POLITECNICO DI TORINO

Corso di Laurea Magistrale
in Ingegneria Matematica

Tesi di Laurea Magistrale

Poisson random measures: an application to simulations of Lévy and max-stable processes

Relatori
prof. Thomas Mikosch
prof. Enrico Bibbona

Candidato
Simone Chiominto

Anno Accademico 2022-2023

Introduction

This thesis is about simulations of two classes of stochastic processes: Lévy processes and max-stable processes. Let us explain what we mean by simulating a stochastic process. Since a stochastic process X is a random function, hence an infinite-dimensional object, it is a difficult object to be managed by a computer. The aim of a simulation is to find an algorithm that computes the possible values of a realization of X onto a finite number n of points $t_1, \dots, t_n$. This means that the outcome of a sampling algorithm is a random vector, a finite-dimensional object, that we would like to be representative of the whole process. In particular, we would like the random vector $\mathbf{X}_n$ (our simulation) to inherit the statistical properties of X. Then, the most ambitious goal is to compute $\mathbf{X}_n$ in such a way that it has the same distribution of $(X(t_1), \dots, X(t_n))$, the restriction of the stochastic process X onto the evaluation points $t_1, \dots, t_n$. We call this an *exact simulation*.

Finding such algorithm is, in the vast majority of cases, not possible. There are several reasons why. One of the reasons, which we will not consider in the body of the thesis, is that computers do not possess a way of representing exactly real numbers, hence simulations of all processes with values in $\mathbb{R}$ are in practice never exact. A second reason is that it is difficult to find a source of real randomness and in fact, in most cases, simulations depend on a pseudo-random number generator, a deterministic algorithm which produces numbers that in some sense "look random".

Even assuming to possess a theoretical computer which can represent exactly real numbers and that owns a true source of randomness, other problems might arise that make an exact simulation a hard task. Let us explain this through an example. Consider for a moment the simplest case possible: simulating a random variable Y. It is well-known that ${F_Y}^{\leftarrow}(U)$, where U is a uniform random variable in $[0, 1]$ and ${F_Y}^{\leftarrow}$ is the generalized inverse of the cumulative distribution function (cdf) of Y, has the same distribution as Y. Assuming to be able to simulate U, this observation gives us a simple procedure to sample from F_Y. However, there are some highly non-trivial conditions that need to be satisfied in order to get an exact sample through this algorithm. We need to be able to compute explicitly the generalized inverse of F_Y, we need to know explicitly the cumulative distribution function, which might require computing the integral of the probability distribution function (pdf), and, if we do not know directly even the pdf, it might be necessary to compute the Fourier transform of the characteristic function ϕ_Y of Y. This shows how even this fairly simple problem is not as trivial as it seemed at first glance. When we move from the univariate case to the multivariate case, complexity increases because we need to take into account the relationships between the various

components.

Two of the aforementioned problems are the ones we will try to solve in this thesis. In the case of the Lévy processes, the independence of increments let us simplify the problem from the multivariate case (computing sample-paths) to the univariate case (computing the independent increments). Problems arise from the fact that we will assume to only know the characteristic function of increments, not their distribution. Fortunately the distribution of the increments of a Lévy process has a semi-parametric form, depending only on a triplet (c, σ, μ) where c, σ are numbers and μ is a potentially infinite measure, called the Lévy measure. This measure is a sort of mean measure of a Poisson process, in the sense that it tells us the expected number of jumps in a subset of the real numbers in the time interval $[0, 1]$. If there are no jumps the problem is reduced to sampling a normal random variable, which is a problem solved decades ago from Box and Muller in [BM58]. The only problem left is to simulate the jumps from μ. A partial solution to this problem is described in this thesis. In fact, we can find some representations of the jumps in the form of a Poisson process with, in general, infinite number of points. A truncation is needed to make this method feasible for a computer and that will always generate some error. Our strategy will be to find a way of letting the biggest jumps appear first, in such a way that error is minimized. In this case there is no chance of getting an exact simulation, however it is possible to estimate the order of convergence of the error.

In the case of max-stable processes, the problem is different. In fact, simulating the marginal distributions is in general easy, since they belong to three well known distributions (Gumbel, Weibull, Fréchet) with known cumulative distribution functions (and known generalized inverse). The problem here is dealing with the mutual dependence between points of the simulation. Again, our strategy involves finding a representation of the process as a Poisson random measure with an infinite number of points.

Differently from the Lévy case, here we are dealing with component-wise maxima instead of sums, then there is hope that truncation will not generate errors. The problem is that in general it is not possible to know for sure the number of iterations that guarantees our simulation to be exact. In the thesis we show two approaches: the first one is changing measure in such a way that it is possible during the algorithm to know exactly whether a truncation leads to an exact-simulation. A second approach is to compute at each iteration the probability that a truncation leads to an error and stopping the algorithm when this probability is considered small enough.

The thesis is organized as follows:

- In chapter 1 we present a summary of the theory of point processes, Poisson processes and Lévy processes.

- In chapter 2 we address the problem of simulating a Lévy process. We start by describing the easy cases in which exact simulations are possible. Then, we introduce some series representations of a pure jump Lévy process to be used for an approximate sampling algorithm. These methods are then tested on the Gamma process, α-stable process and CGMY process.

- In chapter 3 we first introduce the theory of max-stable processes and then study the simulation approaches for stationary max-stable processes. The methods are then tested on the Brown-Resnick process.

Ringraziamenti

Fare i ringraziamenti in italiano è il modo migliore per far capire che non scrivo male solo in inglese ma anche nella mia lingua madre. Preferisco comuque ringraziare tutte le persone che mi sono state vicine in questi ultimi due anni in italiano per poter essere più sincero e diretto.

Il primo grazie per questa tesi non può che essere al mio relatore all'università di Cophenagen Thomas Mikosch. Lo ringrazio per tutte la conoscenza che mi ha trasmesso, l'attezione nel correggere questa tesi e per i bei momenti extra accademici. Ringrazio molto anche il mio relatore interno al Politecnico di Torino Enrico Bibbona per essere sempre stato disponibile quando più necessario e per aver reso possibile il mio soggiorno a Cophenagen.

Ringrazio tutte le persone con cui ho passato il mio tempo a Cophengen in dipartimento, a casa e fuori. In particolare vorrei citare Luca, Ulisses, Tina, Gabriele, Alexis, Leonardo e Michele. Un grazie particolare anche alla proprietaria della splendida villa dove ho soggiornato, Inger.

Sono da ringraziare anche tutte le persone che hanno condiviso con me il loro tempo a Torino. Ringrazio Francesco per tutti gli stimoli dati e le sempre interessanti discussioni; Thomas per lo splendido periodo vissuto nella stessa casa; Federica e il nostro puntuale appuntamento del Giovedì; Tallo per l'immensa gentilezza e disponibilità. Poi non posso non menzionare Michele, Giuseppe, Pierpaolo, Stefano, Giulia.

Un grazie agli amici di sempre Stefano, Gavriel, Giacomo, Flavio Lorenzo e tutti i miei compagni della triennale (con un ringraziamento speciale per Rachele).

Ringrazio infine la mia famiglia per il supporto costante (anche a distanza) e per aver concretamento reso possibile tutto ciò che ho fatto negli ultimi due anni.

Contents

Chapter 1

Point Processes and Lévy processes

1.1 Point Processes

1.1.1 Definition of point process

Let $E \in \mathbb{R}^n$ be a subset of Euclidean space and $\mathcal{E} \coloneqq \mathcal{B}(\mathcal{E})$ the σ-algebra of Borel sets of E. Given a point $x \in E$, the *Dirac Measure* ε_x on $(E, \mathcal{E})$ is defined as

$$\varepsilon_x(A) \coloneqq \mathbf{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \qquad \forall A \in \mathcal{E}.$$

This measure is the building block of the so-called *counting measures* which formalize the idea of measuring a set by counting the numbers of special points it contains.

Definition 1.1.1.1 (Counting Measure)**.** Let $(x_i)_{i\geq 1}$ be a sequence of points in E. We call $m \colon \mathcal{E} \to \mathbb{N}$ a *counting measure* with atoms $(x_i)_{i\geq 1}$ if for all $A \in \mathcal{E}$

$$m(A) = \sum_{i=1}^{\infty} \varepsilon_{x_i}(A) = \#\{i \geq 1 \mid x_i \in A\}.$$

Moreover, m is a *point measure* if for all K compact $m(K) < \infty$.

More in general a *Radon measure* is a measure m on $(E, \mathcal{E})$ such that for all K compact $m(K) < \infty$. Therefore, we can simply define a point measure as a Radon counting measure.

We would like to define a point process as a random point measure. Following the guideline of the definition of a random variable, a point process is a measurable map from the outcome space $(\Omega, \mathcal{F})$ to a suitable measurable space of point measures. Let us call $M_p(E)$ the set of all point measures and $\mathcal{C}$ the collection of subsets of

$M_p(E)$ of the form $\{m \in M_p(E) \mid m(A) \in B\}$ for all $A \in \mathcal{E}$ and $B \in \mathcal{B}([0,\infty))$[1]. By $\mathcal{M}_p(E) := \sigma(\mathcal{C})$ we denote the smallest σ-algebra containing all sets in $\mathcal{C}$. Equivalently, we can define $\mathcal{M}_p(E)$ as the σ-algebra generated by all evaluation maps $m \mapsto m(A)$ for each $A \in \mathcal{E}$. Given this definition of the measurable space of point measures $(M_p(E), \mathcal{M}_p(E))$, the following definition comes naturally.

Definition 1.1.1.2 (Point Process)**.** A *point process* N is a measurable map from the outcome space $(\Omega, \mathcal{F})$ to the measurable space of point measures $(M_p(E), \mathcal{M}_p(E))$.

The choice of $\mathcal{M}_p(E)$ is justified by the fact that, this way, for each $A \in \mathcal{E}$, $N(A) = N(A;\omega)$ is a random variable. This is expressed in the following lemma.

Lemma 1.1.1.3. *Let $(\Omega, \mathcal{F})$ be the outcome space and $(M_p(E), \mathcal{M}_p(E))$ the space of point measures. $N\colon \Omega \to M_p(E)$ is a point process iff for all $A \in \mathcal{E}$, $N(A)$ is a random variable with values in $\mathbb{N}$ and $N(K) < \infty$ for K compact.*

See [Res87] Proposition 3.1 for the proof.

Then, we want to define a probability distribution on the measurable space of point measures $(M_p(E), \mathcal{M}_p(E))$. Given a probability measure $\mathbb{P}$ on the outcome space $(\Omega, \mathcal{F})$, we can easily define the probability distribution $\mathbb{P}_N$ of a point process N on $(M_p(E), \mathcal{M}_p(E))$ as $\mathbb{P}_N = \mathbb{P} \circ N^{-1}$.

Since point processes are infinite-dimensional objects, it is difficult to create simple intuition of the probability distribution $\mathbb{P}_N$. The collection of the *finite dimensional distributions*

$$\mathbb{P}(N(A_1) = k_1, \ldots, N(A_m) = k_m)$$

for all possible choices of $m < \infty$, $A_i \in \mathcal{E}$ and $k_i = 0, 1, 2, \ldots$ for all $i = 1, \ldots, m$ helps to overcome this problem. In fact, this collection of distributions uniquely defines the distribution $\mathbb{P}_N$. The results follow directly from Corollary 2.2.1 in [Res03].

Remark 1.1.1.4. In the following, we will mostly deal with random point measures of the form

$$N := \sum_{i=1}^{\infty} \varepsilon_{X_i},$$

where $(X_n)_{n\geq 1}$ is a sequence of E-valued random variables such that for each $K \in \mathcal{E}$ compact, $N(K) < \infty$ a.s. In general, these random point measures are not point processes since there might be a K compact and $\omega \in \Omega$ such that $\mathbb{P}_N(N(\cdot;\omega)) = 0$ and $N(K;\omega) = \infty$. However, it is always possible to redefine N on these events of null probability in order to make it a point process. From now on we will always consider this little extension to the precise definition of a point process given above.

[1] In the context of the definition of point processes it is sufficient to consider only B equal to the singletons $\{n\}$ for $n \in \mathbb{N}$. In order to make it easier to generalize the definition we directly consider all Borel subsets of the positive real line.

1.1.2 The Laplace functional and weak convergence of point processes

In order to define the *Laplace functional* of a point process we need to give a precise meaning to the stochastic integral over a point process

$$\int_E g \, \mathrm{d}N. \tag{1.1}$$

If we restrict g to the set of measurable non-negative functions, the integral (1.1) is well-defined path-wise in the sense of the Lebesgue-Stieltjes integral and the map $\omega \mapsto \int_E g \, \mathrm{d}N(\omega)$ is measurable. Therefore, the integral (1.1) is a random variable[2]. Considering a representation of $N = \sum_{i=1}^{\infty} \varepsilon_{X_i}$ we can explicitly write the integral as

$$\int_E g \, \mathrm{d}N = \sum_{i=1}^{\infty} g(X_i).$$

We are ready to define the Laplace functional:

Definition 1.1.2.1 (Laplace functional of a point process)**.** Given a point process N, the *Laplace functional* Ψ_N is the measurable functional from the space of measurable functions $g\colon E \to [0, \infty]$ to $\mathbb{R}$ defined as

$$\Psi_N(g) \coloneqq \mathbb{E}\left[e^{-\int_E g \, \mathrm{d}N} \right] = \int_{M_p(E)} e^{-\int_E g \, \mathrm{d}m} \, \mathbb{P}_N(\mathrm{d}m). \tag{1.2}$$

Observation 1.1.2.2. The Laplace functional Ψ_N is well-defined even if $\int_E g \, \mathrm{d}N(\omega) = \infty$ for some $\omega \in \Omega$.

Observation 1.1.2.3. The Laplace functional uniquely determines the distribution of a point process. In fact, taking a simple function $g = \sum_{i=1}^{m} a_i \mathbf{1}_{A_i}(x)$ where $a_i \geq 0$ and $A_i \in \mathcal{E}$, then

$$\Psi_N(g) = \mathbb{E}\left[e^{-\sum_{i=1}^{m} a_i N(A_i)} \right]$$

is the Laplace transform of $(N(A_1), \ldots, N(A_m))$ which uniquely determines its distribution. Hence, the Laplace functional determines uniquely all finite dimensional distributions and therefore $\mathbb{P}_N$.

The Laplace functional is strongly connected to the concept of *weak convergence* of (the distribution of) point processes. In fact, weak convergence of point processes corresponds to point-wise convergence of the Laplace functional.

[2]The value $\int_E g \, \mathrm{d}N(\omega) = +\infty$ is not excluded, but it has probability 0 on compact sets

Definition 1.1.2.4 (Weak convergence of point processes)**.** Let $(N_n)_{n\geq 1}$, N be point processes on $(E, \mathcal{E})$. We say that N_n *weakly converges* to N (we write $N_n \xrightarrow{d} N$) if

$$(N_n(A_1), \dots, N_n(A_m)) \xrightarrow{d} (N(A_1), \dots, N(A_m))$$

for all bounded sets $A_i \in \mathcal{E}$ such that $\mathbb{P}_N(N(\partial A_i) = 0) = 1$, where $i = 1, \dots, m$.

In order to formulate the equivalence between weak convergence of point processes and point-wise convergence of the Laplace functional we need to restrict the domain of the Laplace functional. Let us call $C_K^+(E)$ the space of non-negative continuous functions with compact support, then the following theorem holds.

Theorem 1.1.2.5. *Let $(N_n)_{n\geq 1}$, N be point processes on $(E, \mathcal{E})$. Then, $N_n \xrightarrow{d} N$ iff for every $g \in C_K^+$ we have $\Psi_{N_n}(g) \to \Psi_N(g)$.*

Proof. See [Res87] Proposition 3.19. □

Observation 1.1.2.6. Considering a fixed $g \in C_K^+$, the integral $\int_E g \, \mathrm{d}N$ is a random variable. Therefore, from $\Psi_{N_n}(g) \to \Psi_{N_n}(g)$ we get $\mathbb{E}\left[e^{-z\int_E g \, \mathrm{d}N}\right] \to \mathbb{E}\left[e^{-z\int_E g \, \mathrm{d}N}\right]$ for all positive z. This is equivalent to $\int_E g \, \mathrm{d}N_n \xrightarrow{d} \int_E g \, \mathrm{d}N$.

1.2 Poisson random measures

We now introduce an important and flexible class of point processes called *Poisson processes* or *Poisson random measures*. In analogy with binomial and Poisson distributions, it is possible to obtain this class of point processes as the weak limit of the so-called *binomial processes*. We will observe this property at the end of the section.

1.2.1 Definition of Poisson process and Poisson integral

Definition 1.2.1.1 (Poisson process or Poisson random measure)**.** Let N be a point process on $(E, \mathcal{E})$ and μ be a Radon measure on E. We call N a *Poisson process* or a *Poisson random measure* with mean measure μ (we write N is a PRM(μ)) if:

1. for each $A \in \mathcal{E}$, we have $N(A) \sim \mathrm{Pois}(\mu(A))$[3];
2. for each $m \geq 2$ and $A_1, \dots, A_m \in \mathcal{E}$ disjoint, $N(A_1), \dots, N(A_m)$ are mutually independent.

[3] If $\mu(A) = 0$ this is equivalent to $N(A) = 0$ a.s., if $\mu(A) = \infty$ this is equivalent to $N(A) = \infty$ a.s.

Observation 1.2.1.2. If $\mu(E) < \infty$ we can represent each Poisson random measure as $N = \sum_{i=1}^{\tau} \varepsilon_{X_i}$ where

- $\tau \sim \text{Pois}(\mu(E))$;
- for $i = 1, \dots, \tau$, $X_i \overset{\text{iid}}{\sim} F := \frac{\mu}{\mu(E)}$ and independent of τ.

This representation of N can be extended to the case $\mu(E) = \infty$ exploiting the Radon property of the measure μ. The idea is to write E as a countable union of disjoint, bounded sets E_i, each with $\mu(E_i) < \infty$. Then, for each i write N restricted to E_i as explained above.

The Laplace Functional of a Poisson Random Measure has a characteristic form.

Proposition 1.2.1.3. *Let N be a PRM(μ). The Laplace functional computed on $g \geq 0$ ans measurable on E is*

$$\Psi_N(g) = \exp\left\{ -\int_E \left(1 - e^{-g(x)}\right) \mu(\mathrm{d}x) \right\}. \tag{1.3}$$

Proof. We write the proof for simple functions, the result for a generic non-negative measurable g follows from standard reasoning in measure theory using monotone convergence. Let g be a simple function

$$g = \sum_{i=1}^{m} a_i \mathbf{1}_{A_i}$$

for $A_1, \dots, A_m \in \mathcal{E}$ disjoint and $a_i \geq 0$. We compute the Poisson integral

$$\int_E g \, \mathrm{d}N = \sum_{i=1}^{m} a_i \int_E \mathbf{1}_{A_i}(x) N(\mathrm{d}x) = \sum_{i=1}^{m} a_i N(A_i),$$

where $N(A_1), \dots, N(A_m)$ are mutually independent because N is a PRM. We use independence to factorize the Laplace functional, which becomes

$$\Psi_N(g) = \mathbb{E}\left[e^{-\sum_{i=1}^{m} a_i N(A_i)} \right] = \prod_{i=1}^{m} \mathbb{E}\left[e^{-a_i N(A_i)} \right]. \tag{1.4}$$

Each factor of (1.4) is the Laplace transform of a Poisson distributed random

variable with mean $\mu(A_i)$ computed at a_i. Therefore, we can write

$$\begin{aligned}\Psi_N(g) &= \prod_{i=1}^{m} \mathbb{E}\left[\, e^{-a_i N(A_i)} \,\right] = \\ &= \prod_{i=1}^{m} e^{-\mu(A_i)(1-e^{-a_i})} = \\ &= \exp\left\{ -\sum_{i=1}^{m} \int_E (1 - e^{-a_i \mathbf{1}_{A_i}(x)}) \mu(\mathrm{d}x) \right\} \\ &= \exp\left\{ -\int_E \left(1 - e^{-g(x)}\right) \mu(\mathrm{d}x) \right\}.\end{aligned}$$

□

From Poisson random measures we can define *Poisson integrals* as special cases of integrals over a point process. Poisson integrals are of particular interest since they inherit some good properties from Poisson random measures. In fact, it is possible to find analytic conditions that ensure Poisson integrals to be a.s. finite. This allows us to extend the definition of Poisson integrals to real valued functions. Other interesting properties are that Poisson integrals of functions with disjoint support are independent and, if two Poisson integrals have 0 correlation, they are independent. For a more in-depth treatment about the topic, see [Mik04] Chapter 7.

Proposition 1.2.1.4. *Let $g \geq 0$ be measurable on E and N be a* PRM(μ). *Then $\int_E g \,\mathrm{d}N$ is a.s. finite iff*

$$\int_E (g(x) \wedge 1)\, \mu(\mathrm{d}x) < \infty.$$

This property allows one to define a generalized Poisson integral over functions that are not restricted to be non-negative. In fact, if g is measurable on E and such that

$$\int_E (|g(x)| \wedge 1)\, \mu(\mathrm{d}x) < \infty$$

then the integral $\int_E g \,\mathrm{d}N$ is a.s. well-defined in the sense of a Lebesgue-Stieltjes integral.

