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.

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 short-fall 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 $E_{k+1} \sim \mathrm{Exp}(-q_{ii})$ 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 $S \setminus \{i\}$ 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 τ be a default-vector with corresponding time-homogeneous Markovian survival-process 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

$$A_{\pi,K} := \left\{ (I_1, \ldots, I_K) : \pi([k]) \subseteq I_k, I_k \subseteq I_{k+1} \; \forall k = 1, \ldots, K \right\},$$

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

$$\begin{aligned}
&\mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I + t + v \mid \boldsymbol{\tau}_J > t + v, \tau_{[d] \setminus J} \leq t + v) \\
&= \mathbb{P}(\boldsymbol{\tau}_I > \boldsymbol{s}_I + t \mid \boldsymbol{\tau}_J > t, \tau_{[d] \setminus J} \leq t).
\end{aligned}$$

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 $E^J_{k+1} \sim \text{Exp}(q_{h(I),h(I \setminus J)})$ 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 $D_{k+1} := \text{argmin}_{\emptyset \neq J \subseteq I} E^J_{k+1}$. Furthermore, define $\boldsymbol{Z}(t) := h(I)$ $\forall t \in (T_k, T_{k+1})$ and $\boldsymbol{Z}(\overline{T}_{k+1}) := h(I \setminus D_{k+1})$.

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:

"terminal one-shot survival probability up to $t_1 + \ldots + t_N$"
$$= \mathbb{P}(\boldsymbol{\tau} > t_1 + \ldots + t_N) = \mathbb{P}(\boldsymbol{\tau}^{(1)} > t_1) \cdot \ldots \cdot \mathbb{P}(\boldsymbol{\tau}^{(N)} > t_N)$$
$$= \text{"terminal iterated survival probability with steps } t_1, \ldots, t_N,\text{"}$$

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 $j \in \mathbb{N}$

$$\boldsymbol{Z}^{(\Delta)}(j+1) \mid \{\boldsymbol{Z}^{(\Delta)}(j) = \boldsymbol{1}\} := 1_{\{\varsigma_{j+1} > \Delta\}},$$

for independent and identically distributed $\boldsymbol{\zeta}_{j+1} \sim C_\Phi(\rho) \oplus (\bar{F}_1, \dots, \bar{F}_d)$, 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]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

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 in-
corporates 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 $\lambda_{\text{IG}} = 1\%$ and $\lambda_{\text{SG}} =
4.5\%$, 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.
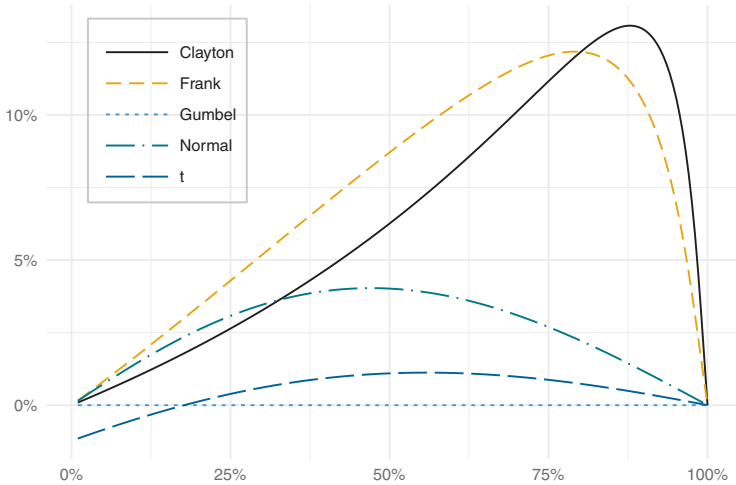
Fig. 1.   Relative deviation of p_T and p_Δ^N in % vs. Kendalls's τ for $T = 5y$, $\lambda_i = 4.5\%$, $i = 1, 2$, $N = 10$, and 3 degrees of freedom for the t-distribution, see [28].

