Limit distributions of the upper order statistics for the Lévy-frailty Marshall-Olkin distribution

Javiera Barrera\textsuperscript{a}, Guido Lagos\textsuperscript{a,b}

\textsuperscript{a}Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez
Av. Diagonal Las Torres 2640, Petalolén, Santiago, Chile
\textsuperscript{b}Industrial Engineering Department, University of Santiago
Av. Ecuador 3769, Estación Central, Santiago, Chile

Abstract

In this paper we give the precise asymptotic behavior of the upper-order statistics of the Lévy-frailty Marshall-Olkin (LFMO) distribution. Our main result is a CLT-type theorem showing that the upper-order statistics converge in distribution after a certain logarithmic centering and scaling, when assuming that the underlying Lévy subordinator process is in the normal domain of attraction of a Stable distribution. Our result is especially useful in network reliability and systemic risk, when modeling the lifetimes of the components in a system using the LFMO distribution, as it allows to give easily computable confidence intervals for the lifetimes of components.

Keywords: Marshall-Olkin distribution, Dependent random variables, Upper order statistics, Extreme-value theory, Reliability

2010 MSC: 62G30, 60K10, Categories and Subject Descriptors: G.3 [Mathematics of Computing]: Probability and Statistics; I.6 [Computing Methodology]: Simulation and Modeling

1. Introduction

In this paper we explore the upper order statistics of the Lévy-frailty (LFMO) subfamily of the multivariate Marshall-Olkin (MO) exponential distribution, also known as the conditionally-iid construction of the Marshall-Olkin distribution. A (general) MO distribution is a positive random vector in \( \mathbb{R}^n \) that is best understood as if each component of the vector represents the “lifetime” of a component in an \( n \)-component system subject to failures as follows. At time zero, all components are working; there are “shocks” that occur after independent exponentially distributed times; each shock hits a subset of components, simultaneously killing all the components in the subset that were not already killed; and killed components stay in that state forever after. The obtained MO copula is a remarkable example of extreme-value copula; see Mai & Scherer (2017). For instance, it is used to capture the dependency structure in rare event settings—see Gudendorf & Segers (2010)—and it is now considered a key tool in...
reliability theory and quantitative risk management; see Engel et al. (2017); Lindskog & McNeil (2003).
In turn, the Lévy-frailty (LFMO) subfamily of the MO distribution was proposed in Mai & Scherer (2009)
as a way to alleviate the high number of parameters needed to specify a “crude” MO distribution. This
alleviation is done by imposing exchangeability between the components’ lifetimes by means of introduc-
ing a Lévy subordinator process that acts as a latent factor. As a result, the LFMO distribution is a
flexible and powerful modeling tool that requires few parameters and is easy to simulate as long as the
Lévy subordinator is as well; see Mai & Scherer (2017).

Our interest in the upper order statistics of the LFMO distribution is threefold. First, we are motivated
by the rich combinatorial structure of the order statistics of the LFMO distribution. As we discuss in
Section 2, the first and average order statistics are relatively easy to address; however, the upper order
statistics exhibit a much more complex behavior that is difficult to grasp analytically. Second, we are also
motivated by the use of order statistics in the field of network reliability when describing the probabilistic
behavior of a system’s lifetime. In particular, using the so-called Samaniego signature result, one can
express the probability of a system being working in terms of the tail probability of the order statistics
of the component’s failure times; see Marichal et al. (2011). Third and last, the study of the stochastic
behavior of extremal order statistics is a classical topic in statistics and probability theory. Indeed,
these quantities are essential tools in statistical inference, where they have a centuries-old history of
applications, see, e.g., David (2006); and are also the essential object of study in extreme-value theory;
see, e.g., Finkenstadt & Rootzén (2003). All in all, it is important to explore the possible regimes of the
order statistics of the components of the LFMO distribution, given its wide application as a model for
simultaneous failures of components in systems.

Main contributions. The main contributions of our work are threefold.

First and foremost, we derive the precise asymptotic behavior of the upper order statistics of the
LFMO distribution as the dimension of the space grows. Indeed, we show that after certain logarithmic
centering and scaling, all the upper order statistics converge in distribution to a nondegenerate limit
random variable, and we give this limit distribution in an analytical closed-form expression in terms of a
stable random variable. Additionally, our results are particularly convenient for computational purposes:
all the required constants are usually easy to compute, the limit random variables are also easy to simulate,
and our results give a probabilistic upper bound on the time horizon required to simulate the underlying
Lévy subordinator of the LFMO distribution. From this perspective, our result is also a contribution
to a body of work in applied and computational probability concerned with providing computationally
tractable approximations for the analysis of probabilistic tools and models; quintessential examples of
this type are heavy-traffic limits in queues, fluid limits and diffusion approximations; see, e.g., Asmussen
& Glynn (2007); Asmussen (2003).

Second, our main result gives important qualitative insight into the effects of choosing the underlying
Lévy subordinator involved in the LFMO distribution. Indeed, our results hold under mild conditions, namely, that the Lévy subordinator lies in the normal domain of attraction of a stable distribution, and we show that there is a critical behavior change in terms of the index of stability $\alpha \in (0, 2]$ of the limit stable random variable. In detail, and using the perspective of the LFMO distribution as modeling failure times in a system with $n$ components, we show that $\alpha = 1$ is a critical value since for $\alpha > 1$, the instant when the last component fails grows as $O(\log n)$ and concentrates around that value. This is qualitatively the same behavior as when the lifetimes of the components are iid. On the other hand, when $\alpha \leq 1$, there is no longer a concentration phenomenon.

Third, in network reliability and systemic risk, our results allow us to give confidence intervals for the last failure times of the system’s components when using the LFMO distribution to model the failure times of components. Moreover, our results can potentially lead to give qualitative insight on the probabilistic behavior of more generally defined system failure times when using the Samaniego signature decomposition; see, e.g., Marichal et al. (2011).

**Literature review.** The study of extreme values and order statistics has been historically motivated by the study of floods, droughts, fatigue failures and other engineering applications, see David (2006). This has become a classical field, and its most relevant theorem is the Fisher-Tippett-Gnedenko result, which gives the general asymptotic behavior for the maximum and minimum of an iid sample. In the process of understanding the dependence structure generated by copula a natural step is to study the asymptotic behavior of its extremes values; this was done for the Archimedian copulas in Wüthrich (2005), but until now it has not been done for the MO model. A work tangentially related to ours is Mai (2018), who analyzes the extreme-value copulas that arise when using certain univariate extreme-value distributions as marginals in a multivariate copula. However, their work focuses on a fixed-dimensional copula, whereas our analysis is asymptotic in the number of dimensions.

Regarding the MO distribution, its study has been driven mostly by its application to reliability and systemic risk modeling, especially in the last decade or so; see Cherubini et al. (2015). This model was first proposed by Marshall & Olkin (1967) as a means to generalize the lack of memory property to two-dimensional distributions, and it was later generalized to any finite dimension. It is precisely the memoryless property that has popularized this model, as it allows efficient simulation techniques; see, e.g., Mai & Scherer (2017) and Botev et al. (2016) for rare event simulation. Nonetheless, an important shortfall of the general MO model is that it requires to specify a number of parameters that is exponential in the number of components when no further assumptions or simplifications are made. In some real-world situations, the nature of the problem can lead to natural simplifications that reduce the number of parameters, e.g., in portfolio-credit risks or in insurance; see Bernhart et al. (2015). In this line, Matus et al. (2018) propose a Lasso selection model to obtain the parameters. Another way to address the explosive number of parameters is to impose exchangeability of the components’ lifetimes, as is the case
of the LFMO model proposed by Mai & Scherer (2009) and that we will discuss in detail in Section 2.

Notation. Throughout the paper, we will use the following notation. Given a collection of random variables \((\xi_n)_n\) and \(\xi_\infty\), we write \(\xi_n \Rightarrow \xi_\infty\) when the sequence \((\xi_n)_n\) converges in distribution to \(\xi_\infty\), i.e., the distributions of the random variables \((\xi_n)_n\) converge weakly to the distribution of \(\xi_\infty\). Additionally, for a real random variable \(\xi\), we denote by \(E\xi\) its expected value and \(\text{Var}(\xi)\) its variance. As usual, we denote by \(\mathbb{R}_+\) the nonnegative real numbers. In addition, for two functions \(f\) and \(g\), we write \(f(x) \sim g(x)\) when \(\lim_{x \to +\infty} f(x)/g(x) = 1\); we write \(f(x) = o(g(x))\) when \(\lim_{x \to +\infty} f(x)/g(x) = 0\); and we write \(f(x) = O(g(x))\) when \(\limsup_{x \to +\infty} f(x)/g(x) < +\infty\). Lastly, for a set \(A\) we denote as \(1_{\{x \in A\}}\) the function in \(x\) that is equal to one when \(x \in A\) and zero otherwise.