To give an intuitive idea of Poisson integrals, we consider consider the case in which $\mu(E) < \infty$. As seen in Observation 1.2.1.2, we can write $N = \sum_{i=i}^{\tau} \varepsilon_{X_i}$ where $\tau \sim \mathrm{Pois}(\mu(E))$ and for $i = 1, \ldots, \tau$, $X_i \overset{\text{iid}}{\sim} F = \frac{\mu}{\mu(E)}$ random variables independent of τ. Given g we can write the Poisson integral $\int_E g \,\mathrm{d}N = \sum_{i=1}^{\tau} Z_i$ where $Z_i = g(X_i)$

are iid random variables independent of τ with common probability distribution $G := \frac{\mu}{\mu(E)} \circ g^{-1}$.

Poisson integrals inherit some properties from Poisson random measures. Of particular interest is that if $\{g_i\}_{i=1,\dots,m}$ have disjoint support $\{\int_E g_i \, \mathrm{d}N\}_{i=1,\dots,m}$, then they are mutually independent. This property follows from the factorization of the joint Laplace-Stieltjes transform of the random vector $(N(A_1), \dots, N(A_m))$ when $A_1, \dots, A_m \in \mathcal{E}$ are disjoint. We will heavily exploit this property to prove the Lévy-Itô representation of a Lévy process.

As anticipated we now see the weak convergence of binomial processes to Poisson processes.

Example 1.2.1.5 (Weak limit of the binomial process)**.** Let $X_1^{(n)}, \dots, X_n^{(n)}$ be iid random variable with common probability distribution F_n. Then we define the *binomial process* N_n as

$$N_n(A) = \sum_{i=1}^{n} \varepsilon_{X_i}(A) \sim \mathrm{Bin}(n, F_n(A)), \quad A \in \mathcal{B}(\mathbb{R})$$

Then, if we consider a partition $A_1, \dots, A_m$ of $\mathbb{R}$, $m > 1$, $A_i \in \mathcal{B}(\mathbb{R})$ for all $i = 1, \dots, m$, we get that

$$(N_n(A_1), \dots, N_n(A_m)) \sim \mathrm{Mult}(n; F_n(A_1), \dots, F_n(A_m)),$$

where $\mathrm{Mult}(n; p_1, \dots, p_m)$ is the multinomial distribution.

We compute the Laplace functional of N_n. Similarly to the Laplace functional of a PRM we compute the Laplace functional of a binomial process for a simple function, then the result can be extended to the general case using monotone convergence.

Let $m \geq 1$ and $g(x) = \sum_{i=1}^{m} a_i \mathbf{1}_{A_i}(x)$, where $a_i \geq 0$ and $A_i \in \mathcal{B}(\mathbb{R})$ for all $i = 1, \dots, m$. We can assume without loss of generality that $A_1, \dots, A_m$ is a partition of $\mathbb{R}$. Then,

$$\begin{aligned}
\Psi_{N_n}(g) =& \mathbb{E}\left[e^{-\int_{\mathbb{R}} g \, \mathrm{d}N_n} \right] \\
=& \mathbb{E}\left[e^{-\sum_{i=1}^{n} g(X_i)} \right] \\
=& \mathbb{E}\left[e^{-\sum_{j=1}^{m} a_j N_n(A_j)} \right] \\
=& \left(\sum_{i=1}^{n} e^{-a_j} F_n(A_j) \right)^n = \left(\int_{\mathbb{R}} e^{-g(x)} F_n(\mathrm{d}x) \right)^n,
\end{aligned}$$

where in the fourth equality we computed the Laplace transform of a multinomial distribution.

Now, consider a sequence of probability distribution $(F_n)_{n\geq 1}$ and a Radon measure μ such that (nF_n) converges vaguely to μ. Then

$$\begin{aligned}\Psi_{N_n}(g) &= \left(\int_{\mathbb{R}} e^{-g(x)} F_n(\mathrm{d}x)\right)^n \\ &= \left(1 - \int_{\mathbb{R}} (1 - e^{-g(x)}) F_n(\mathrm{d}x)\right)^n \\ &= \left(1 - \frac{1}{n}\int_{\mathbb{R}} (1 - e^{-g(x)})\, nF_n(\mathrm{d}x)\right)^n \to \exp\left\{-\int_E \left(1 - e^{-g(x)}\right)\mu(\mathrm{d}x)\right\},\end{aligned}$$

which is the Laplace functional of a PRM(μ). Hence, $(N_n)_{n\geq 1}$ converge weakly to a Poisson Random measure.

1.2.2 Operations on Poisson random measures

In order to deeply understand how Poisson random measures work it is necessary to see how they perform under a change of the space on which they are defined. It is also interesting to show operations that preserve the properties Poisson processes. In this context we will define the sum of independent PRMs and independent marking of PRMs.

How a Poisson integral behaves under transformations on the space E on which the PRM is defined is well explained by Proposition 2.1 from [Ros01]

Proposition 1.2.2.1. *Let M and N be Poisson point processes on measurable spaces S and T_0, with mean measures μ and ν, respectively. Suppose that T_0 is a Borel subset of some Borel space T and that for some measurable mapping $h : S \to T$,*

$$\nu = \mu \circ h^{-1} \quad \text{on } \mathcal{B}(T_0)$$

Then

$$N \overset{d}{=} M \circ h^{-1}$$

Assume in addition that N is defined on a probability space that is rich enough, that is, there exists on this probability space a uniform Unif$(0,1)$ *random variable that is independent of N, and also that*

$$M = \sum_{i=1}^{\infty} \varepsilon_{S_i}$$

for some S-valued random elements S_i, $i \geq 1$. Then there exists a sequence $(\tilde{S}_i)_{i\geq 1}$ of S-valued random elements defined on the same probability space as N such that

$$(\tilde{S}_i)_{i\geq 1} \overset{d}{=} (S_i)_{i\geq 1}$$

and

$$N = \sum_{i=1}^{\infty} \varepsilon_{h(\tilde{S}_i)} \quad a.s.$$

This property will be of great use in the context of simulations of Poisson processes. In fact, it is sometimes easier to simulate a process initially on a different space and then to transform the points according to our needs.

We conclude this section on Poisson processes describing two operations that preserve the property of being a Poisson process.

Proposition 1.2.2.2 (Aggregation of independent PRMs)**.** *Let $m \in \mathbb{N}$ and $\{N_i\}_{i=1,\dots,m}$ be mutually independent* PRM(μ_i) *for $i = 1, \dots, m$ on $(E, \mathcal{E})$. Then $N := N_1 + \cdots + N_m$ is a* PRM(μ) *with $\mu = \mu_1, \dots, \mu_m$.*

Proposition 1.2.2.3 (Independent Marking of a PRM)**.** *Let $N_X = \sum_{i=1}^{\infty} \varepsilon_{X_i}$ be a* PRM(μ) *on E_1 and $(Y_i)_{i\geq 1} \overset{iid}{\sim} F$ be a sequence of E_2-valued random vectors. If $(X_i)_{i\geq 1}$ and $(Y_i)_{i\geq 1}$ are independent then $N := \sum_{i=1}^{\infty} \varepsilon_{(X_i,Y_i)}$ is a* PRM$(\mu \times F)$ *on $E := E_1 \times E_2$.*

1.3 Lévy Processes

In this section we introduce a flexible class of stochastic processes. Lévy processes are additive processes with stationary and independent increments. This class of processes includes processes with quite different behaviors such as continuous processes such as the Brownian motion and discontinuous processes such as the compound Poisson process. This flexibility has found great use for modelling in applications.

1.3.1 Definition of a Lévy process

We start by defining Lévy processes.

Definition 1.3.1.1 (Lévy Process)**.** An $\mathbb{R}^d$-valued stochastic process $\mathbf{X}$ is called a d-dimensional *Lévy process* if the following conditions are satisfied:

1. it has independent increments, i.e., for all $n \geq 0$ and $0 = t_0 < t_1 < \cdots < t_n$ the random vectors $\mathbf{X}(t_i) - \mathbf{X}(t_{i-1})$ for $i = 1, \dots, n$ are independent;

2. it starts at the origin, i.e. $\mathbf{X}(0) = \mathbf{0}$ a.s.;

3. it has stationary increments, i.e. for all $h, t \geq 0$ and $s > t$, $\mathbf{X}(s) - \mathbf{X}(t) \overset{\mathrm{d}}{=} \mathbf{X}(s + h) - \mathbf{X}(t + h)$

4. it is stochastically continuous, i.e. for any $\varepsilon > 0$, when $t \downarrow 0$ we have $\mathbb{P}(\|\mathbf{X}(s+t) - \mathbf{X}(s)\| > \varepsilon) \to 0$.

It is proved (for example in [Sat99]) that these four conditions ensure that we can define a version of $\mathbf{X}$ such that for each $\omega \in \Omega$, $(\mathbf{X}(t;\omega))_{t\geq 0}$ has right continuous path and well-defined left limits (we write *càdlàg* from the French "*continue à droite, limite à gauche*").

Notable examples of Lévy processes are:

- linear functions: $\mathbf{X}(t) = \mathbf{c}t$, $t \geq 0$;
- the standard Wiener process (or standard Brownian motion): $X(t) = W(t)$
- compound Poisson processes: $\mathbf{X}(t) = \int_{[0,t]\times\mathbb{R}^d} \mathbf{x} N(\mathrm{d}t, \mathrm{d}\mathbf{x})$ where N is a Poisson random measure on $(\mathbb{R} \times \mathbb{R}^d)$

These three examples are of fundamental importance, since any Lévy process can be written as a sort of composition of the three aforementioned processes. This is informally the statement of an important theorem known as the Lévy-Itô representation of Lévy processes.

From now on we restrict ourselves to the treatment of the one-dimensional case. The multidimensional case follows easily, with the downside of more complex and less transparent notation.

1.3.2 Infinitely divisible distribution and Lévy-Itô decomposition

Given a Lévy process X, we study the distribution of $X(t)$ at a fixed time $t \geq 0$. A first notable property of Lévy processes is that the distribution of $X(t)$ is uniquely determined by that of $X(1)$. This is evident from the characteristic function of $X(t)$.

Proposition 1.3.2.1. *Let X be a Lévy process and $t \geq 0$. The characteristic function of $X(t)$ is*

$$\phi_{X(t)}(s) = \mathbb{E}\left[e^{\mathrm{i}\,sX(t)}\right] = \left(\mathbb{E}\left[e^{\mathrm{i}\,sX(1)}\right]\right)^t.$$

Proof. Let us prove the assertion in three steps. First we prove it for $t \in \mathbb{N}$, then for $t \in \mathbb{Q}$, finally for $t \in \mathbb{R}$. The result is trivially true for $t = 0$ and $t = 1$. If $t = \mathbb{N} \setminus \{0, 1\}$, we can write $X(t) = \sum_{i=1}^{t}(X(i) - X(i-1))$, then the characteristic

function is

$$\begin{aligned}\phi_{X(t)}(s) =& \mathbb{E}\left[e^{\mathrm{i}\,sX(t)}\right] \\ =& \mathbb{E}\left[e^{\mathrm{i}\,s\sum_{j=1}^{t}(X(j)-X(j-1))}\right] \\ =& \prod_{j=1}^{t}\mathbb{E}\left[e^{\mathrm{i}\,s(X(j)-X(j-1))}\right] \\ =& \mathbb{E}\left[e^{\mathrm{i}\,sX(1)}\right]^t,\end{aligned}$$

where we use the independence of increments in the third equality and the stationarity of increments of a Lévy Process in the last one.

Consider $t \in \mathbb{Q}$; we write $t = \frac{n}{m}$ for $n, m \in \mathbb{N}$. Then, using a similar argument as above we can write

$$\phi_{X(1)}(s) = \left(\phi_{X\left(\frac{1}{m}\right)}(s)\right)^m$$

and

$$\phi_{X\left(\frac{n}{m}\right)}(s) = (\phi_{X\left(\frac{1}{m}\right)}(s))^n.$$

Hence,

$$\phi_{X\left(\frac{n}{m}\right)}(s) = (\phi_{X(1)}(s))^{\frac{n}{m}} = (\phi_{X(1)}(s))^t.$$

If $t \in \mathbb{R}$ we can find a rational sequence $t_n \downarrow t$ for $n \to \infty$. Then we have trivially

$$\phi_{X(t_n)}(s) = (\phi_{X(1)}(s))^{t_n} \to (\phi_{X(1)}(s))^t \tag{1.5}$$

and

$$\phi_{X(t_n)}(s) = \mathbb{E}\left[e^{\mathrm{i}\,sX(t_n)}\right] = \mathbb{E}\left[e^{\mathrm{i}\,s(X(t_n)-X(t))}\right]\mathbb{E}\left[e^{\mathrm{i}\,sX(t)}\right] \to \mathbb{E}\left[e^{\mathrm{i}\,sX(t)}\right], \tag{1.6}$$

where we used in the second equality the independence of increments and in the third one the stochastic continuity[4]. Comparing (1.5) and (1.6), we get the result. □

From the previous theorem we understand that many properties of the Lévy process X depend on the distribution of $X(1)$. A first general property of this marginal distribution is that for each $m > 0$ we have $X(1) = \sum_{i=1}^{m}(X\left(\frac{i}{n}\right) - X\left(\frac{i-1}{n}\right))$ where $X\left(\frac{i}{n}\right) - X\left(\frac{i-1}{n}\right)$ are iid random variables. This means that the distribution of $X(1)$ must be part of a special class of distributions called *infinitely divisible distributions*.

Definition 1.3.2.2 (Infinitely divisible distributions)**.** The distribution of a random variable Y is said to be *infinitely divisible* if for all $m \geq 0$, there exists iid random variables $Y_1, \ldots, Y_m$ such that $Y \overset{\mathrm{d}}{=} Y_1 + \cdots + Y_m$.

[4] $X(t_n) - X(t) \overset{\mathrm{d}}{=} X(t_n - t) \overset{\mathbb{P}}{\to} 0$, hence it converges in distribution to 0. Then the characteristic function must converge to $e^0 = 1$.

The converse is also true: given an infinitely divisible distribution F we can find a Lévy process X such that $X(1) \sim F$. This means that studying Lévy processes is essentially equivalent to studying the properties of infinitely divisible distributions.

Many well known distributions are infinitely divisible, for instance Poisson, Gamma, Gaussian, α-stable and many other distributions with unbounded support. The most important result about infinitely divisible distributions is the so-called *Lévy-Khintchine formula* which states that the characteristic function of any infinitely divisible distribution must have a specific structure. Before presenting the result, we define a class of measures that will be of great importance for the theory of Lévy processes and infinitely divisible distributions.

Definition 1.3.2.3 (Lévy measure)**.** A measure μ on $\mathbb{R}$ is called a *Lévy measure* on $\mathbb{R}$ if it satisfies the following conditions:

- $\mu(\{0\}) = 0$;
- $\int_{\mathbb{R}}(x^2 \wedge 1)\mu(\mathrm{d}x) < \infty$.

Theorem 1.3.2.4 (Lévy-Khintchine formula)**.** *Let Y be an infinitely divisible random variable, i.e., a random variable with an infinitely divisible distribution. Its characteristic function is given by*

$$\phi_Y(s) = \exp\left\{\mathrm{i}\, sc - \sigma^2\frac{s^2}{2} - \int_{\mathbb{R}}\left(1 - e^{\mathrm{i}\, sx} + \mathrm{i}\, sx\mathbf{1}_{[0,1]}(|x|)\right)\mu(\mathrm{d}x)\right\}, \qquad (1.7)$$

where $c \in \mathbb{R}$, $\sigma \geq 0$, and μ is a Lévy measure on $\mathbb{R}$. The triplet (c, σ, μ) uniquely determines the distribution of Y, and it is called characteristic triplet*.*

Let us now consider a Lévy process X with characteristic function

$$\phi_{X(t)}(s) = \exp\left\{-t\int_{\mathbb{R}}\left(1 - \mathrm{e}^{\mathrm{i}\, sx} + \mathrm{i}\, sx\mathbf{1}_{[0,1]}(|x|)\mu(\mathrm{d}x)\right)\right\}.$$

We want to show that X can be written as the limit of compound Poisson processes. The first problem we need to address is that, given a Lévy measure μ on $\mathbb{R}$, a PRM(Leb $\times\, \mu$) on $[0, +\infty) \times \mathbb{R}$ is in general not well-defined since a Lévy measure μ is not a Radon measure generally. We will overcome this obstacle by changing the topology of $\mathbb{R}$ in such a way that we will consider as bounded sets those that are far from the origin. In some sense we will interchange the role of ∞ and 0. With this topology it makes sense to define N as PRM(Leb $\times\, \mu$) on $[0, +\infty) \times \mathbb{R}$.[5]

Now, we can state the main result.

[5]The new topology is induced by a metric, hence the resulting topological space is Hausdorff, locally compact and with a countable basis. These conditions allow to define a Poisson random measure.

Theorem 1.3.2.5 (Lévy-Itô representation of Jump Lévy Process)**.** *Let N be a* PRM$(Leb \times \mu)$ *on $[0,\infty) \times \mathbb{R}$ where μ is a Lévy measure. We define*

$$c_\delta := \int_{\{\delta<|x|<1\}} x\mu(\mathrm{d}x) \tag{1.8}$$

and

$$S_\delta(t) := -c_\delta t + \int_{[0,t]\times\{\delta\le|x|\le1\}} xN(\mathrm{d}s,\mathrm{d}x) + \int_{[0,t]\times\{|x|>1\}} xN(\mathrm{d}s,\mathrm{d}x) \tag{1.9}$$

1. *The limit process*
$$S(t) := \lim_{\delta\downarrow 0} S_\delta(t) \tag{1.10}$$
exists a.s. for all $t \ge 0$.

2. *S is a Lévy process with characteristic function*
$$\phi_{S(t)}(s) = \exp\left\{-t\int_{\mathbb{R}}\left(1 - e^{\mathrm{i}\,sx} + \mathrm{i}\,sxI_{[0,1]}(|x|)\mu(\mathrm{d}x)\right)\right\}. \tag{1.11}$$

Before proving the theorem we state *Kolmogorov's three series theorem*, a classical result in probability theory that we need in order to prove Theorem 1.3.2.5. A proof for Kolmogorov's three series theorem can be found in [Bil95].

Theorem 1.3.2.6 (Kolmogorov's three series theorem)**.** *Let $(X_n)_{n\ge1}$ be a sequence of mutually independent random variables. Then, the series $\sum_{n=1}^\infty X_n$ converges a.s. if the following conditions hold for some $A > 0$, and only if the following conditions hold for any $A > 0$:*

1. *$\sum_{n=1}^\infty \mathbb{P}(|X_n| \ge A)$ converges;*
2. *$Y_n := X_n\mathbf{1}_{\{|X_n|<A\}}$, then $\sum_{n=1}^\infty \mathbb{E}\left[\,Y_n\,\right]$ converges;*
3. *$\sum_{n=1}^\infty \mathrm{Var}(Y_n)$ converges.*

We also need Lemma 10.2.1 from [Mik04].

Lemma 1.3.2.7 (Lemma 10.2.1 from [Mik04])**.** *Let $(X_n)_{n\ge1}$ be a sequence of Lévy processes on $[0,\infty)$ and X a stochastic process on $[0,\infty)$ such that the finite-dimensional distributions of X_n converge to those of X: for any choice of $0 \le t_1 < \cdots < t_m$ and $m \ge 1$,*

$$(X_n(t_1),\ldots,X_n(t_m)) \xrightarrow{d} (X(t_1),\ldots,X(t_m)).$$

Then X has stationary independent increments and $X(0) = 0$ a.s.

We can now prove Theorem 1.3.2.5.