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

$$M^{-1} Q M = \mathrm{diag}(-(\lambda_1 + \lambda_2), -\eta_2, -\eta_1, 0),$$

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

$$P_{(1,1),(1,1)}[t] = e^{-(\lambda_1+\lambda_2)\,t},$$

$$P_{(1,1),(0,1)}[t] = \frac{\lambda_1}{\lambda_1+\lambda_2-\eta_2}\left(e^{-\eta_2\,t} - e^{-(\lambda_1+\lambda_2)\,t}\right),$$

$$P_{(1,1),(1,0)}[t] = \frac{\lambda_2}{\lambda_1+\lambda_2-\eta_1}\left(e^{-\eta_1\,t} - e^{-(\lambda_1+\lambda_2)\,t}\right),$$

$$P_{(1,1),(0,0)}[t] = -\frac{\lambda_1}{\lambda_1+\lambda_2-\eta_2}\,e^{-\eta_2\,t} - \frac{\lambda_2}{\lambda_1+\lambda_2-\eta_1}\,e^{-\eta_1\,t}$$
$$+ 1 + \left(\frac{\lambda_1}{\lambda_1+\lambda_2-\eta_2} + \frac{\lambda_2}{\lambda_1+\lambda_2-\eta_1} - 1\right)e^{-(\lambda_1+\lambda_2)\,t},$$

$$P_{(0,1),(0,1)}[t] = e^{-\eta_2\,t}, \quad P_{(0,1),(0,0)}(t) = 1 - e^{-\eta_2\,t},$$

$$P_{(1,0),(1,0)}[t] = e^{-\eta_1\,t}, \quad P_{(1,0),(0,0)}(t) = 1 - e^{-\eta_1\,t},$$

and all other entries of P being zero. In particular, we calculate

$$\mathbb{P}(\tau_1 > t_1, \tau_2 > t_2)$$
$$= \begin{cases} P_{(1,1),(1,1)}(t_1)\left(P_{(1,1),(1,1)}(t_2-t_1) + P_{(1,1),(0,1)}(t_2-t_1)\right), & t_2 \geq t_1 \\ P_{(1,1),(1,1)}(t_2)\left(P_{(1,1),(1,1)}(t_1-t_2) + P_{(1,1),(1,0)}(t_1-t_2)\right), & t_1 > t_2 \end{cases}$$
$$= \begin{cases} \frac{\lambda_2-\eta_2}{\lambda_1+\lambda_2-\eta_2}\,e^{-(\lambda_1+\lambda_2)\,t_2} + \frac{\lambda_1}{\lambda_1+\lambda_2-\eta_2}\,e^{-\eta_2\,t_2-(\lambda_1+\lambda_2-\eta_2)\,t_1}, & t_2 \geq t_1 \\ \frac{\lambda_1-\eta_1}{\lambda_1+\lambda_2-\eta_1}\,e^{-(\lambda_1+\lambda_2)\,t_1} + \frac{\lambda_2}{\lambda_1+\lambda_2-\eta_1}\,e^{-\eta_1\,t_1-(\lambda_1+\lambda_2-\eta_1)\,t_2}, & t_1 > t_2. \end{cases}$$

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 τ 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 $\tilde{\tau}$ exists with $\tilde{\tau}_{[d]} \overset{d}{=} \tau$. Then, for $i \in [d]$, it holds that (cf. Eq. (8))

$$\tilde{\tau}_i = \min\{\tilde{E}^I : i \in I\}$$
$$= \min\{\min\{\tilde{E}^I, \tilde{E}^{I \cup \{d+1\}}\} : i \in I \subseteq [d]\},$$

where $\tilde{E}^I, \emptyset \neq I \subseteq [d+1]$, are the independent exponential random shocks from the shock model representation of $\tilde{\tau}$. In particular, it follows for the rates of τ 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* τ *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 $\text{Poi}(\lambda(t_k - t_{k-1}))$ 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 $\Upsilon^{(I)}(t) := \infty \mathbf{1}_{\{E_I \leq t\}}$ 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.
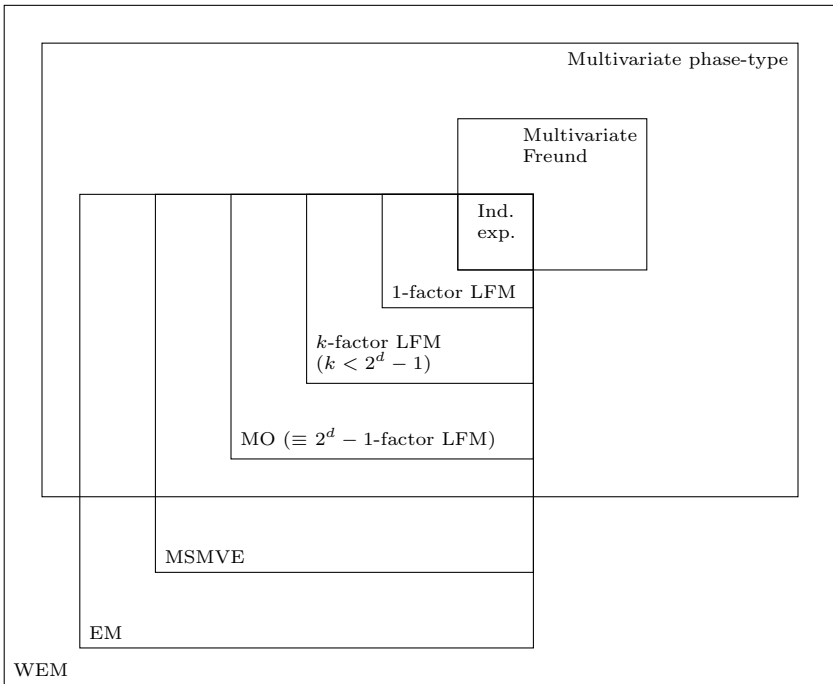