Organization of this paper. In Section 2 we give the preliminary concepts that contextualize our work. For that, in Section 2.1 we introduce the MO and LFMO distributions; in Section 2.2 we give an introductory analysis of the order statistics of the LFMO distribution; and in Section 2.3 we give a Markov chain and cutoff perspective on the LFMO order statistics. In Section 3, we state our main result regarding the asymptotic behavior of the upper order statistics of a multivariate LFMO distribution. In Section 4, we show simulation results that computationally test the convergence of our results presented in Section 3. Finally, in Section 6, we discuss our results and summarize our main contributions.

2. Preliminaries

In this section we give an overview of the mathematical models and tools we consider in this paper. For that, in Section 2.1 we define the Marshall-Olkin (MO) distribution and the Lévy-frailty (LFMO) sub-family; later, in Section 2.2, we give a broad analysis of the order statistics of the LFMO distribution; and in Section 2.3 we give a Markov chain and cutoff perspective on the LFMO order statistics. In Section 3, we state our main result regarding the asymptotic behavior of the upper order statistics of a multivariate LFMO distribution. In Section 4, we show simulation results that computationally test the convergence of our results presented in Section 3. Finally, in Section 6, we discuss our results and summarize our main contributions.

2.1. Marshall-Olkin and Lévy-frailty Marshall-Olkin distributions

Definition 1 (Marshall-Olkin distribution). A random vector \(T \in \mathbb{R}^n\) is said to have a Marshall-Olkin (MO) distribution in \(\mathbb{R}^n\) if its components \((T_i)_i\) are defined as

\[ T_i = \min\{Z_V : V \subset \{1, \ldots, n\}, i \in V\}. \]

where \((Z_V)_V\) is a family of independent exponential random variables with rate \(\lambda_V\) for each \(V \subset \{1, \ldots, n\}\). We assume that \(\lambda_V \in \mathbb{R}_+\), with the convention that \(Z_V = \infty\) if \(\lambda_V = 0\), and additionally assume that \(\max_{V : i \in V} \lambda_V > 0\) for each \(i\).
An intuitive interpretation of Definition 1 consists of thinking of $T_i$ as the lifetime of the $i$-th component in a system with $n$ components that can be either working or not working. All components are initially working, and for each subset $V \subseteq \{1, \ldots, n\}$ at time $Z_V$, there is a “shock” that hits all components of $V$. Once a shock hits a working component, it stops working and stays in that state forever. In this way, the time $T_i$ at which a component $i$ stops working is the first time that a shock hits it. We remark that all the $T_i$ are marginally exponentially distributed. Additionally, the assumption that some shocks’ rates could be zero is equivalent to assuming that those shocks will never occur, which is used in certain settings for modeling reasons.

**Definition 2** (Lévy-Frailty Marshall-Olkin distribution). A random vector $T \in \mathbb{R}^n$ is said to have a Lévy-frailty Marshall-Olkin (LFMO) distribution in $\mathbb{R}^n$ if its components $(T_i)_i$ are defined as

\[
T_i := \inf\{t \geq 0 : S_t \geq \varepsilon_i\}, \quad i = 1, \ldots, n, \tag{1}
\]

where $S = (S_t)_{t \geq 0}$ is a Lévy subordinator with $S_0 = 0$ and $\{\varepsilon_1, \ldots, \varepsilon_n\}$ is a collection of iid exponential random variables, the “triggers”, with unit parameter and which are independent of $S$.

An intuitive interpretation of Definition 2 consists again of thinking of $T_i$ as being the lifetime of the $i$-th component in an $n$-component system in which all components are working at time zero but component $i$ fails the first time that the Lévy subordinator $S$ upcrosses the trigger $\varepsilon_i$ associated with $i$. This construction is equivalent in distribution to defining a MO distribution with the following shock rates $(\lambda_V)_V$:

\[
\lambda_V = \sum_{i=0}^{\vert V \vert - 1} \binom{\vert V \vert - 1}{i} (-1)^i \left[ \psi(n - \vert V \vert + i + 1) - \psi(n - \vert V \vert + i) \right],
\]

for all $V \subset \{1, \ldots, n\}$, where $\Psi(x) = -\log \mathbb{E}e^{-xS_1}$ is the Laplace exponent of the Lévy subordinator $S$; see, e.g., Cherubini et al. (2015). Note that the latter parameters $\lambda_V$ only depend on $\vert V \vert$, thus in particular making the components exchangeable, see Mai & Scherer (2013).

We remark that the LFMO distribution can be viewed as the result of imposing on the MO distribution the existence of a “latent” stochastic variable such that, conditional on the value of it, the random variables $T_1, \ldots, T_n$ are iid. It turns out that the Lévy subordinator $S$ characterizes such a latent variable, and conditional on the value of $S$, we have that $T_1, \ldots, T_n$ are iid with $\mathbb{P}(T_i > t \mid S) = e^{-S_t}$ for all $t \geq 0$. This outcome inspires the conditionally-iid and Lévy-frailty terminology; see, e.g., Mai (2014). For the proof of the correspondence between the two constructions, see Theorem 3.2 in Mai & Scherer (2017, ch. 3).

We also remark that the MO distribution in $n$ dimensions of Definition 1 requires specifying a number of $2^n - 1$ parameters $\lambda_V$, whereas the LFMO distribution only requires parameterizing the Lévy subordinator $S$. In this sense, the LFMO distribution is a way to reduce the parametric complexity of the MO distribution by means of introducing the latent variable $S$. In fact, most characteristics of the LFMO model are expressed in terms of $S_1$ and its Laplace exponent $\Psi(x) = -\log \mathbb{E}e^{-xS_1}$, which
have, respectively, rich probabilistic and analytical structures that allow to study and exploit the model in great detail; see, e.g., Bernhart et al. (2015); Hering & Mai (2012). Finally, it is worth mentioning that Engel et al. (2017) studies the LFMO construction when the triggers are nonhomogeneous, that is, when the triggers in Definition 2 are independent and exponential but with different rates. For further details about the LFMO distribution, see Mai & Scherer (2009, 2017).

2.2. Order statistics of the LFMO distribution

The order statistics of the CI MO distribution are, in general, not easy to perform calculations with, despite having a rich combinatorial structure. To illustrate this fact, we now give a straightforward result regarding the marginal distribution of the order statistics of the LFMO model. It is proved by applying Rényi’s representation of the order statistics of iid exponential random variables, see e.g. (Nagaraja, 2006, eq. (11.2)), together with the distribution of the sum of arbitrary independent exponential random variables, see (Nagaraja, 2006, Lemma 11.3.1).

Proposition 1. Let $T$ in $\mathbb{R}^n$ be an LFMO distributed random vector and denote by $T_{m:n}$ the $m$-th increasing order statistic of $T$. For all $m = 1, \ldots, n$ and $t \geq 0$, it holds that

$$
\mathbb{P}(T_{m:n} > t) = \sum_{k=n-m+1}^{n} e^{-\psi(k)t} \prod_{l=n-m+1}^{n} \left(1 - \frac{k}{l}\right)^{-1} \sum_{k=n-m+1}^{n} e^{-\psi(k)t} \binom{n}{k} \binom{k-1}{n-m} (-1)^{k-n+m-1},
$$

(2)

where $\Psi$ is the Laplace exponent of $S_1$, i.e., $\Psi(x) = -\log \mathbb{E} e^{-xS_1}$.

From (2), it follows that the first order statistic $T_{1:n}$ takes a simple form, as it is marginally distributed as an exponential random variable with mean $1/\Psi(n)$. Additionally, Fernández et al. (2015, Lemma 3.3) gives a simple argument to obtain the limit distribution of the average order statistics. In contrast, the upper order statistic $T_{n:n}$ is much more complex; for instance, its mean is

$$
\mathbb{E} T_{n:n} = \sum_{k=1}^{n} \binom{n}{k} \frac{(-1)^{k-1}}{\psi(k)}.
$$

(3)

Intuitively, though, if $\mathbb{E} S_1 < +\infty$, then $\mathbb{E} T_{n:n}$ should asymptotically behave as $(\log n + \gamma) / \mathbb{E} S_1$, where $\gamma$ is the Euler-Mascheroni constant, since $\mathbb{E} S_1 = t \mathbb{E} S_1$ and the mean of the maximum of $n$ iid exponential random variables of parameter 1 is $\sum_{k=1}^{n} 1/k \approx \log n + \gamma$. On the other hand, if $\mathbb{E} S_1 = \infty$ and the Lévy subordinator is stable, so in particular $\alpha < 1$, then $\mathbb{P}(S_{t_{n:n}} > (\log n + \gamma)) = \mathbb{P}(S_t > (\log n + \gamma)/t_{n}^{1/\alpha}) \to 1$ when $t_{n,n}$ is such that $t_{n,n}/(\log n)\alpha \to +\infty$. This result suggests that, for large $n$, with probability close to one, the trajectory of the subordinator will climb sufficiently fast that it surpasses all exponential triggers—thus killing all components—in an interval of size $o(\log n)$.


