, P. Grzegorzewski and O. Hryniewicz (Springer International Publishing, 2017) pp. 205–212.
54. Y. Sun, R. Mendoza-Arriaga and V. Linetsky, Marshall–Olkin distributions, subordinators, efficient simulation, and applications to credit risk, Available at SSRN: https://ssrn.com/abstract=1702087 or http://dx.doi.org/10.2139/ssrn.1702087 (December, 2016).
55. J.-F. Mai and M. Scherer, Reparameterizing Marshall–Olkin copulas with applications to sampling, *Journal of Statistical Computation and Simulation* **81**, 59–78 (2011).
56. J.-F. Mai and M. Scherer, H-extendible copulas, *Journal of Multivariate Analysis* **110**, 151–160 (2012).