Fig. 2.   Venn-diagram of (selected) multivariate exponential, Phase-type distributions, and distributions fulfilling the WEM-property. See Chap. 2.2 as well as [41], [37], [35], [13] for details.

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 $\{1, \ldots, |S|\}$). 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 $j \in S\setminus\{i\}$ 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 $\{\boldsymbol{E}_l : l \leq k\}$, $\boldsymbol{E}_{k+1} := (E^1_{k+1}, \ldots, E^d_{k+1})'$.*
  - *(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) $\mathbb{P}_k(\min_{j \in S} E_{k+1}^j > t) = \exp\{q_{ii}t\} \ \forall t > 0$.

(II) $\mathbb{P}_k(\operatorname{argmin}_{j \in S\setminus\{i\}} E_{k+1}^j = j^\star) = -q_{ij^\star}/q_{ii} \ \forall j^\star \in S\setminus\{i\}$.

(III) The random variables $\min_{j \in S} E_{k+1}^j$ and $\operatorname{argmin}_{j \in S\setminus\{i\}} E_{k+1}^j$ 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 $\{Z(T_k) = i\}$, $\min_{j \in S} E_{k+1}^j$ 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:

$$\mathbb{P}_k\left(\operatorname*{argmin}_{j \in S\setminus\{i\}} E_{k+1}^j = j^\star\right) = \mathbb{P}_k\left(E_{k+1}^{j^\star} < \min_{j \in S\setminus\{i,j^\star\}} E_{k+1}^j\right)$$

$$\overset{(\star)}{=} \mathbb{E}_k\left[\exp\left\{-E_{k+1}^{j^\star} \sum_{j \in S\setminus\{i,j^\star\}} q_{ij}\right\}\right]$$

$$\overset{(\dagger)}{=} \frac{q_{ij^\star}}{q_{ij^\star} + \sum_{j \in S\setminus\{i,j^\star\}} q_{ij}} \overset{(\ddagger)}{=} -\frac{q_{ij^\star}}{q_{ii}},$$

where $(\star)$ follows using the tower property conditioning on $E_{k+1}^{j^\star}$, $(\dagger)$ 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:

$$\mathbb{P}_k\left(\min_{j \in S\setminus\{i\}} E_{k+1}^j > t, \operatorname*{argmin}_{j \in S\setminus\{i\}} E_{k+1}^j = j^\star\right)$$

$$\overset{(\star)}{=} \mathbb{E}_k\left[1_{\{E_{k+1}^{j^\star} > t\}} \exp\left\{-E_{k+1}^{j^\star} \sum_{j \in S\setminus\{i,j^\star\}} q_{ij}\right\}\right]$$

$$\overset{(\dagger)}{=} -\frac{q_{ij^\star}}{q_{ii}} \exp\{q_{ii}t\},$$

where $(\star)$ follows using the tower property conditioning on $E_{k+1}^{j^\star}$ and $(\dagger)$ follows using that for an exponential random variable E with rate $\eta > 0$

one has for $t, x > 0$

$$\mathbb{E}\left[1_{\{E>t\}}\exp\{-xE\}\right] = \int_t^\infty \eta \exp\{-(x+\eta)v\}dv$$
$$= \frac{\eta}{\eta+x}\exp\{-(x+\eta)t\}.$$

$\square$

Acknowledgments

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).