In perspective, the previous behaviors of the upper order statistics are not obvious to deduce from equation (3). However, in our main result in the following section, we not only corroborate them but also give the precise asymptotic distribution of all the upper order statistics of the LFMO distribution.

2.3. Markov process perspective and its abrupt convergence

The MO distribution can also be studied from the perspective of the Markov process it induces when modeling the evolution in time of components that “fail” or “break down”; see, e.g., Yuge et al. (2016). Indeed, for instance in the case of the LFMO distribution, one can define a continuous time Markov process modeling the number of components that have failed up to each time instant. The centering and scaling sequences obtained in the following section were in fact obtained by exploring the resemblance between the order statistics of the LFMO distribution and the abrupt convergence phenomenon that arises for parallel Markov chains.

The cutoff phenomenon is an abrupt convergence behavior that has been observed to occur in some sequences of ergodic Markov chains in continuous or discrete time, say \( X^1, X^2, \ldots \). Consider a distance \( d \) between probability distributions, e.g., total variation, \( L^p \), separation, Entropy, Hellinger, etc. Defining \( d_n(t) := d(X^n_t, X^n_\infty) \), i.e., \( d_n(t) \) is the distance between the distribution of \( X^n \) at time \( t \) and its stationary distribution, then it is said that there is a cutoff phenomenon if there is a sequence \( (t_n)_n \) such that, broadly speaking, before time \( t_n \) the distance to stationarity is big and after this instant is very small. The phenomenon can be described at three increasing levels of complexity: cutoff, window cutoff and profile cutoff; see Diaconis et al. (2006); Barrera & Ycart (2014) and references therein for precise definitions and discussion. When the distance is characterized by a coupling time \( T_n \), i.e., \( d_n(t) = \mathbb{P}(T_n > t) \), (for example for the total variation and separation distances) then for the sequence \( (T_n)_n \) the first cutoff level corresponds to a weak law of large numbers around \( t_n \) and the third level corresponds to the convergence in distribution of \( (T_n)_n \) under a suitable renormalization. For reversible Markov processes the cutoff phenomenon it is well understood, see Chen et al. (2008) for the cutoff result and Basu et al. (2015) for the profile cutoff result. On the other hand, for non-reversible Markov processes no general result is known yet, but there are precise results for interesting chains such as random card shuffles, see Chen & Saloff-Coste (2008). A key ingredient to prove window and profile cutoff results have been concentration inequalities and central limit theorem for coupling and hitting times; see Hermon et al. (2018) and Martínez & Ycart (2001).

From this perspective, our main result in the following section is in fact a cutoff result on the time of absorption of the Markov process. More specifically, our result gives the cutoff profile in the total variation distance for the sequence \( X^1, X^2, \ldots \) of Markov processes, where \( X^n = (X^n_t : t \geq 0) \) models the number of components that have failed up to each time in a system with \( n \) components whose lifetimes have a LFMO distribution.
3. Main result

In this section, we give the main result our paper, which shows the asymptotic behavior of the upper order statistics of the LFMO distribution. We will consider the following two hypotheses.

**Hypothesis (A,):** \( \var{S_1} = +\infty \), i.e., the variance of \( S_1 \) is infinite, and \( \prob{S_1 > t} \sim A/t^\alpha \) as \( t \to +\infty \) for some \( \alpha \in (0, 2) \) and \( A > 0 \).

**Hypothesis (B):** \( 0 < \var{S_1} < +\infty \).

Additionally, when working under Hypothesis (A,), we will use the constant \( C_\alpha \) defined as

\[
C_\alpha := \begin{cases} 
\frac{1-\alpha}{\Gamma((2-\alpha)\cos(\pi\alpha/2))} & \text{if } \alpha \in (0, 1) \cup (1, 2) \\
2/\pi & \text{if } \alpha = 1.
\end{cases}
\]  

Finally, we will use the parameterization and notation of Whitt (2002, Section 4.5.1) for stable random variables. The following is the main result of the paper.

**Theorem 1.** Consider a random vector \( T \in \mathbb{R}^n \) having a Lévy-frailty Marshall-Olkin (LFMO) distribution in \( n \) dimensions with underlying Lévy subordinator \( S \), as in Definition 2. Denote by \( T_{m:n} \) the \( m \)-th increasing order statistic of the components of \( T \), i.e., \( \{T_1, ..., T_n\} = \{T_{1:n}, ..., T_{n:n}\} \) and \( T_{1:n} \leq ... \leq T_{n:n} \).

1. Assume that \( \ex{S_1} < +\infty \) and either Hypothesis (A,) holds for some \( \alpha \in (1, 2) \) or Hypothesis (B) holds, in which case we set \( \alpha := 2 \). Let \( (m_n) \) be any integer sequence satisfying \( 1 \leq m_n \leq n \) for all \( n \) and \( \log(n - m_n) = o ((\log n)^{1/\alpha}) \) as \( n \to +\infty \). Then, as \( n \to +\infty \), we have that

\[
\frac{T_{m_n:n} - \log n/\ex{S_1}}{(\log n)^{1/\alpha}/\ex{S_1}} \Rightarrow \Sigma,
\]  

where \( \Sigma \) is a Stable_{\alpha}(\sigma, -1, 0) random variable under Hypothesis (A,) and a Normal(0, \( \sigma^2 \)) random variable under Hypothesis (B), with \( \sigma \) defined as

\[
\sigma := \begin{cases} 
\left( \frac{4}{C_\alpha \ex{S_1}} \right)^{1/\alpha} & \text{under Hypothesis (A,)} \\
\sqrt{\var{S_1}/\ex{S_1}} & \text{under Hypothesis (B)},
\end{cases}
\]  

and \( C_\alpha \) defined as in (4).

2. Assume that \( \ex{S_1} = +\infty \) and Hypothesis (A,) holds for some \( \alpha \in (0, 1] \). Let \( (m_n) \) be an integer sequence satisfying \( 1 \leq m_n \leq n \) for all \( n \) and \( n - m_n = o (n^\rho) \) as \( n \to +\infty \) for all \( \rho \in (0, 1) \). Then, as \( n \to +\infty \), it holds that

\[
\frac{T_{m_n:n}}{(\log n)^{1/\alpha}} \Rightarrow \frac{1}{\Sigma^{\alpha}},
\]  

where \( \Sigma \) is a Stable_{\alpha}(\sigma, 1, 0) random variable with \( \sigma := (A/C_\alpha)^{1/\alpha} \) and \( C_\alpha \) is as defined in (4) above.
The proof of the previous result is shown in Section 5. The key idea is to exploit the conditionally iid property to write the distribution of the left-hand side of (5), conditional on the path of $S$, as a deterministic function evaluated in a certain “zoom-out” centering and scaling of $S$. We then show that the thus obtained sequence of deterministic functions converges pointwise and that the “zoom-out” version of $S$ converges in distribution. We conclude that both limits hold simultaneously by using a generalized version of the continuous mapping theorem for weak convergence.

Remarks on Theorem 1.

1. First, we note that Hypotheses (A) and (B) describe all the non-trivial cases in which the distribution of $S_1$ is in the normal domain of attraction of a stable distribution; that is, the cases in which there exists a sequence $(\mu_n)_n$ and a $\rho > 0$ such that $(S_n - \mu_n)/n^\rho$ converges in distribution as $n \to +\infty$ to a stable distribution. See, e.g., Whitt (2002, Section 4.5) for further details.

2. The case of $\text{Var}(S_1) = 0$ corresponds to the trivial case $S_t = S_1 t$ with $S_1$ a deterministic constant. In this case the components $T_i$ are iid distributed exponential random variables with mean $1/S_1$, and it is easily shown that, e.g., $S_1 (T_{n:n} - \log n/S_1)$ converges in distribution to a standard Gumbel distribution as $n \to +\infty$.