Proof of Theorem 1.3.2.5. In order to prove the first point we use the Three Series Theorem. Without loss of generality we prove the convergence for $t = 1$. Let $(\delta_n)_{n\in\mathbb{N}}$ be a monotone decreasing sequence such that $\delta_0 = 1$ and $\delta_n \downarrow 0$ and define

$$Z_n = \int_{[0,1]\times\{\delta_n\le|x|\le\delta_{n-1}\}} x\,(N(\mathrm{d}t,\mathrm{d}x) - \mathrm{d}t\,\mu(\mathrm{d}x))$$

The sequence $(Z_n)_{n\ge1}$ is made such that

$$S_{\delta_n}(1) = \sum_{i=1}^{n} Z_i + \int_{[0,t]\times\{|x|>1\}} xN(\mathrm{d}s,\mathrm{d}x)$$

then, the convergence of $S_{\delta_n}(1)$ is equivalent to the convergence of Z_n because $N([0,t]\times[1,\infty]) < \infty$ a.s. We use Kolmogorov's Three Series Theorem considering $A = 1$ on $(Z_n)_{n\ge1}$ and we get

1. $\sum_{n=1}^{\infty}\mathbb{P}(|Z_n|\ge1)$ converges. In fact for all $n \ge 1$, $\mathbb{E}\,[\,Z_n\,] = 0$. Then

$$\mathbb{P}(|Z_n|\ge1) = \mathbb{P}(|Z_n - \mathbb{E}\,[\,Z_n\,]|\ge1) \le \mathrm{Var}(Z_n),$$

where we used Chebyshev's inequality. Finally

$$\begin{aligned}\sum_{n=1}^{\infty}\mathbb{P}(|Z_n|\ge1) &\le \sum_{n=1}^{\infty}\mathrm{Var}(Z_n)\\ &= \sum_{n=1}^{\infty}\int_{\{\delta_n\le|x|\le\delta_{n-1}\}} x^2\,\mu(\mathrm{d}x)\\ &= \int_{\{0\le|x|\le1\}} x^2\,\mu(\mathrm{d}x) < \infty.\end{aligned}$$

2. Call $Y_n := Z_n\mathbf{1}_{\{|Z_n|<1\}}$, then $\sum_{n=1}^{\infty}\mathbb{E}\,[\,Y_n\,]$ converges. In fact

$$0 = \mathbb{E}\,[\,Z_n\,] = \mathbb{E}\left[\,Z_n\mathbf{1}_{\{|Z_n|<1\}} + Z_n\mathbf{1}_{\{|Z_n|\ge1\}}\,\right],$$

then

$$|\mathbb{E}\,[\,Y_n\,]| = \left|\mathbb{E}\left[\,Z_n\mathbf{1}_{\{|Z_n|\ge1\}}\,\right]\right|.$$

We get

$$\begin{aligned}\sum_{n=1}^{\infty}\mathbb{E}\left[Y_n\right] &\leq \sum_{n=1}^{\infty}\left|\mathbb{E}\left[Y_n\right]\right| \\ &\leq \sum_{n=1}^{\infty}\left|\mathbb{E}\left[Z_n\mathbf{1}_{\{|Z_n|\geq 1\}}\right]\right| \\ &\leq \sum_{n=1}^{\infty}\mathbb{E}\left[|Z_n|\,\mathbf{1}_{\{|Z_n|\geq 1\}}\right] \\ &\leq \sum_{n=1}^{\infty}\mathbb{E}\left[|Z_n|^2\right] \\ &= \sum_{n=1}^{\infty}\mathrm{Var}(Z_n) \leq \infty.\end{aligned}$$

3. $\sum_{n=1}^{\infty}\mathrm{Var}(Y_n)$ converges since $\sum_{n=1}^{\infty}\mathrm{Var}(Z_n)$ converges.

Now we prove the second point. Since for all t, $S_\delta(t)$ converges to $S(t)$ a.s., it converges also in distribution in the sense of the finite-dimensional distributions. Then, we can use Lemma 1.3.2.7 and say that S has independent stationary increments and $S(0) = 0$ a.s.

Then, we compute the characteristic function of $S(t)$ for a fixed $t \geq 0$. From the independence of Poisson integrals with disjoint support we have

$$\begin{aligned}\phi_{S_\delta(t)}(s) =&\mathbb{E}\left[\exp\left\{-t\int_{\{\delta<|x|\leq 1\}}(1-e^{\mathrm{i}\,sx}+\mathrm{i}\,sx)\,\mu(\mathrm{d}x)\right\}\right] \\ &\times\mathbb{E}\left[\exp\left\{-t\int_{\{|x|\geq 1\}}(1-e^{\mathrm{i}\,sx})\,\mu(\mathrm{d}x)\right\}\right] \\ =&\mathbb{E}\left[\exp\left\{-t\int_{\{|x|>\delta\}}(1-e^{\mathrm{i}\,sx}+\mathbf{1}_{[0,1]}(|x|)\;\mathrm{i}\,sx)\,\mu(\mathrm{d}x)\right\}\right]\end{aligned}$$

which converges to (1.11) for $\delta \downarrow 0$.

We only need to prove the stochastic continuity. From (1.11) we have

$$\phi_{S(t)}(s) = (\phi_{S(1)}(s))^t \to 1, \quad t \downarrow 0,$$

hence $S(t) \xrightarrow{\mathrm{d}} 0$ which is a constant, therefore $S(t) \xrightarrow{\mathbb{P}} 0$ as $t \downarrow 0$.

□

It is important to remark the necessity of centering S_δ with the term tc_δ. It is necessary since in general for a Lévy measure we have that, from Proposition 1.2.1.4, the term

$$\int_{[0,t]\times\{\delta<|x|<1\}} xN(\mathrm{d}s,\mathrm{d}x)$$

explodes a.s. as $\delta \downarrow 0$. If μ satisfies $\int_{\mathbb{R}}(|x| \wedge 1)\mu(\mathrm{d}x) < \infty$ centering is not needed.

A simple corollary of Theorem 1.3.2.5 is the generalization of the Lévy-Itô representation to a general Lévy process.

Theorem 1.3.2.8. *Let X be a Lévy process. Then,*

1. *we can write*
$$X(t) = ct + \sigma B(t) + S(t) \tag{1.12}$$
where $c \in \mathbb{R}$, $\sigma > 0$, B is the standard Brownian motion, S is defined as in (1.10) and B, S are independent.

2. *The decomposition (1.12) is uniquely determined by the characteristic triplet (c, σ, μ) of the infinitely divisible distribution of $X(1)$ of the Lévy-Khintchine formula (1.7).*

This theorem allows us to interpret a Lévy process as the sum of a deterministic drift, a Brownian (and therefore continuous a.s.) component and a jump component. Therefore, we define a *pure jump Lévy process* as a Lévy process with characteristic triplet $(c, 0, \mu)$ which means that there is no Brownian component.

It is worth noticing that when the centering procedure is needed to define the limit of compound Poisson processes, even if the value c of the characteristic triplet is uniquely defined, it loses the meaning of slope of the linear drift. In fact, when computing the compensation constant c_δ in (1.8) we arbitrarily restrict the domain to the set $[-1, 1]$, but we could have chosen every other compact neighborhood of the origin equivalently. This different choice would have changed the value c, hence its interpretation as the mean slope of the process is meaningless

1.3.3 Some examples of Lévy Processes

We have already considered some important examples of Lévy processes: the linear drift, the standard Brownian motion and the Poisson compound process, which, as shown in Theorem 1.3.2.8, are the building blocks of any Lévy process. We now consider some other examples and we compute explicitly their characteristic triplet (c, σ, μ).

Example 1.3.3.1 (Gamma Lévy process)**.** Consider a random variable Y which is distributed following a Gamma(α, β) distribution, where α is the shape parameter and β the rate parameter. Its characteristic function is given by

$$\phi_Y(s) = \left(1 - \frac{\mathrm{i}\, s}{\beta}\right)^{-\alpha}.$$

We see immediately that Gamma(α, β) is an infinitely divisible random variable. Indeed, for each $m > 0$ we can write

$$\left(1 - \frac{\mathrm{i}\, s}{\beta}\right)^{-\alpha} = \left(\left(1 - \frac{\mathrm{i}\, s}{\beta}\right)^{-\frac{\alpha}{m}}\right)^{m},$$

hence Y has the same distribution of the sum of m independent random variable, each one Gamma$(\frac{\alpha}{m}, \beta)$ distributed.

Since Gamma(α, β) is an infinitely divisible distribution, we can define a Lévy process X which has $X(1) \sim$ Gamma(α, β). We call it a *Gamma Process*. In order to write the characteristic triplet, we notice that the characteristic function of a Gamma(α, β) random variable can be written equivalently as

$$\phi_{X(1)}(s) = \left(1 - \frac{\mathrm{i}\, s}{\beta}\right)^{-\alpha} = \exp\left\{-\alpha \int_0^\infty (1 - e^{\mathrm{i}\, sx}) \frac{e^{-\beta x}}{x}\, \mathrm{d}x\right\}$$

Then, we let μ be

$$\mu(\mathrm{d}x) = \frac{e^{-\beta x}}{x}\, \mathrm{d}x,$$

and we compute c as

$$c = \alpha \int_0^1 \mathrm{e}^{-\beta x}\, \mathrm{d}x = \frac{\alpha}{\beta}(1 - e^{-\beta x}).$$

The characteristic triplet (c, σ, μ) of a Gamma Lévy process is:

$$\begin{aligned} c &= \frac{\alpha}{\beta}(1 - e^{-\beta x}); \\ \sigma &= 0; \\ \mu(\mathrm{d}x) &= \frac{e^{-\beta x}}{x}\, \mathrm{d}x. \end{aligned}$$

Example 1.3.3.2 (α-stable process)**.** Consider for $0 < \alpha < 2$ the collection of α-stable distributions $S_\alpha(\sigma, \beta, \tau)$, defined by their characteristic functions

$$\begin{aligned} \phi_Y(s) &= \exp\left\{-\sigma\, |s|^\alpha \left(1 - \mathrm{i}\, \beta \tan\left(\frac{\pi\alpha}{2}\right)\mathrm{sgn}(z)\right) + \mathrm{i}\, \tau s\right\} \qquad \text{for } \alpha \neq 1, \\ \phi_Y(s) &= \exp\left\{-\sigma\, |s| \left(1 - \mathrm{i}\, \beta \frac{2}{\pi}\mathrm{sgn}(z) \log(|z|)\right) + \mathrm{i}\, \tau s\right\} \qquad \text{for } \alpha = 1, \end{aligned}$$

where $\sigma > 0$ is the scale parameter, $\beta \in [-1, 1]$ is the skewness parameter and τ the translation parameter. The infinite divisibility of α-stable distributions is

evident for *symmetric α-stable distributions*, i.e. when $\beta = \tau = 0$. In this particular case the characteristic function is

$$\phi_Y(s) = e^{-\sigma^\alpha |s|^\alpha},$$

therefore, if $Y \sim S_\alpha(\sigma, 0, 0)$ and $m > 0$ we can write Y as sum of m iid random variables with distribution $S_\alpha(\sigma\, m^{-1/\alpha}, 0, 0)$.

An *α-stable process* is a Lévy process X in which $X(1)$ has an $S_\alpha(\sigma, \beta, 0)$ distribution and its characteristic triplet is $(c, 0, \mu_\alpha)$, where the Lévy measure is given by

$$\mu_\alpha(\mathrm{d}x) := \begin{cases} C_+ x^{-(\alpha+1)}\,\mathrm{d}x & x > 0 \\ C_- |x|^{-(\alpha+1)}\,\mathrm{d}x & x < 0 \end{cases}$$

with

$$C_\pm := C_\alpha \frac{1 \pm \beta}{2}\sigma^\alpha, \qquad C_\alpha := \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\frac{\alpha}{2})}. \tag{1.13}$$

Example 1.3.3.3 (Tempered stable process)**.** A modification of an α-stable process with exponentially light tails is the *tempered stable process*. It has characteristic triplet $(0, 0, \mu)$ where the Lévy measure is given by

$$\mu(\mathrm{d}x) := \begin{cases} C_+\, e^{-Mx}\, x^{-(Y+1)}\,\mathrm{d}x & x > 0 \\ C_-\, e^{Gx}\, |x|^{-(Y+1)}\,\mathrm{d}x & x < 0 \end{cases}$$

for C, G, $M > 0$, $0 \le Y < 2$.

If $Y > 0$ and $C_+ = C_- := C$, the corresponding Lévy process is called the *CGMY process* from the initials of its creators Carr, German, Madan, Yor [Car+02].

Example 1.3.3.4 (Inverse Gaussian process)**.** Let B be the standard Brownian motion, we define the stopping time[6]

$$\tau_s := \inf\{t > 0 : B(t) + bt > s\}$$

as the first time that the Brownian motion with drift b crosses s. Given the continuity property of the Brownian motion we have $B(\tau_s) + b\tau_s = s$ a.s.. This process has a.s. non-decreasing paths, thus it is an example of *subordinators*, i.e. Lévy processes with a.s. non-decreasing paths. The characteristic triplet is

$$\begin{aligned} c &= -\frac{2}{b}\int_0^b \frac{e^{y^2/2}}{\sqrt{2\pi}}\,\mathrm{d}y; \\ \sigma &= 0; \\ \mu(\mathrm{d}x) &= \frac{1}{\sqrt{2\pi x^3}} e^{-\frac{b^2 x}{2}}\,\mathrm{d}x \quad \text{on } x > 0. \end{aligned}$$

[6] with respect to the filtration generated by $B(t)$ with $t \le s$

For a given s we can also write an analytic expression for the pdf of τ_s:

$$\mu_s(\mathrm{d}x) = \frac{s}{\sqrt{2\pi x^3}} \exp\left\{ sb - \frac{1}{2}\Big(\frac{s^2}{x} + b^2 x\Big)\right\}.$$

Chapter 2

Simulation of Lévy processes

We start by considering approaches that are taken to generate sample paths of two very simple Lévy processes: the Brownian motion and the homogeneous Poisson process.

Let B be a standard Brownian motion and $t_n^h := nh$ be the points of an equi-spaced grid on the time axis. We call B_n^h a sample of B at time t_n^h. Since increments of the Brownian motion are independent, simulations of paths is based on generating the Gaussian increments

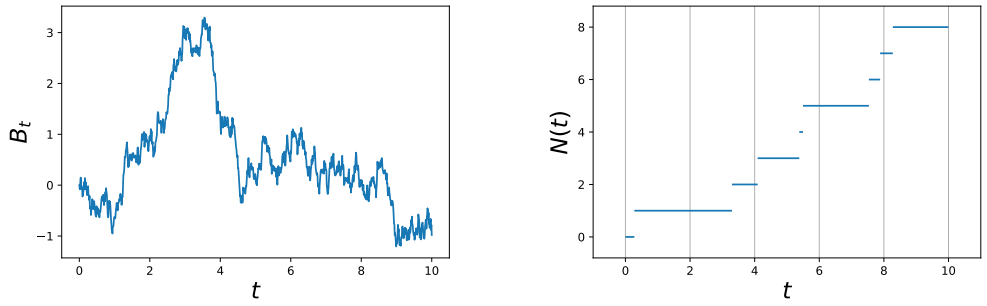
$$\Delta_n^h B := B_n^h - B_{n-1}^h \sim \mathcal{N}(0, h)$$

for which there exist efficient algorithms such as the *Box-Muller* algorithm [BM58]. Then, the value B_n^h is computed as the sum of all the increments. This method provides an exact simulation of the Brownian motion at discrete time points.

The approach taken when computing the homogeneous Poisson process N with rate λ is completely different. We are in fact interested in generating the exact time at which a jump occurs. There exist two possible methods:

- sample the interarrivals times T_i directly as iid exponential random variables with rate λ;
- if we are only interested in simulating N until a given time T, Observation 1.2.1.2 suggests a sampling algorithm. First, we can sample $N(T) \sim \text{Pois}(\lambda T)$, then we generate $N(T)$ jump times Γ_i as iid random variables with common distribution $\text{Unif}(0, T)$.

In the following, we approach simulations of a Lévy process similarly to simulations of Brownian motion, namely we consider a grid of time points (usually equi-spaced) in which we want to simulate the value of a path of the Lévy process. Therefore, we seldom consider the problem of exactly identifying the time of a (big) jump, even if it is sometime stressed the possibility of doing that. Given a



(a) Brownian motion (b) Poisson process

Figure 2.1: Sample paths of standard Brownian motion and homogeneous Poisson process

grid, the generation of the paths of a Lévy process only requires to sample the increments. Then, the problem of simulation of Lévy processes became equivalent to the problem of sampling from an infinitely divisible distribution. This, depending on the properties of the distribution, can be more or less challenging.

In order to understand when sampling is more or less difficult, let us make a classification of infinitely divisible distributions depending on their Lévy measure μ.

- $\mu(\mathbb{R}) < \infty$. In this case the jump part of the Lévy process is a compound Poisson process $S(t) = \int_{[0,t]\times\mathbb{R}} xN(\mathrm{d}s, \mathrm{d}x)$, where N is a PRM(Leb $\times$ μ). Simulations of these processes are very easy if it is easy to sample from $\frac{\mu}{\mu(\mathbb{R})}$. We talk about this case in Subsection 2.1.3.

- $\int_{\mathbb{R}} (|x| \wedge 1)\,\mathrm{d}x < \infty$ but μ is not finite. Here Lévy paths have infinitely many jumps but, despite this, the jump part of the process still has finite variation paths which helps with the convergence of the methods.

- $\int_{\mathbb{R}} (|x| \wedge 1)\,\mathrm{d}x = \infty$. In this situation Lévy paths have infinitely many jumps and the jump part of the process has infinite variation path.

Even if distributions of the second and third kind are in general more difficult to sample from, there are some lucky cases in which generation of random variates from them if fairly easy. Examples are distributions of which we know explicitly the pdf such as the Gamma or the Lognormal distribution. There exist other examples in which, although there is no analytic expression of the distribution, there are some specific algorithms that make sampling easy. This is the case for α-stable distributions.

2.1 Sampling algorithms

This section deals with some sampling algorithms used for generating random variates from infinitely divisible distributions, hence for simulation of Lévy process paths. The first few algorithms are used for some lucky distributions in which sampling is fairly easy and efficient. The last class of algorithms is thought for dealing with the general case.

2.1.1 Explicit marginal distribution

The first class of algorithms is about sampling from infinitely divisible distributions of which we know an analytic expression of the distribution. Many well-known distributions belong to this category, for instance Normal, Gamma, Log-normal, Inverse Gaussian, Student's t, geometric, etc....

For many of these distributions there exist ad-hoc algorithms, but even if that is not the case it is always possible to resort to *inversions* or *Acceptance-Rejection* methods. An overview of these methods can be found at [AG07] Chapter 2.

2.1.2 α-stable distribution

Sampling from α-stable distributions looks like a non-trivial task. In fact, there exists an analytic form of the pdf only for symmetric distribution with $\alpha = 1$ (the Cauchy distribution) and when $\alpha = \frac{1}{2}$ and $\beta = \pm 1$ (Lévy distribution). However, an efficient sampling algorithm for any distribution $S_\alpha(1, \beta, 0)$ for all values of $\alpha \in (0, 2)$ and $\beta = [-1, 1]$ has been invented by Chambers, Mallows and Stuck [CMS76]. The algorithm for $Y \sim S_\alpha(1, \beta, 0)$ is the following:

- sample $U \sim \text{Unif}(-\frac{\pi}{2}, \frac{\pi}{2})$ and an independent $W \sim \text{Exp}(1)$
- If $\alpha \neq 1$ compute

$$Y = \left(1 + \zeta^2\right)^{\frac{1}{2\alpha}} \frac{\sin(\alpha(U + \xi))}{\cos(U)^{\frac{1}{\alpha}}} \left(\frac{\cos(U - \alpha(U + \xi))}{W}\right)^{\frac{1-\alpha}{\alpha}},$$

  else if $\alpha = 1$ compute

$$Y = \frac{1}{\xi}\left[\left(\frac{\pi}{2} + \beta U\right)\tan(U) - \beta \log\left(\frac{\frac{\pi}{2} W \cos(U)}{\frac{\pi}{2} + \beta U}\right)\right],$$

  where

$$\zeta := -\beta \tan\left(\frac{\pi\alpha}{2}\right)$$

and

$$\xi := \begin{cases} \frac{1}{\alpha}\arctan(-\zeta) & \text{for } \alpha \neq 1 \\ \frac{\pi}{2} & \text{for } \alpha = 1. \end{cases}$$

For $Y \sim S_\alpha(\sigma, \beta, \mu)$ it is sufficient to generate from the preceding algorithm an $X \sim S_\alpha(1, \beta, 0)$ and define Y as:

$$Y = \begin{cases} \sigma X + \mu & \text{for } \alpha \neq 1 \\ \sigma X + \frac{2}{\pi}\beta\sigma\log(\sigma) + \mu & \text{for } \alpha = 1. \end{cases}$$

Observation 2.1.2.1. The Chambers-Mallows-Stuck algorithm for $\alpha = 2$ and $\beta = 0$ reduces to the *Box-Muller* algorithm for sampling from Normal distributions.

2.1.3 Finite Lévy measure

If the Lévy measure of a Lévy process is finite, the jump part reduces to a compound Poisson process. In fact, call S the pure jump part of a Lévy process with a finite Lévy measure μ and N be a PRM(Leb $\times\, \mu$), for each $t > 0$, then we can write

$$S(t) = \int_{[0,t]\times\mathbb{R}} xN(\mathrm{d}s, \mathrm{d}x) = \sum_{i=1}^{m} Z_i$$

where $m \sim \text{Pois}(t \times \mu(\mathbb{R}))$ and for $Z_i \overset{\text{iid}}{\sim} F := \frac{\mu}{\mu(\mathbb{R})}$ independent of m.

If we want to simulate S until a fixed time $T > 0$, a straightforward procedure is the following:

1. sample $m \sim \text{Pois}(T \cdot \mu(\mathbb{R}))$;
2. for $i = 1, \ldots, m$ sample the jump times $U_i \overset{\text{iid}}{\sim} \text{Unif}(0, T)$[1];
3. for $i = 1, \ldots, m$ sample the jump sizes $Z_i \overset{\text{iid}}{\sim} F$;
4. for any $t \leq T$ compute $S(t) = \sum_{i=1}^{m} Z_i \mathbf{1}_{\{U_i \leq t\}}$.