3. When both $E S_1 = +\infty$ and Hypothesis (A) hold, then necessarily, $\alpha \in (0, 1]$. On the other hand, when $S_1$ satisfies $\mathbb{P}(S_1 > t) \sim A/t^\alpha$ as $t \to +\infty$ for some positive $A$ and an $\alpha > 2$, then necessarily, $\text{Var}(S_1) < +\infty$, in which case Hypothesis (B) holds.

4. The distribution of the random variable on the right-hand side of (7) converges as $\alpha \searrow 0$ to an exponential distribution with mean $1/\sigma^\alpha$; see, e.g., Maller & Schindler (2018). Informally then, under the assumptions of part 2. of Theorem 1, for large $n$ and $\alpha$ close to zero, the distribution of $T_{m:n}$ is “similar” to an exponential distribution with mean $(\log n)^{1/\alpha}/\sigma^\alpha$.

5. Note that the components of $T$ are dependent; however, their marginals are identically distributed as an exponential distribution with mean $1/\Psi(1)$, where $\Psi(1) = -\log \mathbb{E} e^{-S_1}$. Therefore, if the components are instead assumed to be iid distributed with the same marginal as before then now $\Psi(1) (T_{n:n} - \log n/\Psi(1))$ converges in distribution, as $n \to +\infty$, to a standard Gumbel distribution.

6. Theorem 1 suggests that the rate at which the random variable $T_{m:n}/\log n$ converges to $1/E S_1$ is $(\log n)^{-(\alpha - 1)/\alpha}$. Indeed, Theorem 1 states that intuitively, and with a slight abuse of notation, when $n$ is large, the random variable $T_{m:n}/\log n$ is approximately distributed as $1/E S_1 + (\log n)^{(1-\alpha)/\alpha}/E S_1 \cdot \text{Stable}_\alpha(\sigma, -1, 0)$. In particular, the convergence is thus slower when $\alpha$ is close to 1, which is the behavior that we observe empirically in the computational experiments presented in Section 4.

7. A direct application of Theorem 1 is to obtain confidence intervals for the upper order statistics $T_{m:n}$ when $n$ is considered sufficiently large. This approach is especially useful in network reliability.
and systemic risk, where the LFMO distribution is used to model failure times of components in a system with simultaneous failures. From this perspective, Theorem 1 can give confidence intervals for the last failures of the components in the system. Moreover, using the so-called Samaniego signature result (see Marichal et al. (2011)), one may be able to derive confidence intervals for fairly general definitions of systemic failure times; this is a future research direction to be explored.

4. Computational experiments

To empirically test the convergence in Theorem 1, in this section, we show the results of performing a large number of Monte Carlo simulations of the random variables on the left- and right-hand sides of (5). To simulate the left-hand side of (5), we sample a multivariate random vector $T$ having a LFMO distribution. This process is easily done using the property that $T$ is equal in distribution to the joint distribution of the times at which the subordinator $S$ up-crosses the collection of iid exponentially distributed “triggers”, as defined in equation (2). On the other hand, to simulate the right-hand side of (5), we use the stable random variable samplers for Julia available at White (2013), that uses Notation 1 of Nolan (2018, Section 1.3). We remark, however, that there are multiple parameterizations for stable random variables—see the discussion in Nolan (2018, Section 1.3)—so care is needed to produce samples consistent with the parameterization of Whitt (2002, Section 4.5.1), which is the one we use in this paper.

In Figure 1, we show the empirical cumulative distribution function resulting from $10^6$ Monte Carlo simulations of the left- and right-hand sides of (5), in the case of choosing $m_n := n$ in Theorem 1, and for several values of $n$. Here, the Lévy subordinator $S$ is a compound Poisson process (CPP) with rate $\lambda = 1$ and with Pareto($\alpha$) step sizes, and we test multiple values of $\alpha$.

Intuitively, to approach the limit, we require at the least that the centered and scaled Lévy process behaves similar to the corresponding limiting stable distribution. Moreover, as we already discussed, the last exponential trigger is located at $\log n$. Therefore, to prove our main result, we use the convergence of the centered and scaled Lévy subordinator at a time sequence that grows as $O(\log n)$. When $\alpha = 2$, the Berry-Esseen theorem guarantees that the convergence rate must be $(\log n)^{-1/2}$. However, in the case in which $\alpha \in (1, 2)$, it is known that convergence is given by a regularly varying function and there is no lower limit for the speed: any slowly varying function tending to zero can serve as the rate function; see De Haan & Peng (1999). Therefore, a very slow rate of convergence can be obtained for $\alpha$ in $(1, 2)$, and thus it is not surprising that the convergence shown in Figure 1 is not fast. Our experiments suggest that the rate of convergence depends on $\alpha$. In particular, for $\alpha = 1.05$, the convergence is slower than for the other values of $\alpha$ analyzed. This is probably because the case $\alpha = 1.05$ is very close to the critical value $\alpha = 1$, which separates two very different asymptotic behaviors.
5. Proof of the main result

In this section, we present the proof of our main result, Theorem 1. For that purpose, in Section 5.1, we present the main plan and lemmas used in the proof, and in Section 5.2, we give the actual proof.

5.1. Elements of the proof

The plan of the proof for both parts 1. and 2. of Theorem 1 is the same and is as follows.

We start by considering the sequence of random variables for which we want to show that converge, i.e., the ones on the left-hand sides of (5) and (7) for parts 1. and 2., respectively, and write their tail distribution function conditional on the trajectory of the underlying Lévy subordinator $S$. The first step of the plan is to show that we can construct a sequence of deterministic real functions, say $(T_n)_n$, and a sequence of real random variables, say $(\xi_n)_n$, such that the conditional tail probability at hand is almost surely equal to $T_n(\xi_n)$. Here, the random variables $\xi_n$ are related to a certain “zoom-out” centering and scaling of the Lévy subordinator $S$, and the functions $T_n$ are related to the cumulative distribution function of a binomial random variable.

The second step in the proof plan is to show that the sequence $(\xi_n)_n$ converges in distribution. We
argue this by using the Kolmogorov-Gnedenko results that generalize the central limit theorem in the heavy-tailed setting; see, e.g., Whitt (2002, Section 4.5).

The third step in the plan is to show that the sequence \((f_n)_n\) converges pointwise almost everywhere. To argue this, we will need to use the following result, whose proof we defer to the appendix.

**Lemma 1.** Let \((p_n)_n\) be a sequence in \((0,1)\) and denote by \(b(n, p_n)\) a binomial random variable with parameters \((n, p_n)\). Let \((k_n)_n\) be sequence of nonnegative integers.

1. If \(np_n \to 0\) as \(n \to +\infty\) then \(P(b(n, p_n) \leq k_n) \to 1\).
2. If \(np_n \to +\infty\) and \(p_n \to 0\) as \(n \to +\infty\) then \(P(b(n, p_n) \leq k_n) \to 0\) and \(\lim_n (k_n - np_n)/\sqrt{np_n} = -\infty\) are equivalent.

Finally, the fourth and last step of the proof plan is to argue that \(T_n(\xi_n)\) converges in distribution since \(T_n\) and \(\xi_n\) converge pointwise and in distribution, respectively. For this, we will use Lemma 2 below, which is a generalized version of the continuous mapping theorem for weak convergence.

**Lemma 2** *(Theorem 4.27 of Kallenberg (2002)).* Let \(\xi\) and \((\xi_n)\) be random variables on a metric space \(X\) satisfying \(\xi_n \Rightarrow \xi\) as \(n \to +\infty\). Consider another metric space \(Y\), and let \(T\) and \((T_n)\) be measurable mappings from \(X\) to \(Y\). Assume that for some measurable set \(\Xi \subseteq X\), it holds that \(\xi \in \Xi\) almost surely and that for all \(x \in \Xi\) and all sequences \((x_n)\) in \(X\) such that \(x_n \to x\) we have that \(T_n(x_n) \to T(x)\). Then, it holds that \(T_n(\xi_n) \Rightarrow T(\xi)\) as \(n \to +\infty\).

### 5.2. Proof of Theorem 1

We now prove Theorem 1. To clarify the exposition, we present the proofs of parts 1. and 2. separately.

**Proof of Theorem 1 part 1.** We will actually show that for all \(t \in \mathbb{R}\), we have that

\[
P \left( T_{m_n:n} > \frac{\log n + t (\log n)^{1/\alpha}}{E S_1} \left| \left( S_s \right)_{s \in \mathbb{R}_+} \right. \right) \to 1_{\{\sigma \Sigma_\infty + t \leq 0\}} \tag{8} \]

as \(n \to +\infty\) if \(n - m_n = o\left((\log n)^{1/\alpha}\right)\), where \(\Sigma_\infty\) is a Stable\(_{\alpha}(1,1,0)\) random variable and \(\sigma\) is as defined in (6). This result immediately implies part 1. of Theorem 1. Indeed, since both random variables on the left- and right-hand sides of display (8) have bounded support, then the convergence also holds when taking expected value; thus, we obtain that if \(n - m_n = o\left((\log n)^{1/\alpha}\right)\), then for all \(t\), we have, after rearranging terms,

\[
P \left( \frac{T_{m_n:n} - \log n/ES_1}{(\log n)^{1/\alpha}/ES_1} > t \right) \to P(-\sigma \Sigma_\infty > t) \tag{9} \]

as \(n \to +\infty\). We conclude by noting that \(\Sigma_\infty\) is distributed as a Stable\(_{\alpha}(\sigma,-1,0)\) random variable under Hypothesis \((A_\alpha)\)—see Whitt (2002, Section 4.5.1)—and as a Normal\((0,\sigma^2)\) random variable under Hypothesis \((B)\), and that the limit (9) holding for all \(t\) characterizes the limit in distribution (5).
The first part of the proof consists of showing that for all \( t \in \mathbb{R} \) and all \( n \) sufficiently large such that \( \log n + t(\log n)^{1/\alpha} \geq 0 \), we have that

\[
P \left( T_{m_n:n} > \frac{\log n + t(\log n)^{1/\alpha}}{\mathbb{E}S_1} \right| (S_s)_{s \in \mathbb{R}_+} \right) = 1 - f_n(\sigma \Sigma_n + t) \tag{10}
\]

almost surely. Here, the (deterministic) functions \((f_n)_n\) are defined as

\[
f_n(x) := \begin{cases} 
\mathbb{P} \left( b(n, e^{-x(\log n)^{1/\alpha}}) \leq n - m_n \right) & \text{if } x \geq -\frac{\log n}{(\log n)^{1/\alpha}} \\
\mathbb{P} \left( b(n, 1) \leq n - m_n \right) & \text{otherwise,}
\end{cases}
\tag{11}
\]

with \( b(n, p) \) denoting a binomial random variable with parameters \( n \) and \( p \). The sequence of random variables \((\Sigma_n)_n\) is defined as

\[
\Sigma_n := \left( S_{u_n - u_n \mathbb{E}S_1} \right) \frac{1 + t(\log n)^{-(\alpha - 1)/\alpha}}{\sigma (u_n \mathbb{E}S_1)^{1/\alpha}},
\tag{12}
\]

with \( \sigma \) as in (6) and where \( u_n = u_n(t) \) is defined as

\[
u_n := \frac{\log n + t(\log n)^{1/\alpha}}{\mathbb{E}S_1}.
\tag{13}
\]

We remark that, by the definition of \( u_n \), the condition that \( n \) is sufficiently large such that \( \log n + t(\log n)^{1/\alpha} \geq 0 \) is equivalent to the condition \( u_n \geq 0 \). We also remark that the condition \( u_n \geq 0 \) implies that the argument \( \sigma \Sigma_n + t \) in the term \( f_n(\sigma \Sigma_n + t) \) is always in the domain \( x \geq -\log n/(\log n)^{1/\alpha} \) in the definition of \( f_n \), i.e., \( \sigma \Sigma_n + t \geq -\log n/(\log n)^{1/\alpha} \) almost surely; this is direct by using that \( S \) has nondecreasing paths with \( S_0 = 0 \) and using the definitions (6), (12) and (13) of \( \sigma, \Sigma_n \) and \( u_n \), respectively.

Indeed, to prove the characterization (10) of the term \( \mathbb{P} \left( T_{m_n:n} > \frac{\log n + t(\log n)^{1/\alpha}}{\mathbb{E}S_1} \right| (S_s)_{s \in \mathbb{R}_+} \right) \), we first use that \( S \) is a Lévy subordinator with \( S_0 = 0 \), and that given \((S_s)_{s \in \mathbb{R}_+}\), we know that the times \( T_1, \ldots, T_n \) are iid with the event \( \{ T_k > u \} \) having probability \( e^{-S_k} \), i.e., \( \mathbb{P}(T_k > u \mid (S_s)_{s \in \mathbb{R}_+}) = e^{-S_k} \). This implies that for all \( u \geq 0 \) and all \( k = 1, \ldots, n \) we have

\[
\mathbb{P} \left( T_{k:n} > u \mid (S_s)_{s \in \mathbb{R}_+} \right) = \mathbb{P} \left( b(n, e^{-S_k}) > n - k \mid (S_s)_{s \in \mathbb{R}_+} \right).
\]

In particular, for \( k = m_n \) we have

\[
\mathbb{P} \left( T_{m_n:n} > u \mid (S_s)_{s \in \mathbb{R}_+} \right) = 1 - \mathbb{P} \left( b(n, e^{-S_k}) \leq n - m_n \mid (S_s)_{s \in \mathbb{R}_+} \right). \tag{14}
\]

Now, note that

\[
e^{-S_k} = \exp \left( -\sigma \left( \frac{S_u - u \mathbb{E}S_1}{\sigma (u \mathbb{E}S_1)^{1/\alpha}} \right) (u \mathbb{E}S_1)^{1/\alpha} + \log n - u \mathbb{E}S_1 \right) / n.
\]
In particular, plugging-in \( u = u_n \), with \( u_n \) as defined in (13) and \( n \) sufficiently large such that \( u_n \geq 0 \), we obtain that
\[
\exp(-S_{u_n}) = \exp\left(-\sigma \frac{S_{u_n} - u_n\mathbb{E}S_1}{\sigma(u_n\mathbb{E}S_1)^{1/\alpha}} \left(\log n + t(\log n)^{1/\alpha}\right)^{1/\alpha} - t(\log n)^{1/\alpha}\right) / n
\]
\[
= \exp\left(-\left[\sigma \frac{S_{u_n} - u_n\mathbb{E}S_1}{\sigma(u_n\mathbb{E}S_1)^{1/\alpha}}\left(1 + t(\log n)^{-(\alpha-1)/\alpha}\right)^{1/\alpha} + t\right] (\log n)^{1/\alpha}\right) / n
\]
\[
= \exp\left(-(\sigma\Sigma_n + t) (\log n)^{1/\alpha}\right) / n,
\]
where in the last equation, we used the definition (12) of \( \Sigma_n \). Plugging in this expression for \( e^{-S_{u_n}} \) and \( u = u_n = (\log n + t(\log n)^{1/\alpha})/\mathbb{E}S_1 \) in equation (14), we obtain that, almost surely,
\[
\mathbb{P}\left(T_{m_n,n} > \frac{\log n + t(\log n)^{1/\alpha}}{\mathbb{E}S_1} \left| (S_s)_{s \in \mathbb{R}_+}\right.\right)
\]
\[
= 1 - \mathbb{P}\left(b(n, e^{-(\sigma\Sigma_n + t)(\log n)^{1/\alpha}}/n) \leq n - m_n \right| (S_s)_{s \in \mathbb{R}_+}\right)
\]
\[
= 1 - f_n(\sigma\Sigma_n + t),
\]
where in the last equality we used the definition (11) of \( f_n \), and that for all \( n \) sufficiently large such that \( u_n \geq 0 \) we have that \( \Sigma_n \) is measurable with respect to the sigma-algebra generated by \( (S_s)_{s \in \mathbb{R}_+} \). This proves equation (10).

The second part of the proof consists of showing that the sequence \((\Sigma_n : n \geq 1)\), with \( \Sigma_n \) as defined in (12), satisfies the convergence in distribution
\[
\Sigma_n \implies \Sigma_{\infty}
\] (15)
as \( n \to +\infty \), where \( \Sigma_{\infty} \) is a Stable(\( \alpha, 1, 0 \)) random variable under Hypothesis (A\( \alpha \)) and a Normal(0,1) random variable under Hypothesis (B). Indeed, first note that the term \((1 + t(\log n)^{-(\alpha-1)/\alpha})^{1/\alpha} \) in the definition of \( \Sigma_n \) converges to 1 as \( n \to +\infty \) since \( \alpha \) in particular satisfies \( \alpha > 1 \). On the other hand, using the definition (6) of \( \sigma \), we have that as \( u \to +\infty \), the random variable \((S_u - u\mathbb{E}S_1)/\sigma(u\mathbb{E}S_1)^{1/\alpha}\) converges in distribution to a Stable(\( \alpha, 1, 0 \)) random variable under Hypothesis (A\( \alpha \)) and a Normal(0,1) random variable under Hypothesis (B); this holds by Whitt (2002, Theorem 4.5.2) in the case of Hypothesis (A\( \alpha \)) and by the central limit theorem in the case of Hypothesis (B). We conclude the convergence (15) by noting that \( u_n \to +\infty \) as \( n \to +\infty \).