It is also interesting to see the procedure for a general Lévy process X with both a Brownian component B and a compound Poisson component S. There are two possible approach. The first is to compute the sample path on a predefined time grid $\{t^h\}_{h=1,\ldots,m}$, that we assume equi-spaced, as follows:

1. simulate B as explained at the beginning of the chapter;

[1]That the distribution of U_i is uniform conditional to m follows from Observation 1.2.1.2

2. sample the increments of the Poisson compound process

$$S(t_n^h) - S(t_{n-1}^h) = \int_{t_{n-1}^h}^{t_n^h} \int_{\mathbb{R}} xN(\mathrm{d}s, \mathrm{d}x) \overset{\mathrm{d}}{=} \int_0^{t_1^h} \int_{\mathbb{R}} xN(\mathrm{d}s, \mathrm{d}x)$$

using the preceding algorithm;

3. compute X on the grid as the sum of the Brownian and the compound Poisson component.

Using this approach, the information about the precise times in which a jump occurs is lost. The second method deals with this problem enriching the (equispaced) grid with the jump times and computing the increments of the Brownian motion also on the enriched grid. A more extended explanation of this procedure is in Section 3.5 of [Gla04].

2.1.4 Series representation of a Lévy process

There exists a class of sampling methods which exploits the representation of a Lévy process in form of a series. For using these methods we will only assume that the characteristic triplet (c, σ, μ) of an infinitely divisible distribution is known, hence we know how to write its characteristic function in the form given by the Lévy-Khintchine formula (1.7). We discuss this class of methods in the next section.

2.2 Series representation of a Lévy process

In this section we will discuss a range of methods which involve writing a pure jump Lévy process $X = \{X(t) \mid t \in [0,1]\}$ as an infinite sum of random variables. This is done by exploiting the property that the pure jump Lévy process S with characteristic triplet $(0, 0, \mu)$ can be seen as limit of compound Poisson processes

$$S_{1/n}(t) = -c_{1/n}t + \int_{[0,t]\times\{\frac{1}{n}\le|x|\le 1\}} xN(\mathrm{d}s, \mathrm{d}x) + \int_{[0,t]\times\{|x|\ge 1\}} xN(\mathrm{d}s, \mathrm{d}x)$$

where $c_{1/n}$ and $S_{1/n}$ are defined in (1.8) and in (1.9) respectively, and N is a (generalized) PRM(Leb $\times$ μ) that can be written explicitly as

$$N = \sum_{i=1}^{\infty} \varepsilon_{(U_i, J_i)}, \tag{2.1}$$

where, for $i \ge 1$, $U_i \overset{\text{iid}}{\sim} \mathrm{Unif}(0,1)$ is the i-th jump time and $(J_i)_{i\ge 1}$ are iid random variables independent of the jump times representing the jump sizes. Each pure

jump Lévy process X with characteristic triplet $(c, 0, \mu)$ can be written as the a.s. limit of

$$X_n(t) := c\,t + S_{1/n}(t)$$

which can be written more explicitly using (2.1) as

$$X_n(t) = \sum_{i \in \Lambda_n} J_i\, \mathbf{1}_{\{U_i \le t\}} - t\, b_n$$

where $\Lambda_n := \{i \ge 1 \mid |J_i| \ge n^{-1}\}$ is the random subset of indices whose jump sizes are greater than n^{-1} in absolute value and $b_n := c_{1/n} - c$. This representation becomes useful for simulation when we replace the random set Λ_n with a nonrandom set of indices of the form $\{1, \dots, m\}$, and write X as the series

$$X(t) = \lim_{m \to \infty} \sum_{i=1}^{m} (J_i \mathbf{1}_{\{U_i \le t\}} - t c_i) \quad a.s.,$$

where $(c_i)_{i \ge 1}$ is a suitable sequence of centers.

Let us discuss some ways of sampling the sequence $(J_i)_{i \ge 1}$, the problem of computing the sequence of centers $(c_i)_{i \ge 1}$ will be discussed later. From now on we will use the following notation

- $(U_i)_{i \ge 1}$ is an iid sequence of $\mathrm{Unif}(0, 1)$ random variables
- $(\Gamma_i)_{i \ge 1}$ is the sequence of the partial sums of iid standard exponential random variables which are independent of $(U_i)_{i \ge 1}$.

Proposition 1.2.2.1 will be used in all of the following methods, and we stick to its notation. However, this leads to an ambiguity on the letter μ, which might refer to both the Lévy measure and the mean measure of M. Thus, we decided to denote here as $\mathcal{Q}$ the Lévy measure of the distribution from which we want to sample.

2.2.1 Inverse Lévy measure method

Suppose the Lévy measure $\mathcal{Q}$ is concentrated on $(0, \infty)$. We define the *inverse Lévy measure* as the function $\mathcal{Q}^{\leftarrow}$ such that

$$\mathcal{Q}^{\leftarrow}(\gamma) := \inf\{x > 0 \mid \mathcal{Q}([x, \infty)) < \gamma\}, \quad \gamma > 0.$$

Call $M := \sum_{i=1}^{\infty} \varepsilon_{(U_i, \Gamma_i)}$ a PRM(μ) on the space $S = [0, 1] \times (0, \infty)$ where $\mu = \mathrm{Leb} \times \mathrm{Leb}$. Let $T = [0, 1] \times (0, \infty)$ and $h : S \to T$ be the function defined by

$$h(u, \gamma) := (u, \mathcal{Q}^{\leftarrow}(\gamma)).$$

Write $J_i = \mathcal{Q}^{\leftarrow}(\Gamma_i)$. From Proposition 1.2.2.1 we get

$$N = \sum_{i=1}^{\infty} \varepsilon_{h(U_i, \Gamma_i)} = \sum_{i=1}^{\infty} \varepsilon_{(U_i, J_i)} \quad \text{a.s.}$$

2.2.2 Thinning method

Let $F \ll \mathcal{Q}$ be a probability distribution on $\mathbb{R}$ and, for $i \geq 0$, $V_i \overset{\text{iid}}{\sim} F$ independent of $(U_i)_{i\geq 1}$ and $(\Gamma_i)_{i\geq 1}$. Consider a PRM(μ) $M := \sum_{i=1}^{\infty} \varepsilon_{(U_i,\Gamma_i,V_i)}$ on the space $S = [0,1] \times (0,\infty) \times \mathbb{R}$ where $\mu = \text{Leb} \times \text{Leb} \times F$. Set $T = [0,1] \times [0,\infty)$ and $h : S \to T$ the function defined by

$$h(u,\gamma,v) := (u, H(\gamma,v)),$$

where

$$H(\gamma,v) = \begin{cases} v & \text{if } \frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}F}(v) \geq \gamma \\ 0 & \text{otherwise} \end{cases}$$

or equivalently $H(\gamma,v) = v\mathbf{1}_{\{\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}F}(v)\geq\gamma\}}$. Then, renaming $J_i = H(\Gamma_i, V_i)$, from Proposition 1.2.2.1 we get

$$N = \sum_{i=1}^{\infty} \varepsilon_{h(U_i,\Gamma_i,V_i)} = \sum_{i=1}^{\infty} \varepsilon_{(U_i,J_i)} \quad \text{a.s.}$$

which is a PRM(ν) on S with $\nu = \text{Leb} \times \mathcal{Q}$.

2.2.3 Rejection method

Let X_0 be a Lévy process on $\mathbb{R}$ such that its Lévy measure $\mathcal{Q}_0$ satisfies

$$\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathcal{Q}_0} \leq 1.$$

We represent the (generalized) Poisson Process associated to X_0 as

$$N_0 = \sum_{i=1}^{\infty} \varepsilon_{(U_i,J_i^0)}$$

for a suitable choice of iid random variables $(J_i)_{i\geq 1}$ independent of $(U_i)_{i\geq 1}$. We define a PRM(μ) M on the space $S = [0,1] \times \mathbb{R} \times [0,1]$ where $\mu = \text{Leb} \times \mathcal{Q}_0 \times \text{Leb}$. Then, for $T = [0,1] \times \mathbb{R}$ and $h : S \to T$ we define

$$h(u,j^0,v) := (u, H(j^0,v)),$$

where

$$H(j^0,v) := \begin{cases} j^0 & \text{if } \frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathcal{Q}_0}(j^0) \geq v \\ 0 & \text{otherwise} \end{cases}$$

or equivalently $H(j^0, v) := j^0 \mathbf{1}_{\{\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathcal{Q}_0}(j^0) \geq v\}}$. Now, for $i \geq 1$ we set $V_i \overset{\text{iid}}{\sim} \text{Unif}(0,1)$ independent of $(U_i)_{i\geq 1}$ and $(J_i^0)_{i\geq 1}$. We rename $J_i = H(J_i^0, V_i)$. From Proposition 1.2.2.1 we get

$$N = \sum_{i=1}^{\infty} \varepsilon_{h(U_i, J_i^0, V_i)} = \sum_{i=1}^{\infty} \varepsilon_{(U_i, J_i)} \quad \text{a.s.}$$

which is a PRM(ν) on S with $\nu = \text{Leb} \times \mathcal{Q}$.

2.2.4 Shot noise method

The following method is based on the idea of writing an infinitely divisible distribution as the marginal distribution of a (generalized) shot noise process. A shot noise process can be imagined as a stochastic process where the present value of the process is computed as the sum of effects originating from a discrete set of past events and where the effect of an event in past is a decreasing function of time. This means that the more time passes from the moment of the event, the less is its effect on the present time. To write this idea in mathematical form we need the concept of probability kernel.

Definition 2.2.4.1 (Probability kernel)**.** Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be two measurable spaces. A map $\sigma : X \times \mathcal{B} \to [0,1]$ is called a *probability kernel* if

1. for each $B \in \mathcal{B}$, $x \mapsto \sigma(x, B)$ is $\mathcal{A}$-measurable;
2. for each $x \in X$, $\sigma(x, \cdot)$ is a probability measure on $(Y, \mathcal{B})$.

Suppose now to write the Lévy measure $\mathcal{Q}$ as

$$\mathcal{Q}(A) = \int_0^{\infty} \sigma(r; A)\, \mathrm{d}r, \quad A \in \mathcal{B}(\mathbb{R}),$$

where σ is a probability kernel. Let $(J_i)_{i\geq 1}$ be a sequence of conditionally independent random variables given $(\Gamma_i)_{i\geq 1}$ such that

$$\mathbb{P}(J_i \in A \mid (\Gamma_i)_{i\geq 1},\, (J_j)_{j\neq i}) = \sigma(\Gamma_i, A), \quad A \in \mathcal{B}(\mathbb{R}).$$

If we interpret J_i as the effect on the present time of an event i happened Γ_i units of time ago, this means that the probability of this effect being of a certain amount depends only on the time it occurred. We define a PRM(μ) M on the space $S = [0,1] \times (0,\infty) \times \mathbb{R}$ where

$$\mu([0,t],[0,\gamma],A) = \int_{u\in[0,t]} \int_{r\in[0,\gamma]} \sigma(r; A)\, \mathrm{d}r\, \mathrm{d}u, \quad t, \gamma > 0,\ A \in \mathcal{B}(\mathbb{R}).$$

Then, for $T = [0,1] \times \mathbb{R}$ we define $h : S \to T$

$$h(u, \gamma, j) \coloneqq (u, j)$$

From Proposition 1.2.2.1 we get

$$N = \sum_{i=1}^{\infty} \varepsilon_{h(U_i,\Gamma_i,J_i)} = \sum_{i=1}^{\infty} \varepsilon_{(U_i,J_i)} \quad \text{a.s.}$$

which is a PRM(ν) on T with $\nu = \text{Leb} \times \mathcal{Q}$.

Coherently with all the previous methods we can write $J_i = H(\Gamma_i, V_i)$ for a suitable function H and a collection of iid random variables $(V_i)_{i\geq 1}$ independent of $(\Gamma_i)_{i\geq 1}$ and $(U_i)_{i\geq 1}$. In order to show this we state the following Lemma from [Kal02]

Lemma 2.2.4.2 (Lemma 2.22 from [Kal02])**.** *Let σ be a probability kernel from a measurable space S to a Borel space T. Then, there exists some measurable function $H : S \times [0,1] \to T$ such that if $V \sim \text{Unif}(0,1)$, then $H(r, V)$ has distribution $\sigma(r, \cdot)$ for every $r \in S$.*

The explicit expression of H is $H(r, v) = \sup\{x \in [0,1] \mid \sigma(s, [0, x]) < t\}$ for all $r \in S$ and $v \in [0,1]$.

Observation 2.2.4.3. All the previous methods can be written in the form of a shot noise. For instance:

- the inverse Lévy measure method $\sigma(r; A) = \varepsilon_{\mathcal{Q}^{\leftarrow}(r)}(A)$;
- the thinning method $\sigma(r; \mathrm{d}x) = F(\mathrm{d}x)\mathbf{1}_{\{\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}F}(x) > r\}}$.

2.2.5 Bondesson's method

The following method is a particular form of the shot noise method where the probability kernel is taken as $\sigma(r; A) = F(\frac{A}{g(r)})$ where F is a probability distribution on $\mathbb{R}$ and $g : [0, \infty) \to [0, \infty)$ is a non-increasing function. It is one of the earliest series representation method in literature, it was proposed by Bondesson in [Bon82].

Set $V_i \overset{\text{iid}}{\sim} F$ for $i \geq 0$ independent of $(\Gamma_i)_{i\geq 1}$ and $(U_i)_{i\geq 1}$. Consider a PRM(μ) $M \coloneqq \sum_{i=1}^{\infty} \varepsilon_{(U_i,\Gamma_i,V_i)}$ on the space $S = [0,1] \times (0, \infty) \times \mathbb{R}$ where $\mu = \text{Leb} \times \text{Leb} \times F$. If we define $T = [0,1] \times [0, \infty)$ and $h : S \to T$ as the function defined by

$$h(u, \gamma, v) \coloneqq (u, g(\gamma)v).$$

then, renaming $J_i \coloneqq g(\Gamma_i)V_i$, from Proposition 1.2.2.1 we get

$$N = \sum_{i=1}^{\infty} \varepsilon_{h(U_i,\Gamma_i,V_i)} = \sum_{i=1}^{\infty} \varepsilon_{(U_i,J_i)} \quad \text{a.s.}$$

which is a PRM(ν) on S where $\nu = \text{Leb} \times \mathcal{Q}$.

2.2.6 How to choose the centering sequence

At the beginning of the section we said that when the Lévy process is of infinite variation we need to center the series with a suitable sequence of $(b_n)_{n\geq 1}$ such that the series representation of a Lévy process

$$X_n(t) = \sum_{i=1}^{n} J_i \, \mathbf{1}_{\{U_i \leq t\}} - t\, b_n$$

converges a.s. to $X(t)$. Here we quickly state a possible choice of this sequence.

First we notice, that in order to address this problem we can restrict ourselves to the shot noise representation since all the other methods can be represented as a shot noise. Given this assumption, the centering sequence proposed in [Ros01] is

$$b_n \coloneqq A(n) - A(n-1)$$

where A is defined as

$$A(s) \coloneqq \int_0^s \int_{|x|\leq 1} x\sigma(r; \mathrm{d}x)\, \mathrm{d}r.$$

This choice has in general nothing in common with c_δ in (1.8). The only exception is the inverse Lévy method and its variants where

$$A(s) = \int_{\mathcal{Q}^{\leftarrow}(s)}^{1} x\mathcal{Q}(\mathrm{d}x) = c_{\mathcal{Q}^{\leftarrow}(s)},$$

while for all the other methods it is more difficult to find an equivalent characterization of $A(s)$. We also notice that the sequence of $(b_i)_{i\geq 1}$ has the good property that it does not depend on the realization of the random sequences $(\Gamma_i)_{i\geq 1}$, $(U_i)_{i\geq 1}$ and $(V_i)_{i\geq 1}$. Hence, it can be computed once before all simulations and be used for the generation of all the random variables.

2.2.7 Speed of convergence

In order to make the series representation useful for simulations, we need to truncate the series at a finite value n. To have an idea of what is a good value of n for simulation purposes we need to predict what might be the value of a chosen error criteria.

We first discuss the problem when X is of finite variation. Without loss of generality we study the error at time 1, thus we change the notation, denoting $X(1)$ as X. The value we want to compute is

$$\mathbb{E}\left[\,|\Delta_n X|\,\right] = \mathbb{E}\left[\,|X - X_n|\,\right] = \mathbb{E}\left[\left|\sum_{i=n+1}^{\infty} J_i\right|\right]$$

in the context of noise shot method this becomes equivalent to computing

$$\mathbb{E}\left[\left|\int_{\Gamma_n}^{\infty}\int_{\mathbb{R}} x\sigma(r;\mathrm{d}x)\,\mathrm{d}r\right|\right] \leq \mathbb{E}\left[\int_{\Gamma_n}^{\infty}\int_{\mathbb{R}} |x|\,\sigma(r;\mathrm{d}x)\,\mathrm{d}r\right]$$

From the right-hand side we notice that the problem can be studied without loss of generality on subordinators, where we have

$$\mathbb{E}\left[\,\Delta_n X\,\right] = \mathbb{E}\left[\int_{\Gamma_n}^{\infty}\int_{\mathbb{R}} x\sigma(r;\mathrm{d}x)\,\mathrm{d}r\right],$$

but since the problem of computing the expectation is usually analytically infeasible we need some estimates. The tool we need is the following standard result from probability theory.

Theorem 2.2.7.1 (Law of iterated logarithm)**.** *Let $(X_n)_{n\geq 1}$ be a sequence of iid random variables with $\mathbb{E}[X_n] = 0$ and $\mathrm{Var}(X_n) = 1$. Let $S_n := X_1 + \cdots + X_n$. Then*

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2n\log\log n}} = 1 \quad a.s. \tag{2.2}$$

See [Pet95] for the proof. Let consider the variable $\Gamma_n = E_1 + \cdots + E_n$ where E_i are standard exponential random variables. We can use theorem 2.2.7.1 on $S_n := \Gamma_n - n = (E_1 - 1) + \cdots + (E_n - 1)$. Then we can say that for n sufficiently big there exist $\varepsilon > 0$ such that

$$n - (1+\varepsilon)\sqrt{2n\log\log n} \leq \Gamma_n \leq n + (1+\varepsilon)\sqrt{2n\log\log n} \quad \text{a.s.} \tag{2.3}$$

Since $e(n) := \int_n^{\infty}\int_{\mathbb{R}} x\sigma(r;\mathrm{d}x)\,\mathrm{d}r$ is a monotone non-increasing function of n we can write

$$e(n + (1+\varepsilon)\sqrt{2n\log\log n}) \leq e(\Gamma_n) \leq e(n - (1+\varepsilon)\sqrt{2n\log\log n}) \quad \text{a.s.}$$

Then, we can compute a superior and inferior bound on the order of convergence.

The speed of convergence depends much on the distribution and on the method used to compute to represent the infinitely divisible random variable. Some examples and numerical tests are provided at the end of the chapter.

In the case of infinite variation the error cannot be computed this way, since the value of $\mathbb{E}[\,\Delta_n X\,]$ does not converge absolutely. However, it is possible to compute the mean squared error

$$\mathbb{E}\left[\,(X - X_n)^2\,\right] = \mathbb{E}\left[\int_{\Gamma_n}^{\infty}\int_{\mathbb{R}} x^2\sigma(r;\mathrm{d}x)\,\mathrm{d}r\right].$$

Using the same argument as before we can estimate $e_2(\Gamma_n) := \int_{\Gamma_n}^{\infty}\int_{\mathbb{R}} x^2\sigma(r;\mathrm{d}x)\,\mathrm{d}r$ as

$$e_2(n + (1+\varepsilon)\sqrt{2n\log\log n}) \leq e_2(\Gamma_n) \leq e_2(n - (1+\varepsilon)\sqrt{2n\log\log n}) \quad \text{a.s.}$$

2.3 Dealing with small jumps

Let us now consider the case of a pure jump Lévy process J. Any method that use the series representation for the simulation of Lévy processes requires a truncation. For instance, when using the *inverse Lévy measure* and its variations, truncating at a value n means that we do not consider jumps that are in absolute value smaller than $\varepsilon(n) := \mathcal{Q}^{\leftarrow}(\Gamma_n)$, we denote this approximation $J_n^{(0)}$.

A possible refining strategy for this approximation is to add to $J_n^{(0)}$ the expected value $\mu_{\varepsilon(n)}$ of the sum of all the (infinite) jumps smaller than $\varepsilon(n)$. Which means that if we call

$$\mu_\varepsilon := \int_{\{|x|<\varepsilon\}} x\mathcal{Q}(\mathrm{d}x),$$

the new approximation is $J_n^{(1)} = J_n^{(0)} + t\mu_{\varepsilon(n)}$.