The third part of the proof consists of showing that the functions \((f_n)_n\) defined in (11) satisfy
\[
\lim_n f_n(x) = 1_{\{x > 0\}} \text{ for all } x \neq 0 \text{ if and only if } n - m_n = o\left((\log n)^{1/\alpha}\right).
\] (16)
Indeed, first note that \( \lim_n f_n(x) = 1 \) holds for all \( x > 0 \). This comes from applying part 1. of Lemma 1 with \( p_n = q_n(x) \) defined as
\[
q_n(x) := e^{-x(\log n)^{1/\alpha}} / n,
\] (17)
since \( nq_n(x) \to 0 \) as \( n \to +\infty \) for all \( x > 0 \). Next, we argue that \( \lim_n f_n(x) = 0 \) holds for all \( x < 0 \) if and only if \( n - m_n = o\left((\log n)^{1/\alpha}\right) \). To do this, we first apply part 2. of Lemma 1 with \( p_n = q_n(x) \) as in (17), since for all \( x < 0 \), we have that \( nq_n(x) \to +\infty \) and \( q_n(x) \to 0 \) since \( \alpha > 1 \); in this way, we obtain that \( \lim_n f_n(x) = 0 \) holds for all \( x < 0 \) if and only if

\[
\lim_n (n - m_n - nq_n(x))/\sqrt{nq_n(x)} = -\infty \quad \text{for all } x < 0. \tag{18}
\]

We now argue that the condition (18) is equivalent to \( n - m_n = o\left((\log n)^{1/\alpha}\right) \) as \( n \to +\infty \). For this, recall that \( x < 0 \) and note that rewriting

\[
\frac{n - m_n - nq_n(x)}{\sqrt{nq_n(x)}} = \frac{n - m_n - e^{-x(\log n)^{1/\alpha}}}{e^{-\frac{2}{3}(\log n)^{1/\alpha}}}
\]

we can see that if \( (n - m_n)/(\log n)^{1/\alpha} \to 0 \) then (18) holds; and on the other hand, if \( \limsup_n (n - m_n)/(\log n)^{1/\alpha} = \epsilon > 0 \), then it is sufficient to plug in \( x = -\epsilon/2 \) in (19) to obtain that \( \limsup_n (n - m_n - nq_n(x))/\sqrt{nq_n(x)} = +\infty \). This proves equivalence (16).

The fourth part of the proof consists of showing that if \( n - m_n = o\left((\log n)^{1/\alpha}\right) \) holds then

\[
f_n \left(\sigma \Sigma_n + t\right) \Rightarrow 1_{\{\sigma \Sigma_n + t > 0\}} \tag{20}
\]

holds as \( n \to +\infty \). Intuitively, this should be true because from the second part of the proof we have \( \sigma \Sigma_n + t \Rightarrow \sigma \Sigma_n + t \) as \( n \to +\infty \), and from the third part of the proof we have that \( f_n(x) \to 1_{\{x > 0\}} \) holds for all \( x \neq 0 \) if \( n - m_n = o\left((\log n)^{1/\alpha}\right) \). To formalize this intuition, we apply Lemma 2, which is a generalized version of the continuous mapping theorem for weak convergence. Indeed, we can apply this result because, first, \( f_n \) is continuous for all \( n \); and second, assuming \( n - m_n = o\left((\log n)^{1/\alpha}\right) \) and defining \( \Xi := \mathbb{R} \setminus \{0\} \), we have \( \mathbb{P}(\sigma \Sigma_n + t \in \Xi) = 1 \) and \( \lim_n f_n(x) = 1_{\{x > 0\}} \) for all \( x \in \Xi \). This gives the limit in distribution (20) under the aforementioned growth condition on \( n - m_n \).

Finally, using the limit (20) in equation (10), we obtain that (8) holds if \( n - m_n = o\left((\log n)^{1/\alpha}\right) \), which is what we wanted to prove, as we argued in the beginning of this proof. This concludes the proof of part 1. of Theorem 1.

We now give the proof of part 2. of Theorem 1. We give a concise version of the proof, as the arguments parallel the main ideas of the proof of part 1.

Proof of Theorem 1 part 2. We will actually show that for all \( t \in \mathbb{R} \), we have that

\[
\mathbb{P}\left(T_{m_n:n} > t (\log n)^{1/\alpha} \mid (S_s)_{s \in \mathbb{R}_+}\right) \Rightarrow 1_{\{\sigma \Pi_n t^{1/\alpha} < 1\}} \tag{21}
\]
as \( n \to +\infty \) if \( n - m_n = o(n^\rho) \) for all \( \rho \in (0, 1) \), where \( \Pi_\infty \) is a Stable\(_\alpha(1, 1, 0) \) random variable and \( \sigma \) is defined as

\[
\sigma := \left( \frac{A}{C_\alpha} \right)^{1/\alpha}. \tag{22}
\]

The result (21) immediately implies part 2. of Theorem 1. Indeed, both random variables on the left-and right-hand sides of display (21) have bounded support, so the convergence also holds when taking the expected value, thus obtaining after rearranging terms

\[
P \left( \frac{T_{m_n:n}}{(\log n)^{1/\alpha}} > t \right) \to P \left( \frac{1}{(\sigma \Pi_\infty)^{1/\alpha}} > t \right) \tag{23}
\]
as \( n \to +\infty \). We conclude by noting that \( \sigma \Pi_\infty \) is distributed as a Stable\(_\alpha(\sigma, 1, 0) \) random variable—see Whitt (2002, Section 4.5.1)—and that the limit (23) holding for all \( t \in \mathbb{R} \) characterizes the limit in distribution (7), which is what we want to prove.

The first part of the proof consists of showing that for all \( t \in \mathbb{R} \) and all \( n \geq 1 \), we have that

\[
P \left( T_{m_n:n} > t (\log n)^{1/\alpha} \middle| (S_s)_{s \in \mathbb{R}_+} \right) = 1 - g_n(\sigma \Pi_n t^{1/\alpha}), \tag{24}
\]
almost surely. Here, the (deterministic) functions \((g_n)_n\) are defined as

\[
g_n(x) := \begin{cases} P(b(n, 1/n^x) \leq n - m_n) & \text{if } x \geq 0 \\ P(b(n, 1) \leq n - m_n) & \text{otherwise}, \end{cases} \tag{25}
\]
and the sequence of random variables \((\Pi_n)_n\) is defined as

\[
\Pi_n := \frac{S_{v_n}}{\sigma v_n^{1/\alpha}}, \tag{26}
\]
with \( \sigma \) as defined in (22) and

\[
v_n = v_n(t) := (\log n)^\alpha t. \tag{27}
\]
Indeed, substituting \( u := v_n \) into equation (14) we obtain

\[
P \left( T_{m_n:n} > (\log n)^\alpha t \middle| (S_s)_{s \in \mathbb{R}_+} \right) = 1 - P \left( b(n, e^{-S_{v_n}}) \leq n - m_n \middle| (S_s)_{s \in \mathbb{R}_+} \right),
\]
and rewriting \( e^{-S_{v_n}} \) as

\[
e^{-S_{v_n}} = \exp \left( -\frac{S_{v_n}}{\sigma v_n^{1/\alpha}} \log n \right) = \exp \left( -\sigma \Pi_n t^{1/\alpha} \log n \right) = 1/n^{\sigma \Pi_n t^{1/\alpha}}
\]
we obtain the characterization (24) of \( P \left( T_{m_n:n} > t (\log n)^{1/\alpha} \middle| (S_s)_{s \in \mathbb{R}_+} \right) \).