A more refined proposal is to approximate small jumps with a Brownian Motion with drift $\mu_{\varepsilon(n)}$ and variance $\sigma^2_{\varepsilon(n)} := \int_{\{|x|<\varepsilon(n)\}} x^2\mathcal{Q}(\mathrm{d}x)$. To make this proposal mathematically rigorous we need to prove that

$$\frac{J - J_n^{(0)} - t\mu_{\varepsilon(n)}}{\sigma_{\varepsilon(n)}} \xrightarrow{\mathrm{d}} B, \tag{2.4}$$

where B is a standard Brownian motion. The problem here is that for $n \to \infty$ the variance goes to 0 and we might need some control over the speed of its convergence. In [AR01] it is shown that the condition $\sigma_{c\sigma_\varepsilon \wedge \varepsilon} \sim \sigma_\varepsilon$ when $\varepsilon \to 0$ for all $c > 0$ is sufficient and necessary for (2.4). In absence of atoms near 0 this is equivalent to the more intuitive condition

$$\frac{\varepsilon}{\sigma_\varepsilon} \to 0, \tag{2.5}$$

i.e., σ_ε goes to 0 slower than ε. This condition is verified, for instance, by α-stable and CGMY distributions, but not by Gamma distributions. Based on this we define a new possible approximation of J as

$$J_n^{(2)} = J_n^{(0)} + t\mu_{\varepsilon(n)} + \sigma_{\varepsilon(n)}B.$$

2.4 Three examples of series representation

In this section we discuss simulations of Gamma random variables, α-stable random variables and CGMY random variables using the series representation of Lévy processes. When possible we discuss the speed of convergence and we show that in the case of α-stable and CGMY distributions the speed convergence depends on the parameter α (Y in the case of CGMY), hence this method might be competitive only for very small value of those parameters.

2.4.1 Gamma distribution

The first example we provide is the representation of Gamma random variables which has characteristic triplet $(0, 0, \mathcal{Q})$, where

$$\mathcal{Q}(\mathrm{d}x) = a\frac{e^{x/b}}{x},$$

a is the shape parameter and b is the rate parameter of a Gamma distribution.

In [Ros01] there are provided four different representations. The first uses the *inverse Lévy measure method*, we have

$$X(t) = \sum_{i=1}^{\infty} bE_1^{-1}\left(\frac{\Gamma_i}{a}\right)\mathbf{1}_{\{U_i \le t\}}$$

where E_1^{-1} is the inverse of the *exponential integral function*

$$E_1(x) = \int_1^{\infty} \frac{e^{-tx}}{t}.$$

The second representation is based on the *Bondesson's method* using $V_i \sim \mathrm{Exp}(1)$ and $g(\gamma) = e^{-\frac{\gamma}{a}}$. The representation is the following

$$X(t) = \sum_{i=1}^{\infty} be^{\frac{\Gamma_i}{a}}\, V_i\, \mathbf{1}_{\{U_i \le t\}}$$

The third representation is based on the *thinning method* using $V_i \sim \mathrm{Exp}(b)$, hence $F(\mathrm{d}x) = \frac{1}{b}e^{\frac{x}{b}}$ and the Radon-Nicodym derivative is

$$\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}F}(x) = \frac{ba}{x}.$$

Then, calling

$$\{\tau_1 \le \tau_2 \le \cdots\} = \left\{ i \ge 1 \mid \frac{ba}{V_i} \le \Gamma_i \right\}$$

we get the representation

$$X(t) = \sum_{i=1}^{\infty} V_{\tau_i}\, \mathbf{1}_{\{U_{\tau_i} \le t\}}.$$

The fourth representation is based on the *rejection method* using as Lévy measure

$$\mathcal{Q}_0(\mathrm{d}x) = \frac{ab}{x(x+b)}$$

from which we can sample using the inverse Lévy measure method with

$$\mathcal{Q}_0{}^{\leftarrow}(\gamma) = \frac{b}{e^{\frac{\gamma}{a}} - 1}.$$

The jumps are selected using the Radon-Nicodym derivative

$$\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathcal{Q}_0}(x) = e^{-\frac{x}{b}}\big(1 + \frac{x}{b}\big)$$

then, calling

$$\{\tau_1 \leq \tau_2 \leq \cdots\} = \left\{ i \geq 1 \mid e^{-\frac{J_i^0}{b}}\big(1 + \frac{J_i^0}{b}\big) \leq V_i \right\}$$

we get the representation

$$X(t) = \sum_{i=1}^{\infty} J_{\tau_i}^0 \, \mathbf{1}_{\{U_{\tau_i} \leq t\}}.$$

The Gamma Lévy process is a subordinator; it is easy to discuss the order of convergence of the error for two of the former methods.

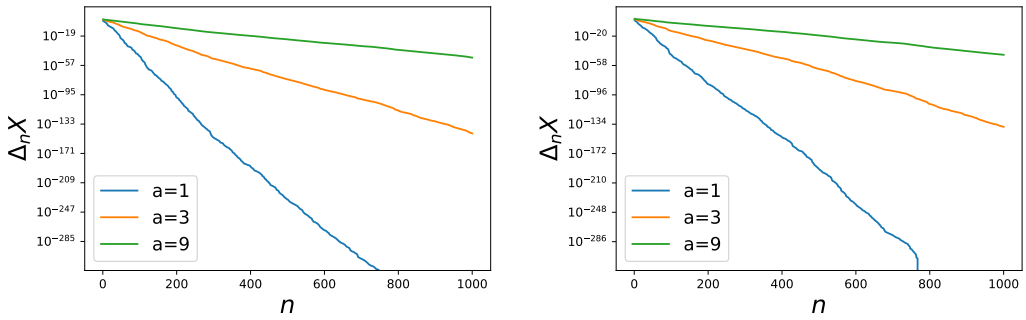
In the case of *Bondesson's method*, calling $\gamma_n = n - (1+\varepsilon)\sqrt{2n \log\log n}$, we have for n sufficiently big and for $\varepsilon > 0$

$$\begin{aligned}
\mathbb{E}\,[\,\Delta_n X\,] &\leq \int_{\gamma_n}^{\infty} \int_0^{\infty} x F\left(\frac{\mathrm{d}x}{g(\gamma)}\right) \mathrm{d}\gamma \\
&= \int_{\gamma_n}^{\infty} \int_0^{\infty} x \frac{1}{b} e^{-\frac{x}{b} e^{\frac{\gamma}{a}}} \,\mathrm{d}x \,\mathrm{d}\gamma \\
&= b \int_{\gamma_n}^{\infty} \int_0^{\infty} y e^{-y e^{\frac{\gamma}{a}}} \,\mathrm{d}y \,\mathrm{d}\gamma \\
&= b \int_{\gamma_n}^{\infty} e^{-\frac{\gamma}{a}} \int_0^{\infty} e^{-y e^{\frac{\gamma}{a}}} \,\mathrm{d}y \,\mathrm{d}\gamma \\
&= b \int_{\gamma_n}^{\infty} e^{-\frac{\gamma}{a}} \,\mathrm{d}\gamma \\
&= abe^{-\frac{\gamma_n}{a}} \sim \mathcal{O}(e^{-\frac{n-(1+\varepsilon)\sqrt{2n\log\log n}}{a}}),
\end{aligned}$$

so we get exponential convergence.

In the case of the *rejection method* we have

$$\begin{aligned}
\mathbb{E}\,[\,\Delta_n X\,] &\leq \int_0^{\mathcal{Q}_0{}^{\leftarrow}(\gamma_n)} x \mathcal{Q}(\mathrm{d}x) \\
&= \int_0^{\frac{b}{e^{\gamma_n/a} - 1}} a e^{\frac{x}{b}} \,\mathrm{d}x \\
&= \frac{a}{b}(e^{\frac{1}{(e^{\gamma_n/a} - 1)}} - 1) \sim \mathcal{O}(e^{-\frac{n-(1+\varepsilon)\sqrt{2n\log\log n}}{a}}),
\end{aligned}$$



(a) Bondesson's Method (b) Rejection Method

Figure 2.2: Semilog-y plot of the error for Bondesson's and Rejection method for Gamma random variables

which is again an exponential convergence. The exponential order of convergence and the dependence on the shape parameter a is shown in figure 2.2. It is evident in figure 2.3 how well Bondesson's method approximates the distribution even for small value of n.

2.4.2 α-stable distribution

The Lévy measure of α-stable[2] processes have a very simple form, this property allows us to find a series representation based on a simple variation of the *inverse Lévy measure method*. Consider an α-stable Lévy process X with Lévy Measure

$$\mathcal{Q}(\mathrm{d}x) = \left(C_+ \mathbf{1}_{\{x>0\}} + C_- \mathbf{1}_{\{x<0\}}\right) |x|^{-(\alpha+1)}\, \mathrm{d}x$$

where C_+ and C_- where defined in (1.13). We call ν the Lévy measure on $(0, +\infty)$ defined as

$$\nu(\mathrm{d}x) := (C_+ + C_-)\, x^{-(\alpha+1)}\, \mathrm{d}x = C_\alpha\, x^{-(\alpha+1)}$$

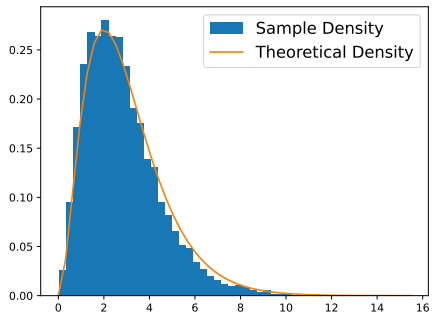
and we consider the inverse Lévy measure

$$\nu^{\leftarrow}(\gamma) = \left(\frac{\alpha\gamma}{C_\alpha}\right)^{-\frac{1}{\alpha}}.$$

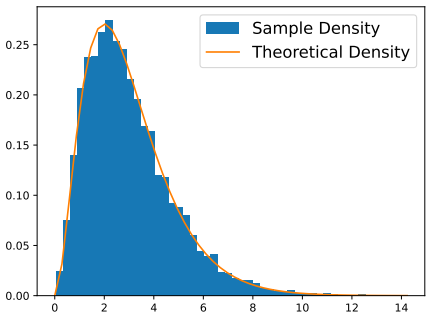
Then, when $\alpha \in (0,1)$, which corresponds to the finite variation case, or $\beta = 0$, which corresponds to the symmetric case, a series representation is

$$X(t) = \sum_{i=1}^{\infty} \nu^{\leftarrow}(\Gamma_i)\, V_i\, \mathbf{1}_{\{U_i<t\}} = \left(\frac{C_\alpha}{\alpha}\right)^{\frac{1}{\alpha}} \sum_{i=1}^{\infty} \Gamma_i^{-\frac{1}{\alpha}}\, V_i\, \mathbf{1}_{\{U_i<t\}}$$

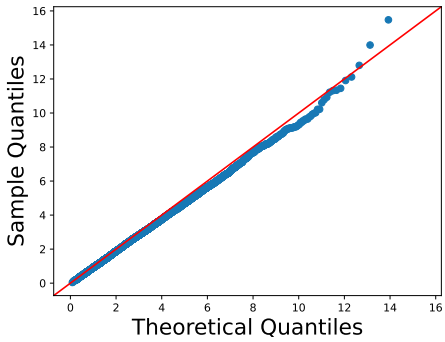
[2]In this subsection we restrict without loss of generality to $S_\alpha(1, \beta, 0)$.



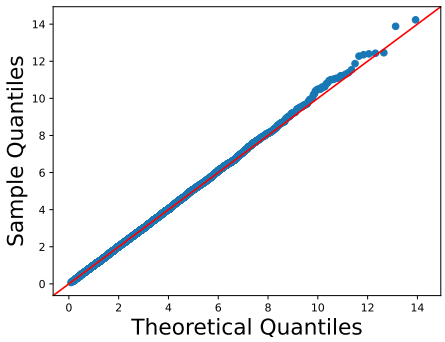
(a) Density $n = 10$



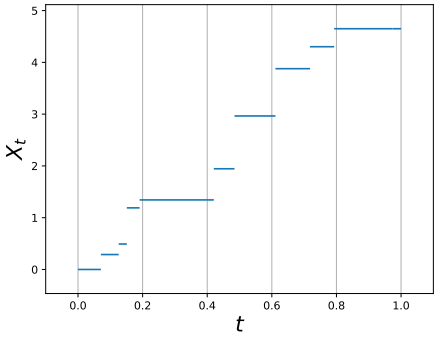
(b) Density $n = 1000$



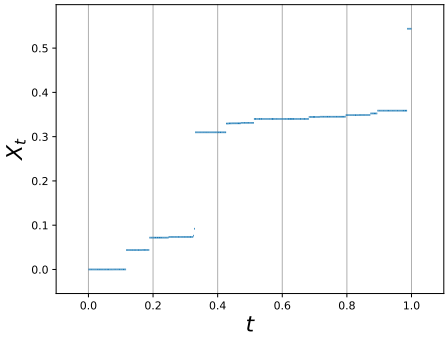
(c) Q-Q plot $n = 10$



(d) Q-Q plot $n = 1000$



(e) Lévy path $n = 10$



(f) Lévy path $n = 1000$

Figure 2.3: Simulations of a Gamma Lévy process X using Bondesson's Series representation truncated at $n = 10$ (left column) and $n = 1000$ (right column). The first row represents the sample density of 10000 samples of $X(1)$, the second row the Q-Q plots and the third one example of sample paths.

where $(V_i)_{i\geq 1}$ are iid random variable independent of $(\Gamma_i)_{i\geq 1}$ and $(U_i)_{i\geq 1}$ such that $\mathbb{P}(V_i = \pm 1) = \frac{1\pm\beta}{2}$.

If $\alpha \in [1,2)$ and $\beta \neq 0$ we need a compensation. The series representation becomes

$$X(t) = \sum_{i=1}^{\infty} \left\{ \left(\frac{C_\alpha}{\alpha}\right)^{\frac{1}{\alpha}} \Gamma_i^{-\frac{1}{\alpha}} V_i \, \mathbf{1}_{\{U_i<t\}} - b_i \right\}$$

with $b_i \coloneqq A(i) - A(i-1)$ and

$$A(n) = \begin{cases} \frac{C_+ - C_-}{1-\alpha} \left[1 - \left(\frac{\alpha\, n}{C_\alpha}\right)^{\frac{1-\alpha}{\alpha}} \right] & \text{if } \alpha \in (1,2) \\ \frac{C_+ - C_-}{\alpha} \log\left(\frac{\alpha\, n}{C_\alpha}\right) & \text{if } \alpha = 1. \end{cases}$$

When $\alpha \in (0,1)$ we can study the order of convergence of the mean absolute error. Calling as before $\gamma_n = n - (1+\varepsilon)\sqrt{2n \log\log n}$, we see that

$$\begin{aligned} \mathbb{E}\left[\, |\Delta_n X| \,\right] &\leq \left| \int_{-\nu^{\leftarrow}(\gamma_n)}^{\nu^{\leftarrow}(\gamma_n)} x \mathcal{Q}(\mathrm{d}x) \right| \leq \int_{-\nu^{\leftarrow}(\gamma_n)}^{\nu^{\leftarrow}(\gamma_n)} |x| \, \mathcal{Q}(\mathrm{d}x) \\ &= \int_0^{\nu^{\leftarrow}(\gamma_n)} x \, \nu(\mathrm{d}x) = C_\alpha \int_0^{\left(\frac{\alpha\gamma_n}{C_\alpha}\right)^{-\frac{1}{\alpha}}} x^{-\alpha} \, \mathrm{d}x \\ &= \frac{C_\alpha^{1/\alpha}}{1-\alpha} (\alpha\gamma_n)^{(1-\frac{1}{\alpha})} \sim \mathcal{O}(n^{(1-\frac{1}{\alpha})}). \end{aligned}$$

Thus, we see that for $\alpha = \frac{1}{2}$ we have a linear convergence, for $\alpha = \frac{1}{3}$ quadratic convergence and so on. We can see this behavior in figure 2.4.

In figure 2.5 we can see the slow convergence in the case of $\alpha = 0.9$. Even with $n = 100000$ the simulations are very far from being accurate.

When $\alpha \in (0,2)$ it is possible to study the mean squared error and with similar computations as before we get

$$\begin{aligned} \mathbb{E}\left[\, \Delta_n X^2 \,\right] &\leq \int_{-\nu^{\leftarrow}(\gamma_n)}^{\nu^{\leftarrow}(\gamma_n)} x^2 \mathcal{Q}(\mathrm{d}x) \\ &= \int_0^{\nu^{\leftarrow}(\gamma_n)} x^2 \, \nu(\mathrm{d}x) = C_\alpha \int_0^{\left(\frac{\alpha\gamma_n}{C_\alpha}\right)^{-\frac{1}{\alpha}}} x^{-\alpha+1} \, \mathrm{d}x \\ &= \frac{C_\alpha^{2/\alpha}}{2-\alpha} (\alpha\gamma_n)^{(1-\frac{2}{\alpha})} \sim \mathcal{O}(n^{(1-\frac{2}{\alpha})}). \end{aligned}$$

We can see this behavior in figure 2.6.

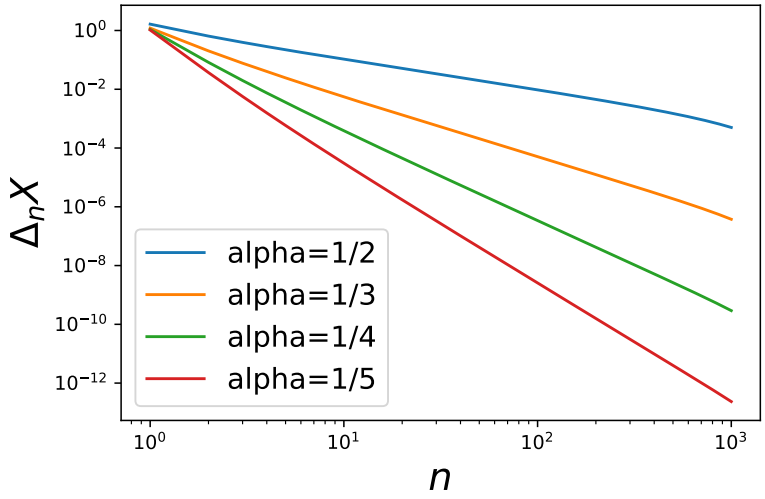


Figure 2.4: Loglog plot of the error for α-stable random variables

For an α-stable variable it is also possible to use the heuristics of approximating small jumps with Brownian motion. We need to compute the values of $\mu_{\nu^{\leftarrow}(n)}$ (only in the case of finite variation) and $\sigma^2_{\nu^{\leftarrow}(n)}$:

$$\mu_{\nu^{\leftarrow}(n)} = \int_0^{(\frac{\alpha n}{C_\alpha})^{-\frac{1}{\alpha}}} (C_+ - C_-)\, x^{-\alpha}\, \mathrm{d}x = \frac{C_+ - C_-}{1-\alpha} \left(\frac{\alpha n}{C_\alpha}\right)^{1-\frac{1}{\alpha}},$$
$$\sigma^2_{\nu^{\leftarrow}(n)} = \int_0^{(\frac{\alpha n}{C_\alpha})^{-\frac{1}{\alpha}}} C_\alpha\, x^{-\alpha+1}\, \mathrm{d}x = \frac{C_\alpha}{2-\alpha} \left(\frac{\alpha n}{C_\alpha}\right)^{1-\frac{2}{\alpha}}.$$

We can see the effect of the correction in figure 2.7 for $X \sim S_{1.9}(1,1,0)$ and 2.8 for $X \sim S_{0.9}(1,1,0)$ in both cases with $n = 10$.

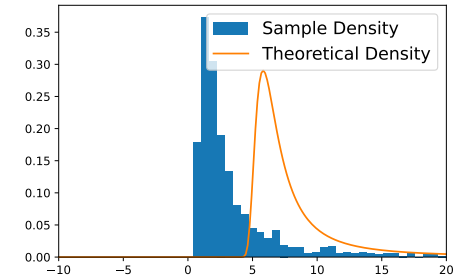
2.4.3 CGMY distribution

The Lévy measure of a CGMY distribution is

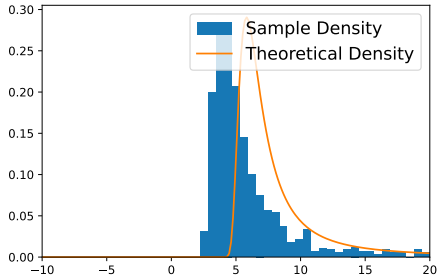
$$\mathcal{Q}(\mathrm{d}x) = C(\mathbf{1}_{\{x>0\}} e^{-Mx} + \mathbf{1}_{\{x<0\}} e^{Gx})\, |x|^{-(Y+1)}\, \mathrm{d}x$$

and can be exploited for a rejection method representation using as $\mathcal{Q}^0$ the Lévy measure of a symmetric α-stable distribution with $\alpha = Y$ and $C_\alpha = \frac{C}{2}$. The Radon-Nikodym derivative is

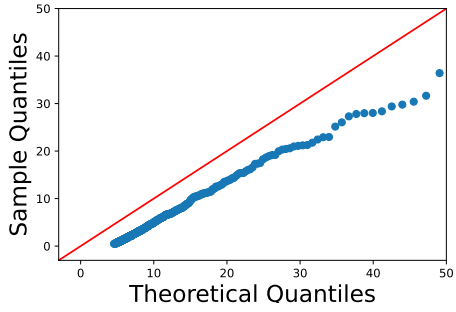
$$\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathcal{Q}^0}(x) = \mathbf{1}_{\{x>0\}} e^{-Mx} + \mathbf{1}_{\{x<0\}} e^{Gx} \leq 1.$$



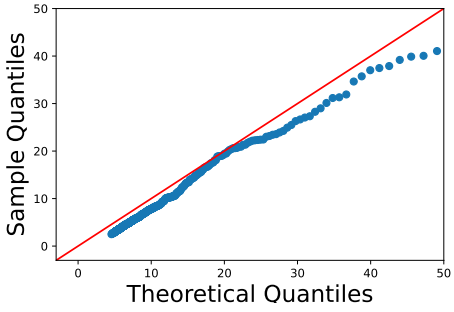
(a) Density $n = 10$



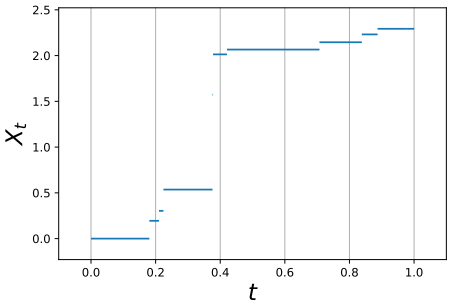
(b) Density $n = 100000$



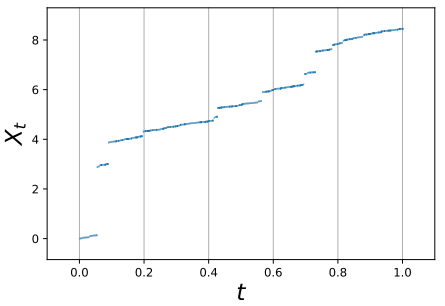
(c) Q-Q plot $n = 10$



(d) Q-Q plot $n = 100000$



(e) Lévy path $n = 10$



(f) Lévy path $n = 100000$

Figure 2.5: Simulations of an α-stable Lévy process X using Inverse Lévy measure series representation truncated at $n = 10$ (left column) and $n = 100000$ (right column). The first row represents the sample density of 1000 samples of X_1, the second row the Q-Q plots and the third an example of the sample paths.

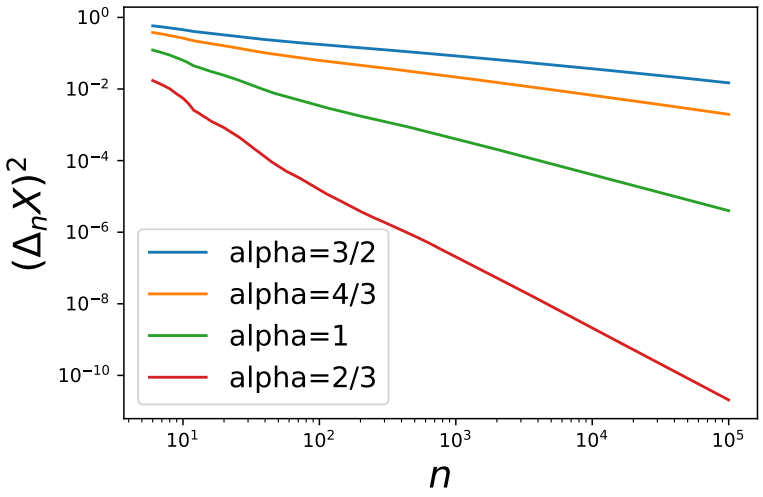
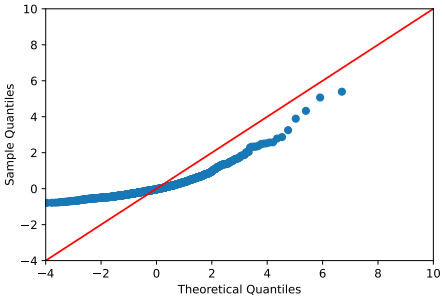
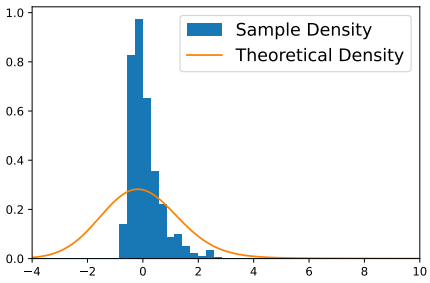


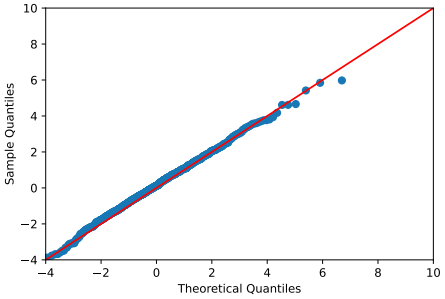
Figure 2.6: Loglog plot of the squared error for α-stable random variables



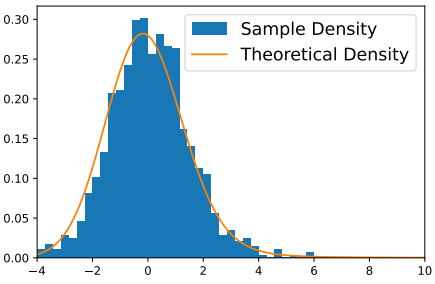
(a) Q-Q plot no correction



(b) Density no correction

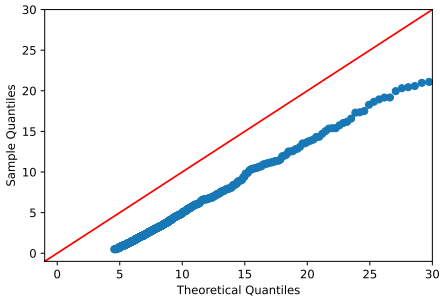


(c) Q-Q plot Brownian correction

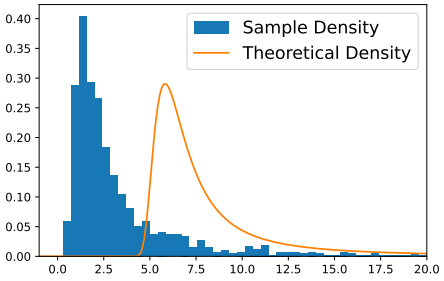


(d) Density Brownian correction

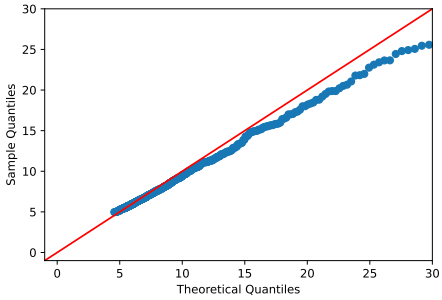
Figure 2.7: Simulations of an α-stable random variables $X \sim S_{1.9}(1, 1, 0)$ without corrections (up) and using the Brownian correction (down).



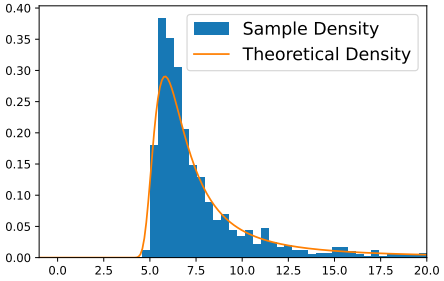
(a) Q-Q plot no correction



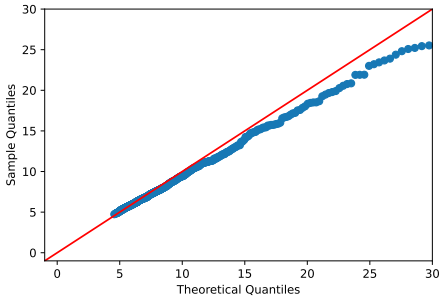
(b) Density no correction



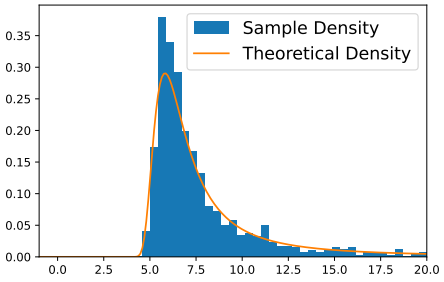
(c) Q-Q plot mean correction



(d) Density mean correction

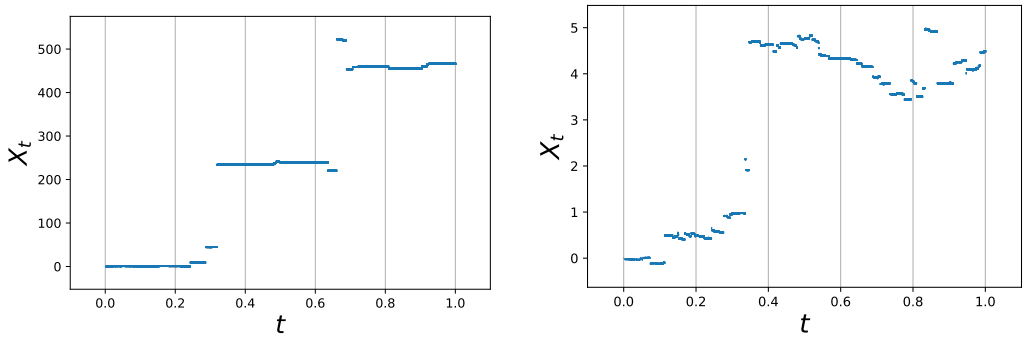


(e) Q-Q plot Brownian correction



(f) Density Brownian correction

Figure 2.8: Simulations of an α-stable random variables $X \sim S_{0.9}(1, 1, 0)$ without corrections (up), correcting with the mean of small jumps (mid) and using both mean and Brownian correction (down).



(a) α-stable Lévy process path (b) CGMY Lévy process path

Figure 2.9: Realization of a sample path of an α-stable Lévy process and the corresponding path of the CGMY Lévy process computed using the rejection method truncated at $n = 10^6$

In figure 2.9 there is a comparison on the path of an α-stable Lévy process and the corresponding path of the CGMY Lévy process computed using the rejection method.

When x is small enough $\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathcal{Q}^0}(x) \sim \mathcal{O}(1)$, hence this method inherits the convergence property of the inverse method for α-stable distributions, as can be seen in figure 2.10.

2.5 Simulation of Lévy-driven stochastic differential equations

In the context of simulation of Lévy processes it is interesting to simulate paths of solutions of stochastic differential equations driven by a Lévy process. In order to do so it is necessary to provide some theoretical background. We need to give meaning to stochastic integrals driven by Lévy processes, give some conditions to the equation in order to guarantee the existence and the uniqueness of a solution and then find some suitable numerical methods and estimate their order of convergence.

The first problem is the following: given a Lévy process X and a stochastic process H, what is the meaning of $\int_0^t H \,\mathrm{d}X$? The approach cannot be naive since Brownian Motion is an example of a Lévy process and this means that path by path Lebesgue-Stieltjes integral might be ill-defined. We know anyway that at least in the case of a compound Poisson process that Lebesgue-Stieltjes integration is possible.

Theory of integration over Lévy process is not in the scope of this thesis, it is

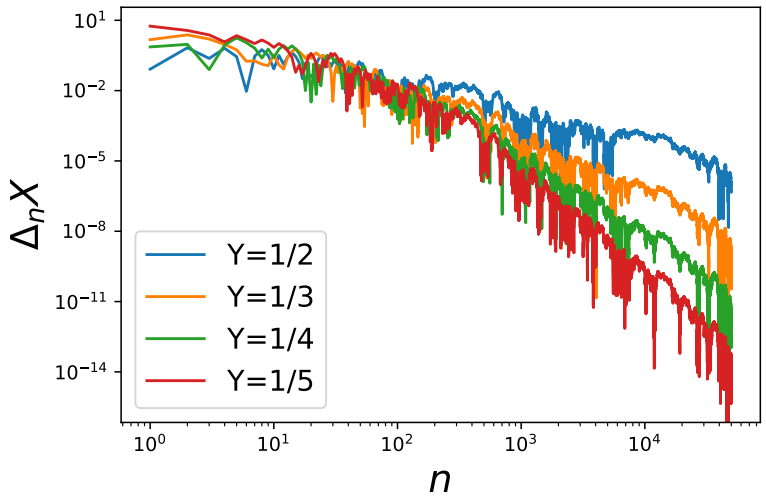


Figure 2.10: Loglog plot of the error for CGMY random variables

anyway important to say that Lévy processes are semimartingales, therefore it is possible to define an integral driven by them that have two fundamental properties: linearity and a form of bounded convergence theorem; see [Pro13].

It is also interesting that there exists some conditions that allows to define the integral in a naive way. The condition required is the boundedness of the p-variation for a value $0 < p < 2$. The p-variation, $0 < p < \infty$, of a real function f on an interval $[a, b]$ is defined as

$$v_p(f) = v_p(f; [a, b]) = \sup_{\kappa} \sum_{i=1}^{n(\kappa)} |f(x_i) - f(x_{i-1})|^p ,$$

where the supremum is taken over all partitions κ of $[a, b]$ and $n(\kappa)$ is the number of elements in the partition. It is possible to prove that symmetric α-stable processes have bounded p-variation for all $p > \alpha$; see [MN00] for a treatment of this topic for stochastic differential equations.

A Lévy-driven stochastic differential equation is usually written in the form

$$\mathrm{d}Y(t) = a(t, Y(t))\,\mathrm{d}t + b(t, Y(t))\,\mathrm{d}X(t) ,$$

which is a notation expressing the following integral equation

$$Y(t) = Y(0) + \int_0^t a(s, Y(s))\,\mathrm{d}s + \int_0^t b(s, Y(s))\,\mathrm{d}X(s) ,$$

where the integrator X is a Lévy Process.

For a general SDE driven by a semimartingale, therefore also for Lévy-driven SDE, we have the following theorem:

Theorem 2.5.0.1. *Let Z be a semimartingale with $Z(0) = 0$, $\mathcal{F}_t$ be the filtration generated by $Z(t)$ and $f : \mathbb{R}_+ \times \Omega \times \mathbb{R} \to \mathbb{R}$ be such that*

1. *for fixed x, $(t, \omega) \mapsto f(t, \omega, x)$ is càglàd[3] and adapted to the filtration generated by Z;*
2. *for each (t, ω), $|(t, \omega, x) - f(t, \omega, y)| \leq K(\omega)\,|x - y|$ for some finite random variable K.*

Let Y_0 be finite and $\mathcal{F}_0$ measurable. Then the equation

$$Y(t) = Y_0 + \int_0^t f(s, \cdot, Y(s-))\,\mathrm{d}Z(s)$$

admits a solution. The solution is unique and it is a semimartingale.

The simplest numerical method for the solution of Lévy-driven SDE is the *Euler Method*: given a discrete (equi-spaced) grid $t_n^h := nh$ the approximation Y_n^h is computed iteratively as

$$Y_n^h = Y_{n-1}^h + a(t_{n-1}^h, Y_{n-1}^h)h + b(t_{n-1}^h, Y_{n-1}^h)[X(t_n^h) - X(t_{n-1}^h)],$$

hence in order to use this method it is necessary to know how to sample the increments of the Lévy process X, which might be possible to do only approximately. Higher order methods are in general not used since they require the computation of the quadratic variation, which is in general not trivial.

We would like to study the error of this method. There exist two types of error criteria:

- *Strong errors.* The following criteria belong to this category: $\mathbb{E}\left[\,\left\|Y_N^h - Y(T)\right\|\,\right]$, $\mathbb{E}\left[\,\left\|Y_N^h - Y(T)\right\|^2\,\right]$ and $\mathbb{E}\left[\,\sup_{0\leq t\leq T}\left\|Y_{[t/h]}^h - Y(t)\right\|\,\right]$;
- *Weak errors.* A weak error criterion have the form $\left|\mathbb{E}\left[\,g(Y_N^h)\,\right] - \mathbb{E}\left[\,g(Y(T))\,\right]\right|$ where g is test function satisfying some smoothness condition.

We concentrate on weak errors of the form $\left|\mathbb{E}\left[\,g(Y_N^h)\,\right] - \mathbb{E}\left[\,g(Y(T))\,\right]\right|$. It has been shown in [PT97] that an Euler scheme for a Lévy-driven equation of the form

$$\mathrm{d}Y(t) = f(Y(t-))\,\mathrm{d}X(t)$$

is of order $\mathcal{O}(n^{-1})$ if

[3] with left-continuous path and well-defined right limits, from the French *"continue à gauche, limit à droite"*

- $f \in C^4$ and all derivatives up to order 4 of f are bounded;
- $g \in C^4$ and moreover $|\partial_I g(x)| = \mathcal{O}(\|x\|^{M'})$ for $|I| = 4$ and some $M' \geq 2$;
- $\int_{\{\|x\|\geq 1\}} \|x\|^\gamma \mathcal{Q}(dx) < \infty$ for $2 \leq \gamma \leq M'^* := \max(2M', 8)$, where $\mathcal{Q}$ is the Lévy measure of X;
- $Y_0 \in L^{M'^*}(\Omega)$.

It is worth noticing that an SDE driven by α-stable processes do not satisfy the third condition.

In [Jac+05] the previous results are extended to the case of Euler schemes where increments are sampled only approximately. This is for example the case when a truncated series representation of a Lévy process is used for sampling from an infinitely divisible distribution. The condition on the sampling algorithm is that, calling ζ_1^n the approximation of the increment for a step wide n^{-1}

$$|\delta_n(h)| = \left| \mathbb{E}\left[h(\zeta_1^n) \right] - \mathbb{E}\left[h\left(Y\left(\frac{1}{n}\right)\right)\right]\right| \leq \frac{K u_n}{n} \|h\|_E$$

for some functions h in a normed space D_E with norm $\|\cdot\|_E$. Then, under some smoothness conditions we get

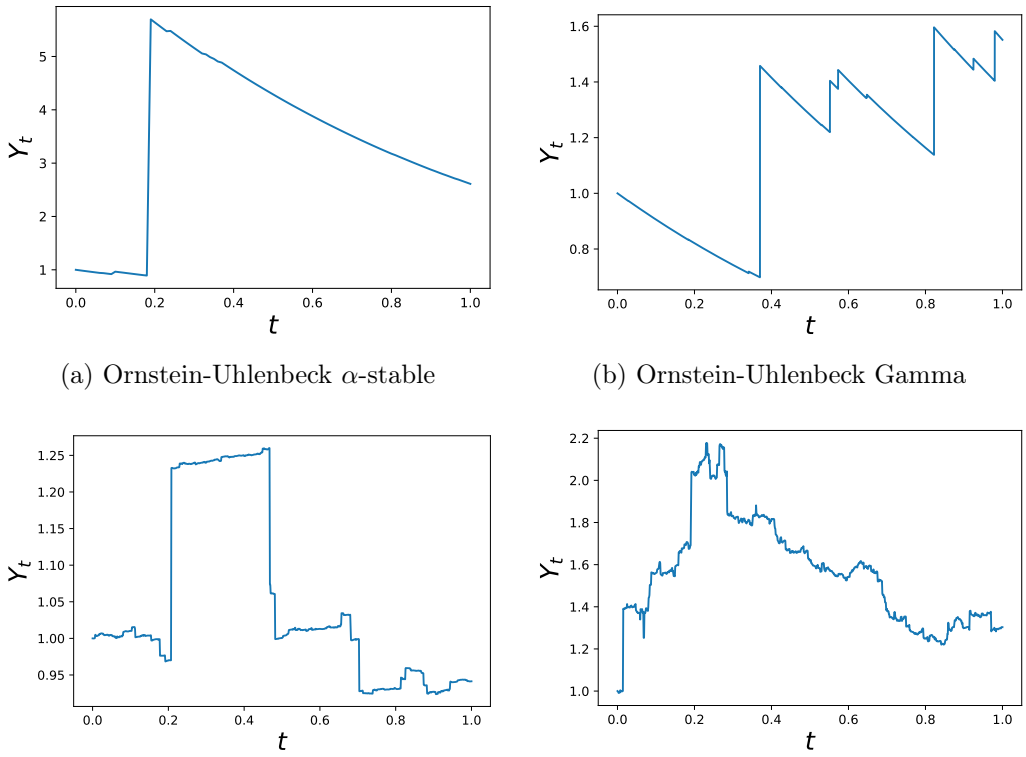
$$\left| \mathbb{E}\left[g(X_N^h) \right] - \mathbb{E}\left[g(X(T)) \right] \right| \sim \mathcal{O}(u_n \wedge n^{-1})$$

Even in this case α-stable processes are excluded.

We now show in figure 2.11 some sample paths of the following two equations using the Euler scheme:

- *Langevin Equation*: $\mathrm{d}Y(t) = -\beta Y(t)\,\mathrm{d}t + \mathrm{d}X(t)$,
- *Geometric Lèvy processs*: $\mathrm{d}Y(t) = cY(t)\,\mathrm{d}t + \gamma Y(t-)\,\mathrm{d}X(t)$.