The second part of the proof is to note that by Whitt (2002, Theorem 4.5.2), the random variable \( \Pi_n \) converges in distribution to a Stable\(_\alpha(1, 1, 0) \) random variable, say \( \Pi_\infty \).
The third part of the proof consists of showing the equivalence

$$\lim_{n} g_n(x) = 1_{\{x > 1\}} \text{ for all } x \in \mathbb{R}_+ \setminus \{0, 1\} \text{ if and only if } n - m_n = o(n^\rho) \text{ for all } \rho \in (0, 1). \quad (28)$$

Indeed, using $p_n := 1/n^\rho$ in part 1. of Lemma 1, we obtain that $\lim_{n} g_n(x) = 1$ for all $x > 1$. Additionally, applying part 2. of Lemma 1 with $p_n$ as before, we obtain that $\lim_{n} g_n(x) = 0$ holds for all $x \in (0, 1)$ if and only if $\lim_{n}(n - m_n - n/n^\rho)/\sqrt{n/n^\rho} = -\infty$ holds for all $x \in (0, 1)$, which in turn is equivalent to $n - m_n = o(n^\rho)$ for all $\rho \in (0, 1)$. The latter equivalence is checked by noting that

$$\frac{n - m_n - n/n^\rho}{\sqrt{n/n^\rho}} = n^{1-x} \left( \frac{n - m_n}{n^{1-x}} - 1 \right),$$

so if there exists $\rho \in (0, 1)$ such that $\lim sup_n (n - m_n)/n^\rho > 0$, then taking $x^* := 1 - \rho/2$, we obtain that $x^* \in (0, 1)$ and $\lim sup_n (n - m_n - n/n^{x^*})/\sqrt{n/n^{x^*}} = +\infty$.

The fourth part of the proof consists of showing that if $n - m_n = o(n^\rho)$ for all $\rho \in (0, 1)$, then

$$g_n \left( \sigma_1 t^{1/\alpha} \Pi_n \right) \Rightarrow 1_{\{\sigma_1 t^{1/\alpha} \Pi_{\infty} > 1\}} \quad (29)$$

as $n \to +\infty$. This comes from applying Lemma 2 with $\Xi := \mathbb{R}_+ \setminus \{0, 1\}$, since from the second part of the proof we have that $\sigma_1 t^{1/\alpha} \Pi_{\infty} \Rightarrow \sigma_1 \Pi_{\infty} t^{1/\alpha}$, with $\mathbb{P}(\sigma_1 \Pi_{\infty} t^{1/\alpha} \in \Xi) = 1$; and from the third part of the proof we have that $\lim_n g_n(x) = 1_{\{x > 1\}}$ for all $x \in \Xi$, under the aforementioned growth condition on $n - m_n$.

Finally, using the limit (29) in equation (24), we obtain that (21) holds if $n - m_n = o(n^\rho)$ for all $\rho \in (0, 1)$, which is what we wanted to prove. This concludes the proof of part 2. of Theorem 1. \hfill \square

6. Conclusions

The Lévy-frailty Marshall-Olkin (LFMO) distribution is an important tool to model the failure times of components in systems where failures can affect several components simultaneously. The dependency between components in the LFMO distribution is introduced via a Lévy subordinator that triggers failures when it up-crosses some exponentially distributed barriers called the triggers. Our main result gives the asymptotic behavior of the upper order statistics of the LFMO distribution. From the reliability and risk modeling perspective, our result gives the asymptotic probabilistic behavior of the last components of the system to fail as the size of the system grows. Our results shows that, in terms of the index of stability $\alpha > 0$ of the Lévy subordinator, $\alpha = 1$ is a critical value or frontier between two regimes that exhibit different qualitative asymptotic behaviors. Indeed, our result shows that for $\alpha > 1$, the asymptotic life of the last component to fail grows as $O(\log(n))$, which is qualitatively the same behavior as when the failure times between components are assumed to be iid, but in the LFMO case we observe larger fluctuations. On the other hand, when $\alpha \leq 1$, the underlying Lévy subordinator will climb sufficiently high in an $O(\log n)$ amount of time to turn off all the components, behaving qualitatively as one component in that scale of time. We believe that these two behaviors should be observed in the general MO model, that is,
if shocks affecting massive subsets of components occur frequently then a behavior similar to the case in which \( \alpha \in (0, 1) \) should be expected; and if massive shocks are rarer then a behavior similar to the case in which \( \alpha > 1 \) should be observed. However, in the general MO setting there is no clear interpretation of the index of stability \( \alpha \). Also, we conjecture that exploiting the connection with the cutoff phenomenon could give the key ideas to prove the result for a general exchangeable case and even for the general MO case.

Finally, we remark that the proof of our result combines the convergence of the binomial representation of the order statistics and the convergence to a stable distribution of a centered and scaled Lévy subordinator. This outcome suggests that bounds on the speed of convergence can be obtained.

Acknowledgements

Javiera Barrera acknowledges the financial support of FONDECYT grant 1161064 and of Programa Iniciativa Científica Milenio NC120062. Guido Lagos acknowledges the financial support of FONDECYT grant 3180767, Programa Iniciativa Científica Milenio NC120062, and the Center for Mathematical Modeling of Universidad de Chile through their grant Proyecto Basal PFB-03.

Appendices

Lemma 3. Let \((q_n) \subset \mathbb{R}\) be a sequence such that \(q_n = o(\sqrt{n})\) as \(n \to +\infty\). Then, \((1 + q_n/n)^n \sim e^{q_n}\) as \(n \to +\infty\), i.e., \(\lim_{n \to +\infty} (1 + q_n/n)^n / e^{q_n} = 1\).

Proof. For all \(n\) sufficiently large, we can write

\[
\frac{(1 + q_n/n)^n}{e^{q_n}} = \exp\left( n \log(1 + q_n/n) - q_n \right)
= \exp\left( \sum_{j \geq 2} (-1)^{j+1} j! \left(q_n/\sqrt{n}\right)^j \frac{1}{n^{j-1}} \right)
= \exp\left( -2 \left(q_n/\sqrt{n}\right)^2 + o \left( (q_n/\sqrt{n})^2 \right) \right),
\]

where we used the series expansion of the log function around 1 to write the term \(\log(1 + q_n/n)\) as a series and where in the last equality, we used that \(q_n/n \to 0\) as \(n \to +\infty\) because \(q_n = o(\sqrt{n})\). We conclude by noting that the last term of the equation goes to one as \(n \to +\infty\) since \(q_n = o(\sqrt{n})\).

Lemma 4. Let \((p_n)\) be a sequence in \((0, 1)\) such that \(p_n \to 0\) and \(np_n \to +\infty\) as \(n \to +\infty\). Denoting by \(b(n, p_n)\) a binomial random variable with parameters \((n, p_n)\), and by \(N\) a standard normal random variable, it holds that

\[
\frac{b(n, p_n) - np_n}{\sqrt{np_n}} \Rightarrow N
\]

(30)
as \( n \to +\infty \).

**Proof.** We will prove that the moment generating function of the random variable in the left of (30) converges to the moment generating function of a standard normal random variable. Indeed,

\[
\mathbb{E} \left[ \exp \left( \frac{t(b(n, p_n) - np_n)}{\sqrt{np_n}} \right) \right] = \left(1 - p_n + p_n e^{t/\sqrt{np_n}} \right)^n e^{-t/\sqrt{np_n}}
\]

\[
= \left(1 + \frac{np_n(e^{t/\sqrt{np_n}} - 1)}{n} \right)^n e^{-t/\sqrt{np_n}}
\]

\[
= \left[ \left(1 + \frac{np_n(e^{t/\sqrt{np_n}} - 1)}{n} \right)^n \right] \exp \left( np_n \left( e^{t/\sqrt{np_n}} - 1 \right) \right)
\]

\[
\cdot \exp \left( np_n \left( e^{t/\sqrt{np_n}} - 1 \right) - t/\sqrt{np_n} \right).
\]

Now, by Lemma 3 we have that the term in the square brackets converges to one, since

\[
\frac{np_n(e^{t/\sqrt{np_n}} - 1)}{\sqrt{n}} = np_n \sum_{j \geq 1} \frac{t^j}{j! (np_n)^{j/2}} = t^{\sqrt{n}} + \frac{1}{\sqrt{n}} \sum_{j \geq 0} \frac{t^{j+2}}{(j+2)! (np_n)^{j/2}},
\]

which converges to zero as \( n \to +\infty \). Additionally, using the series expansion of the exponential function we obtain that

\[
\exp \left( np_n \left( e^{t/\sqrt{np_n}} - 1 \right) - t/\sqrt{np_n} \right) = \exp \left( \frac{t^2}{2} + \frac{1}{\sqrt{np_n}} \sum_{j \geq 0} \frac{t^{j+3}}{(j+3)! (np_n)^{j/2}} \right),
\]

which converges to \( \exp(t^2/2) \) since \( np_n \to +\infty \). This concludes the proof.

We are now able to prove Lemma 1.