They both satisfy the conditions in Theorem 2.5.0.1. The solution of the Langevin equation is also called *Ornstein-Uhlenbeck process* which has the notable property of being positive if $Y_0 > 0$ and X is a subordinator.



(a) Ornstein-Uhlenbeck α-stable

(b) Ornstein-Uhlenbeck Gamma

(c) Geometric Lévy α-stable

(d) Geometric Lévy CGMY

Figure 2.11: Simulation of sample paths of the solution of a Lévy-driven SDE. The upper plots are two Ornstein-Uhlenbeck process paths with $\beta = 1$, in the left one the subordinator is an α-stable process $S_{0.9}(1, 0, 0)$, in the right one it is a Gamma Process where $X(1) \sim \text{Gamma}(3, 0.5)$. The lower plots are two Geometric Lévy processes paths with $c = 0.03$ and $\gamma = 0.2$, in the left one Y is a symmetric α-stable process where $S_{0.9}(0, 0, 0)$, in the right Y is a CGMY Process where $C = G = M = 1$ and $Y = 0.9$.

Chapter 3

Simulations of max-stable processes

In this chapter we discuss the theory behind simulations of max-stable processes. Max-stable processes are a class of stochastic processes that can be obtained as the weak limit of the renormalized point-wise maxima of iid realizations of a stochastic process. Similarly to what is done for the series representation of a Lévy process, it is possible to find representations of max-stable processes in such a way that an approximate simulation algorithm can be implemented easily truncating the representation at a finite number n of terms. However, in contrast to the Lévy case where truncating the series always provides an error, we will see that in this case it is possible to obtain an exact simulation of a max-stable process.

3.1 Max-stable distributions

Before introducing the definition of max-stable processes, we need to explain what max-stable distributions are and to explicitly state their connection with Poisson random measures.

Definition 3.1.0.1 (max-stable distribution)**.** Let $(X_n)_{n\geq 1}$ be a sequence of iid random variables with common distribution H and $(M_n)_{n\geq 1}$ be the sequence of partial maxima

$$M_n \coloneqq \max_{i=1,\dots,n} X_i.$$

The non-degenerate distribution H is said to be *max-stable* if there exist two sequences $(b_n)_{n\geq 1}$ and $(c_n)_{n\geq 1}$ with $c_n > 0$, such that for each $n \geq 1$ integer

$$c_n^{-1}(M_n - b_n) \sim H.$$

The following 3 examples provide a complete characterization of all possible

types[1] of max-stable distributions.

Example 3.1.0.2 (Fréchet distribution)**.** Let $\alpha > 0$ and $H = \Phi_\alpha$ where

$$\Phi_\alpha(x) = e^{-x^{-\alpha}}, \quad x > 0.$$

Φ_α is called the *Fréchet distribution* and it is max-stable. In fact, taking $c_n = n^{\frac{1}{\alpha}}$ and $b_n = 0$ we get

$$\begin{aligned}\mathbb{P}(n^{-\frac{1}{\alpha}} M_n \leq x) =& \mathbb{P}(M_n \leq n^{\frac{1}{\alpha}} x) \\ =& (\mathbb{P}(X_1 \leq n^{\frac{1}{\alpha}} x))^n \\ =& (e^{-(n^{\frac{1}{\alpha}} x)^{-\alpha}})^n = e^{-x^{-\alpha}} = \Phi_\alpha(x).\end{aligned}$$

Example 3.1.0.3 (Gumbel distribution)**.** Let $H = \Lambda$ where

$$\Lambda(x) = e^{-e^{-x}}, \quad c \in \mathbb{R}.$$

Λ is called the *Gumbel distribution* and it is max-stable. In fact, taking $c_n = 1$ and $b_n = \log n$ we get

$$\begin{aligned}\mathbb{P}(M_n - \log n \leq x) =& \mathbb{P}(M_n \leq x + \log n) \\ =& (\mathbb{P}(X_1 \leq x + \log n))^n \\ =& (e^{-e^{-(x+\log n)}})^n = e^{-e^{-x} n^{-1} n} = e^{-e^{-x}} = \Lambda(x).\end{aligned}$$

Example 3.1.0.4 (Weibull distribution)**.** Let $\alpha > 0$ and $H = \Psi_\alpha$ where

$$\Phi_\alpha(x) = e^{-|x|^\alpha}, \quad x < 0.$$

Ψ_α is called the *(negative) Weibull distribution* and it is max-stable. In fact, taking $c_n = n^{-\frac{1}{\alpha}}$ and $b_n = 0$ we get

$$\begin{aligned}\mathbb{P}(n^{\frac{1}{\alpha}} M_n \leq x) =& \mathbb{P}(M_n \leq n^{-\frac{1}{\alpha}} x) \\ =& (\mathbb{P}(X_1 \leq n^{-\frac{1}{\alpha}} x))^n \\ =& (e^{-(n^{-\frac{1}{\alpha}} |x|)^\alpha})^n = e^{-|x|^\alpha} = \Psi_\alpha(x).\end{aligned}$$

.

[1] We say that a distribution F belongs to the type of a distribution H if there exist $a \in \mathbb{R}_+$ and $b \in \mathbb{R}$ such that for all $x \in \mathbb{R}$, $F(x) = H(ax + b)$.

Max-stable distributions have a role in the theory of maxima of iid random variables similar to the role of α-stable distributions in the context of summation of iid random variables. In fact, the generalized central limit theorem tells us that a suitable normalization of the sum of iid random variables converges weakly to an α-stable distribution while, in the context of maxima, it is possible to prove that, in many cases, a suitable renormalization of the maximum of n iid random variables converges weakly to a Fréchet, Gumbel or Weibull distribution as n goes to infinity. Moreover, Fréchet, Gumbel or Weibull distributions are the only non-degenerate distributions for which this is possible. This is the informal statement of the *Fisher–Tippett theorem* [FT28]

Definition 3.1.0.5 (Maximum domain of attraction of a max-stable distribution)**.** Let $H \in \{\Phi_\alpha, \Lambda, \Psi_\alpha\}$ and F be a probability distribution. F is in the *maximum domain of attraction* of H (we write $F \in \mathrm{MDA}(H)$) if there exist two sequences $(a_n)_{n\geq 1}$ and $(d_n)_{n\geq 1}$ with $a_n > 0$ such that for each sequence of iid random variables $(X_n)_{n\geq 1}$ with common distribution F and $x \in \mathrm{supp}(H)$ we have

$$\mathbb{P}\left(\frac{M_n - d_n}{a_n} \leq x\right) \to H(x)$$

where $M_n = \max_{i=1,\dots,n} X_i$.

The properties of the domain of attraction are strictly linked to the weak convergence of binomial point processes to a Poisson random measure. In fact, consider a distribution $F \in \mathrm{MDA}(H)$ with normalization and centering sequence $(a_n)_{n\geq 1}$ and $(d_n)_{n\geq 1}$ respectively and a sequence $(X_n)_{n\geq 1}$ of iid random variables with common distribution F. We can define a sequence of binomial point processes $(N_n)_{n\geq 1}$ as

$$N_n = \sum_{i=1}^{n} \varepsilon_{a_n^{-1}(X_i - d_n)}.$$

We have that

$$N_n(A) = \sum_{i=1}^{n} \varepsilon_{a_n^{-1}(X_i - d_n)}(A) \sim \mathrm{Bin}(n, F_n(A))$$

where $F_n(A) \coloneqq F(a_n A + d_n)$. The link between the maximum domain of attraction of a max-stable distribution and Poisson random measures is expressed in the subsequent theorem.

Theorem 3.1.0.6. *Consider a sequence $(X_n)_{n\geq 1}$ of iid random variables with common distribution F and let H be a max-stable distribution. Then the following statements are equivalent:*

1. $F \in \mathrm{MDA}(H)$;

2. *there exist two sequences* $(a_n)_{n\geq 1}$ *and* $(d_n)_{n\geq 1}$ *with* $a_n > 0$ *such that*

$$n\bar{F}(a_n x + d_n) \to -\log H(x)$$

for every $x \in \operatorname{supp} H$*;* [2]

3. *there exist two sequences* $(a_n)_{n\geq 1}$ *and* $(d_n)_{n\geq 1}$ *with* $a_n > 0$ *such that the binomial processes* $(N_n)_{n\geq 1}$ *converge weakly to* N*, a* $\mathrm{PRM}(\mu_H)$ *with state space* $E = \operatorname{supp} H$ *and mean measure* μ_H *given by*

$$\mu_H((a,b]) = \log H(b) - \log H(a)$$

for $a, b \in E$*.*

Proof. We prove the following: $1 \Leftrightarrow 2$, $2 \Rightarrow 3$, $3 \Rightarrow 1$.

$1 \Leftrightarrow 2$**.** The equivalence can be seen with some algebraic manipulations. The first point means that there exist two sequences $(a_n)_{n\geq 1}$ and $(d_n)_{n\geq 1}$ with $a_n > 0$ such that

$$\mathbb{P}(a_n^{-1}(M_n - d_n) \leq 0) = (F(a_n x + d_n))^n = e^{n\log(1-\bar{F}(a_n x + d_n))} \to H(x)$$

This is equivalent to

$$n\log(1 - \bar{F}(a_n x + d_n)) \to \log(H(x)).$$

A Taylor expansion of $\log(1+x)$ leads to

$$-n(\bar{F}(a_n x + d_n)(1 + \mathrm{o}(1))) \to \log(H(x))$$

which is equivalent to point 2.

$2 \Rightarrow 3$**.** From example 1.2.1.5 it is sufficient to prove that nF_n converges vaguely to μ_h. From point 2 we know that

$$nF_n((x,\infty)) \to \mu_H((x,\infty)).$$

The intervals (x,∞) for $x \in \operatorname{supp}(H)$ generate the Borel sets in $\operatorname{supp}(H)$, hence we get

$$nF_n(A) \to \mu_H(A), \quad A \in \mathcal{B}(\operatorname{supp}(H)).$$

[2] We use the notation $\bar{F}(x)$ to mean $1 - F(x) = \mathbb{P}(X > x)$

$3 \Rightarrow 1$. If $N_n \xrightarrow{\mathrm{d}} N$, then $N_n((x,\infty)) \xrightarrow{\mathrm{d}} N((x,\infty))$ for each continuity point of H (which in this case corresponds to all points in $x \in \operatorname{supp} H$). From this we get in particular that

$$\begin{aligned}\mathbb{P}(a_n^{-1}(M_n - d_n) \leq x) &= \mathbb{P}(N_n((x,\infty)) = 0) \\ &\to \mathbb{P}(N((x,\infty)) = 0) \\ &= e^{-\mu_H((x,\infty))} = H(x).\end{aligned}$$

□

We notice the following connection between Poisson random measures and max-stable distributions. Let H be a max stable distribution and N be a PRM(μ_H) defined as in Theorem 3.1.0.6. Then, if we call $(Y_i)_{i\geq 1}$ the points of N, we get that

$$\sup_{i\geq 1} Y_i \sim H.$$

In fact, calling $y_f = \sup(\operatorname{supp}(H))$,

$$\begin{aligned}\mathbb{P}(\sup_{i\geq 1} Y_i \leq y) =& \mathbb{P}(N((y, y_f)) = 0) \\ =& e^{-\mu_H((y,y_f))} \\ =& e^{-(\log(H(y_f)) - \log(H(y)))} \\ =& e^{\log(H(y))} = H(y)\end{aligned}$$

Using Proposition 1.2.2.1 we can find some explicit representation of Poisson random measures with this property.

Example 3.1.0.7 (A PRM representation of a Fréchet random variable)**.** Let us consider a standard Poisson process N on the positive real line, i.e., a PRM(Leb) with state space the interval $[0,\infty)$. We can write N as

$$N = \sum_{i=1}^{\infty} \varepsilon_{\Gamma_i}$$

where the sequence $(\Gamma_i)_{i\geq 1}$ is obtained as the cumulative sum of standard exponential variables. Let us transform the points via the function h defined as $h(\gamma) = \gamma^{-\frac{1}{\alpha}}$. Then we can define a PRM(ν) on $(0,\infty)$, M, where $\nu := \mathrm{Leb} \circ h^{-1}$. We write this explicitly. Given $0 < a < b < \infty$, we have

$$\nu((a,b]) = \mathrm{Leb}(h^{-1}(a,b]) = a^{-\alpha} - b^{-\alpha} = \mu_{\Phi_\alpha}((a,b])$$

Therefore, we get

$$\max_{i\geq 1} \Gamma_i^{-\frac{1}{\alpha}} = \Gamma_1^{-\frac{1}{\alpha}} \sim \Phi_\alpha.$$

This representation can be extended to a more general one.

Example 3.1.0.8 (A more general PRM representation of a Fréchet random variable)**.** Set $\alpha > 0$ and consider M, a PRM(Leb $\times$ F_V) on $[0,\infty) \times [0,\infty)$ with representation

$$M = \sum_{i=1}^{\infty} \varepsilon_{(\Gamma_i, V_i)},$$

where $(\Gamma_i)_{i\geq 1}$ is defined as before and $(V_i)_{i\geq 1} \overset{\text{iid}}{\sim} F_V$ and independent of $(\Gamma_i)_{i\geq 1}$. We also ask $\mathbb{E}\left[V_1^\alpha\right] < \infty$. We transform the points of the PRM via a function h defined as $h(\gamma, v) = \gamma^{-\frac{1}{\alpha}} v$. As before we can define a new Poisson random measure N on $\mathbb{R}$ with mean measure $\nu := (\text{Leb} \times F_V) \circ h^{-1}$. Given $0 < a < b < \infty$ we compute $\nu((a,b])$:

$$\begin{aligned}
\nu((a,b]) =& (\text{Leb} \times F_V)(h^{-1}(a,b]) \\
=& (\text{Leb} \times F_V)(\{(\gamma, v) \mid a < \gamma^{-\frac{1}{\alpha}} v \leq b\}) \\
=& (\text{Leb} \times F_V)\left(\left\{(\gamma, v) \mid \left(\frac{b}{v}\right)^{-\alpha} \leq \gamma < \left(\frac{a}{v}\right)^{-\alpha}\right\}\right) \\
=& \int_0^\infty \left(\int_{\left(\frac{v}{b}\right)^\alpha}^{\left(\frac{v}{a}\right)^\alpha} \mathrm{d}\gamma\right) F_V(\mathrm{d}v) \\
=& (a^{-\alpha} - b^{-\alpha})\, \mathbb{E}\left[V_1^\alpha\right].
\end{aligned}$$

This means that

$$\begin{aligned}
\mathbb{P}(\sup_{i\geq 1} \Gamma_i^{-\frac{1}{\alpha}} V_i \leq y) =& \mathbb{P}(N((y,\infty)) = 0) \\
=& e^{-\nu((y,\infty))} \\
=& e^{-y^{-\alpha}\mathbb{E}\left[V_1^\alpha\right]} \\
=& \exp\left\{-\left(\frac{y}{\mathbb{E}\left[V_1^\alpha\right]^{\frac{1}{\alpha}}}\right)^{-\alpha}\right\} \\
=& \Phi_\alpha\left(\frac{y}{\mathbb{E}\left[V_1^\alpha\right]^{\frac{1}{\alpha}}}\right) = \Phi_\alpha\left(\frac{y}{\|V_1\|_{L^\alpha}}\right).
\end{aligned}$$

In particular, if $\mathbb{E}\left[V_1^\alpha\right] = 1$, then $\sup_{i\geq 1} \Gamma_i^{-\frac{1}{\alpha}} V_i \sim \Phi_\alpha$.

In the following, we always restrict ourselves to $\mathbb{E}\left[V_1^\alpha\right] = 1$ since we can always get to this easier situation simply through rescaling. It is also possible to restrict ourselves to the case of $\alpha = 1$. In fact, one can obtain any max-stable random variable from a random variable with distribution Φ_1 through simple transformation. Explicitly, let $X \sim \Phi_1$, then

- $X^{\frac{1}{\alpha}} \sim \Phi_\alpha$
- $\log(X) \sim \Lambda$
- $-X^{-\frac{1}{\alpha}} \sim \Psi_\alpha$.

Hence, using the same notation as before and with $\mathbb{E}[V_1] = 1$, we can represent a Gumbel random variable Y as

$$Y \overset{\mathrm{d}}{=} \max_{i \geq 1} (\log(V_i) - \log(\Gamma_i))$$

and a Weibull random variable Z as

$$Z \overset{\mathrm{d}}{=} -\min_{i \geq 1} \left(\frac{\Gamma_i}{V_i} \right)^{\frac{1}{\alpha}}.$$

3.2 Introduction to max-stable processes

We are interested in finding an infinite-dimensional object which has similar properties as max-stable random variables. By this we mean that this random object X should have the property that a suitable renormalization of the point-wise maxima of n independent realizations of X is still distributed as X. This leads to the definition of *max-stable processes*.

Definition 3.2.0.1 (Max-stable process)**.** A stochastic process $X = (X(t))_{t \in T}$ with $T \subseteq \mathbb{R}$ is a *max-stable process* if there exist two sequences $(b_n)_{n \geq 1}$ and $(c_n)_{n \geq 1}$, with $c_n > 0$ for each integer $n \geq 1$, such that, taking n iid copies $X^{(1)}, \ldots, X^{(n)}$ of X, we have

$$c_n^{-1} \left(\max_{i=1,\ldots,n} X^{(i)}(t) - b_n \right)_{t \in T} \overset{\mathrm{d}}{=} X,$$

where $\overset{\mathrm{d}}{=}$ is interpreted in the sense of the finite-dimensional distributions.

It is evident from the preceding definition that, for each $t \in T$, $X(t)$ is either degenerate or a max-stable random variable. Hence, it is possible to define max-stable processes such that the marginals are of a given max-stable distribution type. For example, consider the Fréchet case.

Definition 3.2.0.2 (Max-stable process with Fréchet marginals)**.** A max-stable process X is a *max-stable process with Fréchet marginals* if we can take $b_n = 0$ and $c_n = n^{\frac{1}{\alpha}}$ in definition 3.2.0.1. This means that for each $n \geq 1$ integer and n iid copies $X^{(1)}, \ldots, X^{(n)}$ of X we have

$$n^{-\frac{1}{\alpha}} \left(\max_{i=1,\ldots,n} X^{(i)}(t) \right)_{t \in T} \overset{\mathrm{d}}{=} X,$$

in the sense of the finite-dimensional distributions.[3]

Equivalent definitions can be stated for max-stable processes with Gumbel or Weibull marginals. In the following we treat mostly the Fréchet case with $\alpha = 1$. In fact, as observed in the preceding section it is possible to obtain all the other max-stable distributions through simple transformations of Fréchet distributed random variables with $\alpha = 1$. Equivalently, applying the same transformations point-wise we can obtain max-stable processes with Fréchet, Gumbel or Weibull marginals from a max-stable process with Fréchet marginals with $\alpha = 1$.

We consider a first example of a max-stable process

Example 3.2.0.3. Let $(\Gamma_i)_{i\geq 1}$ be the cumulative sum of iid independent standard exponential and $(U_i)_{i\geq 1}$ a sequence of iid Unif$(0,1)$ random variables independent of $(\Gamma_i)_{i\geq 1}$. Then, consider a stochastic process f from $(\Omega, \mathcal{F}) = ([0,1], \mathcal{B}([0,1]))$ to $\mathbb{R}_+^T$ such that for each $t \in T$, $\mathbb{E}\,[\, f(t;U)\,] < \infty$. We define the stochastic process X as

$$X(t) := \sup_{i\geq 1} \Gamma_i^{-1} f(t;U_i)\,.$$

X is a max-stable process with Fréchet marginals.