**Proof of Lemma 1.** To prove part 1, it is sufficient to show that \( b(n, p_n) = 0 \). This holds since \( \mathbb{P}(b(n, p_n) = 0) = (1 - p_n)^n = (1 - np_n/n)^n \sim e^{-np_n} \sim 1 \) as \( n \to +\infty \), where the asymptotic equalities hold due to Lemma 3 and \( np_n \to 0 \).

We now prove part 2. For that purpose define first \( F_n(t) := \mathbb{P}(b(n, p_n) \leq np_n + \sqrt{np_n}t) \) and \( a_n := (k_n - np_n)/\sqrt{np_n} \). Note that part 2. of Lemma 1 is equivalent to

\[
\limsup F_n(a_n) = 0 \quad \text{if and only if} \quad \limsup a_n = -\infty. \tag{31}
\]

To prove the reverse implication of (31) note that for all \( b \) and all sufficiently large \( n \) we have \( F_n(a_n) \leq F_n(b) \), since \( a_n \leq b \) for all sufficiently large \( n \). Taking then \( \limsup \) we obtain by Lemma 4 that \( \limsup F_n(a_n) \leq \Phi(b) \), where \( \Phi \) is the cumulative distribution function of the standard normal distribution. We conclude that \( \limsup F_n(a_n) = 0 \) by making \( b \searrow -\infty \). Now, to prove the direct implication of (31), let \( \limsup F_n(a_n) = 0 \) and assume ad absurdum that \( \limsup a_n = c > -\infty \). Consider then a subsequence \( (a_{n_j}) \) such that \( a_{n_j} \nearrow c \) as \( j \to +\infty \). Then for all \( \epsilon > 0 \) and all sufficiently large \( j \) it holds that

\[
F_n_j(c - \epsilon) \leq F_n_j(a_{n_j}) \leq F_n_j(c), \tag{32}
\]
since $a_{n_j} \in [c - \epsilon, c]$ for all sufficiently large $j$. Making $j \to +\infty$ in (32) we obtain that

$$\Phi(c - \epsilon) \leq \lim \inf F_{n_j}(a_{n_j}) \leq \lim \sup F_{n_j}(a_{n_j}) \leq \Phi(c),$$

where we used that by Lemma 1 the sequence $(F_{n_j})_j$ converges pointwise to $\Phi$. It follows that by taking $\epsilon \searrow 0$ in equation (33) we obtain that $\lim_j F_{n_j}(a_{n_j}) = \Phi(c)$. Thus, in particular, $\Phi(c)$ is an accumulation point of $(F_n(a_n))_n$, which in turn implies $\Phi(c) \leq \lim \sup F_n(a_n)$. But $\Phi(c) > 0$ because $c > -\infty$, and we had assumed $\lim \sup F_n(a_n) = 0$, which is a contradiction. This shows that necessarily $\lim \sup a_n = -\infty$.

References

Asmussen, S. (2003). *Applied Probability and Queues*. Springer New York.

Asmussen, S., & Glynn, P. W. (2007). *Stochastic Simulation: Algorithms and Analysis* volume 57. Springer.

Barrera, J., & Ycart, B. (2014). Bounds for left and right window cutoffs. *ALEA: Latin American Journal of Probability and Mathematical Statistics, 11*, 445–458.

Basu, R., Hermon, J., & Peres, Y. (2015). Characterization of cutoff for reversible markov chains. In *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms* (pp. 1774–1791). Society for Industrial and Applied Mathematics.

Bernhart, G., Fernández, L., Mai, J.-F., Schenk, S., & Scherer, M. (2015). A survey of dynamic representations and generalizations of the Marshall-Olkin distribution. In *Marshall-Olkin Distributions—Advances in Theory and Applications* (pp. 1–13). Springer.

Botev, Z., L’Ecuyer, P., Simard, R., & Tuffin, B. (2016). Static network reliability estimation under the Marshall-Olkin copula. *ACM Transactions on Modeling and Computer Simulation, 26*, 14.

Chen, G.-Y., & Saloff-Coste, L. (2008). The cutoff phenomenon for randomized riffle shuffles. *Random Structures & Algorithms, 32*, 346–374.

Chen, G.-Y., Saloff-Coste, L. et al. (2008). The cutoff phenomenon for ergodic markov processes. *Electronic Journal of Probability, 13*, 26–78.

Cherubini, U., Durante, F., & Mulinacci, S. (2015). *Marshall-Olkin Distributions — Advances in Theory and Applications: Bologna, Italy, October 2013* volume 141. Springer.

David, H. A. (2006). Topics in the history of order statistics. In *Advances in Distribution Theory, Order Statistics, and Inference* (pp. 157–172). Springer.
De Haan, L., & Peng, L. (1999). Exact rates of convergence to a stable law. *Journal of the London Mathematical Society, 59*, 1134–1152.

Diaconis, P., Saloff-Coste, L. et al. (2006). Separation cut-offs for birth and death chains. *The Annals of Applied Probability, 16*, 2098–2122.

Engel, J., Scherer, M., & Spiegelberg, L. (2017). One-factor Lévy-frailty copulas with inhomogeneous trigger rates. In *Soft Methods for Data Science* (pp. 205–212). Springer.

Fernández, L., Mai, J.-F., & Scherer, M. (2015). The mean of Marshall-Olkin-dependent exponential random variables. In *Marshall-Olkin Distributions–Advances in Theory and Applications* (pp. 33–50). Springer.

Finkenstadt, B., & Rootzén, H. (2003). *Extreme values in finance, telecommunications, and the environment*. CRC Press.

Gudendorf, G., & Segers, J. (2010). Extreme-value copulas. In *Copula theory and its applications* (pp. 127–145). Springer.

Hering, C., & Mai, J.-F. (2012). Moment-based estimation of extendible Marshall-Olkin copulas. *Metrika, 75*, 601–620.

Hermon, J., Peres, Y. et al. (2018). On sensitivity of mixing times and cutoff. *Electronic Journal of Probability, 23*.

Kallenberg, O. (2002). *Foundations of Modern Probability*. Springer Science & Business Media.

Lindskog, F., & McNeil, A. J. (2003). Common Poisson shock models: applications to insurance and credit risk modelling. *Astin Bulletin, 33*, 209–238.

Mai, J.-F. (2014). *Multivariate exponential distributions with latent factor structure and related topics*. Ph.D. thesis Technische Universität München.

Mai, J.-F. (2018). Extreme-value copulas associated with the expected scaled maximum of independent random variables. *Journal of Multivariate Analysis, 166*, 50–61.

Mai, J.-F., & Scherer, M. (2009). Lévy-frailty copulas. *Journal of Multivariate Analysis, 100*, 1567–1585.

Mai, J.-F., & Scherer, M. (2013). Sampling exchangeable and hierarchical Marshall-Olkin distributions. *Communications in Statistics-Theory and Methods, 42*, 619–632.

Mai, J.-F., & Scherer, M. (2017). *Simulating copulas: stochastic models, sampling algorithms, and applications* volume 6. World Scientific.
Maller, R., & Schindler, T. (2018). Small time convergence of subordinators with regularly or slowly varying canonical measure. *arXiv preprint arXiv:1806.09763*.

Marichal, J.-L., Mathonet, P., & Waldhauser, T. (2011). On signature-based expressions of system reliability. *Journal of Multivariate Analysis, 102*, 1410–1416.

Marshall, A., & Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association, 62*, 30–44.

Martínez, S., & Ycart, B. (2001). Decay rates and cutoff for convergence and hitting times of Markov chains with countably infinite state space. *Advances in Applied Probability, 33*, 188–205.

Matus, O., Barrera, J., Moreno, E., & Rubino, G. (2018). On the Marshall-Olkin copula model for network reliability under dependent failures. *IEEE Transactions on Reliability*, (pp. 1–11).

Nagaraja, H. N. (2006). Order statistics from independent exponential random variables and the sum of the top order statistics. In *Advances in Distribution Theory, Order Statistics, and Inference* (pp. 173–185). Springer.

Nolan, J. P. (2018). *Stable Distributions - Models for Heavy Tailed Data*. Boston: Birkhauser. In progress, Chapter 1 online at http://fs2.american.edu/jpnolan/www/stable/stable.html.

White, J. M. (2013). Stable distribution sampler in julia. https://gist.github.com/johnmyleswhite/6045390.

Whitt, W. (2002). *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*. Springer Science & Business Media.

Wüthrich, M. V. (2005). Limit distributions of upper order statistics for families of multivariate distributions. *Extremes, 8*, 339–344.

Yuge, T., Maruyama, M., & Yanagi, S. (2016). Reliability of a $k$-out-of-$n$ system with common-cause failures using multivariate exponential distribution. *Procedia Computer Science, 96*, 968–976.