Let us prove the previous statement. From definition 3.2.0.2 we need to show that for each $m \geq 1$ and collection of times $t_1, \ldots, t_m \in T$ and values $x_1, \ldots, x_m \in \mathbb{R}_+$,

$$\mathbb{P}\big(X(t_1) \leq x_1, \ldots, X(t_m) \leq x_m\big) = \left(\mathbb{P}\left(\frac{X(t_1)}{n} \leq x_1, \ldots, \frac{X(t_m)}{n} \leq x_m\right)\right)^n.$$

To do so, first we write

$$\begin{aligned}\mathbb{P}\big(X(t_1) \leq x_1, \ldots, X(t_m) \leq x_m\big) =& \mathbb{P}\big(\max_{j=1,\ldots,m} \frac{X(t_j)}{x_j} \leq 1\big)\\ =& \mathbb{P}\big(\max_{i\geq 1} \Gamma_i^{-1} \max_{j=1,\ldots,m} \frac{f(t_j;U_i)}{x_j} \leq 1\big).\end{aligned}$$

From example 3.1.0.8 we notice that

$$\sup_{i\geq 1} \Gamma_i^{-1} \max_{j=1,\ldots,m} \frac{f(t_j;U_i)}{x_j}$$

[3]It is important to notice that this does not mean that all marginals are Φ_1 distributed. In fact, it is possible for $X(t)$ to be degenerate at some points t. In order to have all the marginals Fréchet distributed we need stationarity, which, in the following of this thesis, will be in most cases assumed.

can be seen as a Fréchet random variable. Hence,

$$\begin{aligned}\mathbb{P}\big(\sup_{i\geq 1}\Gamma_i^{-1}\max_{j=1,\dots,m}\frac{f(t_j;U_i)}{x_j}\leq 1\big) &= \Phi_1(1)^{\mathbb{E}\left[\max_{j=1,\dots,m}\frac{f\left(t_j;U_i\right)}{x_j}\right]}\\ &= \Phi_1(1)^{n\mathbb{E}\left[\max_{j=1,\dots,m}\frac{f\left(t_j;U_i\right)}{nx_j}\right]}\\ &= \left(\mathbb{P}\big(\max_{i\geq 1}\Gamma_i^{-1}\max_{j=1,\dots,m}\frac{f(t_j;U_i)}{nx_j}\leq 1\big)\right)^n\\ &= \left(\mathbb{P}\left(\frac{X(t_1)}{n}\leq x_1,\dots,\frac{X(t_m)}{n}\leq x_m\right)\right)^n.\end{aligned}$$

The preceding example provides a representation of all max-stable processes with unit Fréchet marginals when T is countable. This result is summarized in this theorem by de Haan in [Haa84]

Theorem 3.2.0.4. *The finite-dimensional distribution of a max-stable sequence $(X(t))_{t\in\mathbb{Z}}$ with unit Fréchet marginals satisfy the relation*

$$\mathbb{P}(X(t_1)\leq x_1,\dots,X(t_m)\leq x_m)=\exp\left(-\int_{\mathbb{R}_+^m}\max_{j=1,\dots,m}\frac{y_j}{x_j}\,G_{t_1,\dots,t_m}(\mathrm{d}\mathbf{y})\right)$$

where $m\geq 1$ and for $j=1,\dots,m$, $x_j>0$ and $G_{t_1,\dots,t_m}$ is the m-dimensional restriction to $\mathbb{R}_+^m$ of a finite measure G on $\mathbb{R}_+^{\mathbb{Z}}$. Moreover, X has a representation

$$X(t)\overset{d}{=}\sup_{i\geq 1}\Gamma_i^{-1}f(t_j;U_i)\,, \tag{3.1}$$

for $t\in\mathbb{Z}$ and a suitable non-negative stochastic process f such that $\mathbb{E}\,[\,f(t;U)\,]=\int_0^1 f(t;u)\,\mathrm{d}u<\infty$.

In the same paper, de Haan proves that representation (3.1) is valid for a general $T\subseteq\mathbb{R}$ under the condition of the process X being stochastically continuous. This representation is not unique. We will use this property to find the most useful representation of a max-stable process for simulation purposes.

Example 3.2.0.5 (Brown-Resnick process)**.** Consider the process $X=(X(t))_{t\in T}$ with representation

$$X(t)=\sup_{i\geq 1}\Gamma_i^{-1}e^{W_i(t)-\frac{1}{2}\sigma^2(t)}, \tag{3.2}$$

where $(\Gamma_i)_{i\geq 1}$ is the cumulative sum of iid standard exponential random variables, $(W_i)_{i\geq 1}$ are iid realization of a Gaussian process on $\mathbb{R}$ with stationary increments,

$\mathbb{E}[W_1(t)]=0$ and $\sigma^2(t) = \mathrm{Var}(W_1(t))$. The process X is called the *Brown-Resnick process*.

It is possible to write X in the form of (3.1) and to show that it is a max-stable process. Moreover, it is also possible to prove that the process is strictly stationary. An intuition of the stationarity can be seen from the fact that for $t \in T$, $\mathbb{E}\left[e^{W_i(t)-\frac{1}{2}\sigma^2(t)}\right] = 1$, hence $X(t) \sim \Phi_1$.

Kabluchko et al. in [KSH09] proved that the distribution of X only depends on the variance function $\sigma^2(t) = \mathrm{Var}(W_1(t))$ and the *variogram* which is the function $\gamma : T \to \mathbb{R}$ defined as

$$\gamma(h) \coloneqq \mathbb{E}\left[(W_1(h+t_0) - W_1(t_0))^2\right].$$

Possible choices of the Gaussian processes with stationary increments are

Brownian Motion. The Brown-Resnick max-stable process was originally considered for W Brownian motion in [BR77]. In this case $\sigma^2(t) = t$ and $\gamma(t) = t$ and $\mathrm{Cov}(W(s), W(t)) = s \wedge t$.

Fractional Brownian motion. Let $H \in (0,1)$. The fractional Brownian motion $W^{(H)}$ is a 0-mean Gaussian process with stationary increments. It has variance function $\sigma^2(t) = \mathrm{Var}(W^{(H)}(t)) = t^{2H}$,variogram $\gamma(t) = |t|^{2H}$ and covariance $\mathrm{Cov}(W(s), W(t)) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H})$.

3.2.1 Crash course in regular variation

In the next section, we address the problem of simulations of max-stable processes. More precisely, we will pay particular attention to strictly stationary max-stable processes on a discrete set T. In order to have the theoretical tools to approach simulations, we introduce here the concept of regular variation of a strictly stationary random sequence. To do so, first we define regular variation in the univariate and multivariate settings. Then, we can introduce the multivariate Fréchet distribution. After that we extend the concept of regular variation to strictly stationary sequences and we see that max-stable processes with unit Fréchet marginals are regularly varying and all the finite-dimenional distributions are multivariate Fréchet. Finally, we state a condition on the distributions of the spectral process that allows us to change the representations of a max-stable process. The concept of regular variation might be extended also to a more general domain; see [DR15].

Definition 3.2.1.1 (Regularly varying function)**.** A measurable function $U : \mathbb{R}_+ \to \mathbb{R}_+$ is *regularly varying* at ∞ with index $\alpha \in \mathbb{R}$ if

$$\lim_{t\to\infty} \frac{U(tx)}{U(x)} = x^\alpha, \quad x > 0.$$

We call α the exponent of variation. If $\alpha = 0$ the function U is said to be *slowly varying*.

Definition 3.2.1.2 (Regularly varying random variable)**.** A positive random variable X with distribution F is said to be *regularly varying* with index $\alpha > 0$ if the survival function $\bar{F}(x) := 1 - F(x) = \mathbb{P}(X > x)$ is regularly varying with exponent of variation $-\alpha$.

Observation 3.2.1.3. Let $X \sim F$ be a regularly varying random variable with exponent α. Suppose we can define a sequence $(a_n)_{n\geq 1}$ such that $\bar{F}(a_n) = \frac{1}{n}$.[4] From the regular variation of $\bar{F}$ we have

$$\lim_{n\to\infty} n\bar{F}(a_n x) = \lim_{n\to\infty} \frac{\bar{F}(a_n x)}{\bar{F}(a_n)} = x^{-\alpha},$$

then, from Theorem 3.1.0.6 we have $X \in \mathrm{MDA}(\Phi_\alpha)$. An equivalent result would have been proved with the less restrictive condition on $(a_n)_{n\geq 1}$

$$\lim_{n\to\infty} n\bar{F}(a_n) \to 1.$$

We extend the concept of regular variation to the multivariate case.

Definition 3.2.1.4 (Regularly varying random vector with positive and identically distributed components)**.** For an integer $d \geq 2$, let us consider a random vector $\mathbf{X} = (X^{(1)}, \ldots, X^{(d)})^T$ with all components positive and identically distributed following a regularly varying distribution F with exponent α.[5] Let us consider a sequence such that $n\bar{F}(a_n x) \to 1$ as $n \to \infty$. The vector $\mathbf{X}$ is *regularly varying*, we write $\mathbf{X} \in \mathrm{RV}(\alpha, \mu)$, if there exists a non-null Radon measure μ on $\mathcal{B}(\mathbb{R}^d_{+,\mathbf{0}})$ such that for every μ-continuity set A we have

$$\mu_n := n\mathbb{P}(a_n^{-1}\mathbf{X} \in A) \to \mu(A), \quad n \to \infty, \tag{3.3}$$

i.e., the sequence of measures $(\mu_n)_{n\geq 1}$ converges vaguely to μ

Observation 3.2.1.5. From (3.3), with an argument similar to the one made in Example 1.2.1.5, we can find a sequence of binomial processes $(N_n)_{n\geq 1}$ defined by

$$N_n(A) = \sum_{i=1}^{n} \varepsilon_{a_n^{-1}\mathbf{X}_i}(A), \quad A \in \mathcal{B}(\mathbb{R}^d_{+,\mathbf{0}})$$

that converges weakly to a $\mathrm{PRM}(\mu)$.

[4] In order to be sure we can define such a sequence we need F continuous

[5] We restrict ourselves to the case of identically distributed components, the reason is that we will deal only with strictly stationary max-stable processes and this makes the theory easier. However, it is possible to relax this hypothesis.

Observation 3.2.1.6. Consider a sequence $(\mathbf{X}_n)_{n\geq 1}$ of iid regularly varying random vectors with positive and identically distributed components and let the sequence $(\mathbf{M}_n)_{n\geq 1}$ be the sequence of component-wise maxima

$$\mathbf{M}_n = (M_n^{(1)}, \dots, M_n^{(d)})^T = (\max_{i=1,\dots,n} X_i^{(1)}, \dots, \max_{i=1,\dots,n} X_i^{(d)})^T.$$

Then, taking a vector $\mathbf{x} \in \mathbb{R}^d_{+,\mathbf{0}}$ we have

$$\mathbb{P}(a_n^{-1}\mathbf{M}_n \in [\mathbf{0}, \mathbf{x}]^c) = \mathbb{P}(N_n([\mathbf{0}, \mathbf{x}]^c) = 0) \to \mathbb{P}(N([\mathbf{0}, \mathbf{x}]^c) = 0) = e^{-\mu([\mathbf{0},\mathbf{x}]^c)}$$

as $n \to \infty$. We call $H \coloneqq e^{-\mu}$ a *multivariate Fréchet distribution* and μ its *exponent measure.*

We show a different characterization of the regular variation property which will be useful for the extension of regular variation to the countable infinite case.

Proposition 3.2.1.7. *For an integer $d \geq 2$, let us consider a random vector $\mathbf{X} = (X^{(1)}, \dots, X^{(d)})^T$ with all components positive and identically distributed with distribution F, a value $\alpha > 0$ and a Radon measure μ on $\mathbb{R}^d_{+,\mathbf{0}}$. Then, the following statements are equivalent:*

1. $\mathbf{X} \in \mathrm{RV}(\alpha, \mu)$
2. $X^{(1)} \in \mathrm{RV}(\alpha)$ *and*
$$\lim_{x\to\infty} \frac{\mathbb{P}(x^{-1}\mathbf{X} \in A)}{\mathbb{P}(X^{(1)} > x)} = \mu(A)$$
for all μ-continuity sets in $\mathcal{B}(\mathbb{R}^d_{+,\mathbf{0}})$

Proof. $2 \Rightarrow 1$**.** Consider a sequence $(a_n)_{n\geq 1}$ such that $\lim_{n\to\infty} n\bar{F}(a_n) \to 1$, then $n\mathbb{P}(a_n^{-1}\mathbf{X} \in A) \to \mu(A)$ for all μ-continuity sets in $\mathcal{B}(\mathbb{R}^d_{+,\mathbf{0}})$. Which means $\mathbf{X} \in \mathrm{RV}(\alpha, \mu)$.

$1 \Rightarrow 2$**.** We only need to show that for all A μ-continuity sets

$$\lim_{x\to\infty} \frac{\mathbb{P}(x^{-1}\mathbf{X} \in A)}{\mathbb{P}(X^{(1)} > x)} = \mu(A).$$

Consider a sequence $(a_n)_{n\geq 1}$ such that $a_n\bar{F}(n) \to 1$. The sequence $(a_n)_{n\geq 1}$ is regularly varying with index $\frac{1}{\alpha}$, hence we can define a function as

$$a(x) \coloneqq x^{\frac{1}{\alpha}} L(x)$$

with L slowly varying and such that $a(x) = a_{[x]}$, where $[x]$ denotes the integer part of x. For every regularly varying function a with positive index there

exist a continuous function $\tilde{a}$ such that $a \sim \tilde{a}$ for $x \to \infty$. Then for every μ-continuity set A we can write

$$\begin{aligned}
\mu(A) &= \lim_{x\to\infty} \frac{\mathbb{P}\left(a(x)^{-1}\mathbf{X} \in A\right)}{\mathbb{P}\left(X^{(1)} > a(x)\right)} \\
&= \lim_{x\to\infty} \frac{\mathbb{P}\left(\frac{\tilde{a}(x)}{a(x)}\frac{\mathbf{X}}{\tilde{a}(x)} \in A\right)}{\mathbb{P}\left(\frac{X^{(1)}}{\tilde{a}(x)} > \frac{a(x)}{\tilde{a}(x)}\right)} \\
&= \lim_{x\to\infty} \frac{\mathbb{P}\left(\tilde{a}(x)^{-1}\mathbf{X} \in A\right)}{\mathbb{P}\left(X^{(1)} > \tilde{a}(x)\right)} \\
&= \lim_{x\to\infty} \frac{\mathbb{P}(x^{-1}\mathbf{X} \in A)}{\mathbb{P}(X^{(1)} > x)}.
\end{aligned}$$

□

We extend this concept to stationary random sequences.

Definition 3.2.1.8 (Regularly varying stationary sequence)**.** A stationary process $(X(t))_{t\in\mathbb{Z}}$ is regularly varying with index $\alpha > 0$ if for every integer $h \geq 0$ there exists a non-null Radon measure μ_h on $\mathbb{R}^h_{+,\mathbf{0}}$ such that

$$\lim_{x\to\infty} \frac{\mathbb{P}(x^{-1}(X(0), \ldots, X(h)) \in A)}{\mathbb{P}(X(0) > x)} = \mu_h(A)$$

for all μ_h-continuity sets in $\mathcal{B}(\mathbb{R}^h_{+,\mathbf{0}})$.

Now, consider a max-stable process X with representation

$$X = \sup_{i\geq 1} \Gamma_i^{-1} V_i(t) \tag{3.4}$$

where, using the usual notation, $(\Gamma_i)_{i\geq 1}$ is the cumulative sum of standard exponentials and V_i are iid realizations of a generic stochastic process V, independent of the sequence $(\Gamma_i)_{i\geq 1}$, that we call *spectral process*.

Proposition 3.2.1.9. *Consider a max-stable process $X = (X(t))_{t\in T}$ with $T \subseteq \mathbb{Z}$ and representation (3.4). Then, for each integer $m > 1$, $t_1, \ldots, t_m \in T$ and $x_1, \ldots, x_m \in \mathbb{R}_+$ we have*

$$\mathbb{P}(X(t_1) < x_1, \ldots, X(t_m) < x_m) = \exp\left\{-\mathbb{E}\left[\max_{i=1,\ldots,m} \frac{V(t_i)}{x_i}\right]\right\}, \tag{3.5}$$

which is a multivariate Fréchet distribution with exponent measure

$$\mu_{t_1,\ldots,t_m}([\mathbf{0}, \mathbf{x}]^c) = \mathbb{E}\left[\max_{i=1,\ldots,m} \frac{V(t_i)}{x_i}\right].$$

Moreover, X is regularly varying with $\alpha = 1$.

Proof. We start by proving (3.5). Fix $m > 1$ integer, $t_1, \dots, t_m \in T$ and $x_1, \dots, x_m \in \mathbb{R}_+$. Then,

$$\begin{aligned}
\mathbb{P}(X(t_1) < x_1, \dots, X(t_m) < t_m) &= \mathbb{P}\left(\max_{j=1,\dots,m} \frac{X(t_j)}{x_j} < 1\right) \\
&= \mathbb{P}\left(\max_{j=1,\dots,m} \sup_{i\geq 1} \frac{\Gamma_i^{-1} V_i(t_j)}{x_j} < 1\right) \\
&= \mathbb{P}\left(\sup_{i\geq 1} \Gamma_i^{-1} \max_{j=1,\dots,m} \frac{V_i(t_j)}{x_j} < 1\right) \\
&= \Phi_1\left(\mathbb{E}\left[\max_{j=1,\dots,m} \frac{V_i(t_j)}{x_j}\right]^{-1}\right) \\
&= \exp\left\{-\mathbb{E}\left[\max_{j=1,\dots,m} \frac{V(t_j)}{x_j}\right]\right\}.
\end{aligned}$$

Now, we prove that X is regularly varying. We show that $X(0) \in \mathrm{RV}(1)$ and for each integer $h \geq 1$ there exists a Radon measure μ_h such that

$$\lim_{y\to\infty} \frac{\mathbb{P}(y^{-1}(X(0), \dots, X(h)) \in [\mathbf{0}, \mathbf{x}]^c)}{\mathbb{P}(X(0) > y)} = \mu_h([\mathbf{0}, \mathbf{x}]^c).$$

We have that

$$\mathbb{P}(X(0) > y) = 1 - \Phi_1(y) \sim y^{-1},$$

then $X(0) \in \mathrm{RV}(1)$. Moreover, by (3.5),

$$\begin{aligned}
\mathbb{P}(y^{-1}(X(0), \dots, X(h)) \in [\mathbf{0}, \mathbf{x}]^c) &= \mathbb{P}((X(0), \dots, X(h)) \in [\mathbf{0}, y\mathbf{x}]^c) \\
&= 1 - \mathbb{P}((X(0), \dots, X(h)) \in [\mathbf{0}, y\mathbf{x}]) \\
&= 1 - \exp\left\{-\mathbb{E}\left[\max_{j=0,\dots,h} \frac{V(j)}{y x_j}\right]\right\} \\
&= 1 - \exp\left\{-\frac{1}{y}\mathbb{E}\left[\max_{j=0,\dots,h} \frac{V(j)}{x_j}\right]\right\} \\
&\sim \frac{\mathbb{E}\left[\max_{j=0,\dots,h} \frac{V(j)}{x_j}\right]}{y}.
\end{aligned}$$

Then,

$$\lim_{y\to\infty} \frac{\mathbb{P}(y^{-1}(X(0), \dots, X(h)) \in [\mathbf{0}, \mathbf{x}]^c)}{\mathbb{P}(X(0) > y)} = \mathbb{E}\left[\max_{j=0,\dots,h} \frac{V(j)}{x_j}\right] = \mu_h([\mathbf{0}, \mathbf{x}]^c).$$

Sets of the form $[\mathbf{0}, \mathbf{x}]^c$ define a π-system, therefore we can extend the measure to a general $A \in \mathcal{B}(\mathbb{R}^{h+1}_{+,\mathbf{0}})$. □

Now, consider two different representations of a max-stable process,

$$X \stackrel{\mathrm{d}}{=} \sup_{i\geq 1} \Gamma_i^{-1} V_i \stackrel{\mathrm{d}}{=} \sup_{i\geq 1} \Gamma_i^{-1} W_i \tag{3.6}$$

where V_i and W_i are iid realizations of two different stochastic processes. In order to have the equality in distribution in the sense of the finite-dimensional distributions we need for each $t_1, \ldots, t_m \in T$ and $x_1, \ldots, x_m$

$$\mu^{(V)}_{t_1,\ldots,t_m}([\mathbf{0}, \mathbf{x}]^c) = \mu^{(W)}_{t_1,\ldots,t_m}([\mathbf{0}, \mathbf{x}]^c),$$

which more explicitly means

$$\mathbb{E}\left[\max_{i=1,\ldots,m} \frac{V(t_i)}{x_i}\right] = \mathbb{E}\left[\max_{i=1,\ldots,m} \frac{W(t_i)}{x_i}\right]. \tag{3.7}$$

.

3.3 Simulation of max-stable processes

In this section we approach simulations of max-stable processes. We will describe three general-purpose methods for max-stable processes: the threshold method, the extremal functions method and the probability threshold method. The threshold method and probability threshold method provide two different ways to truncate the so-called *spectral representation* of a max-stable process. On the contrary, the extremal functions method aims at directly sampling the elements of the spectral representation that contribute to the maximum, which are called *extremal functions*. In this section we do not provide numerical results, they are postponed to the last section of this chapter.

3.3.1 Threshold method

Let us consider a max-stable process X with representation (3.4). The idea of the threshold method is to iteratively simulate the truncated process

$$X^{(j)}(t) = \sup_{1\leq i\leq j} \Gamma_i^{-1} V_i(t)\,. \tag{3.8}$$

Iterations must stop at a value N, which is chosen to be an a.s. finite random variable depending on a threshold τ:

$$N = N_\tau = \min\{j \in \mathbb{N} \mid \Gamma_{j+1}^{-1}\tau < \inf_{t\in T} X^{(j)}(t)\}.$$

This choice means that we stop at the first iteration at which, after sampling Γ_{j+1}, we have $X^{(j+1)} \neq X^{(j)}$ only if $\sup_{t\in T} V_{j+1}(t) > \tau$. Since we asked $V(t)$ to have

finite first moment, the method provides an exact sample as $\tau \to \infty$. This cannot be done in practice, therefore, if the spectral process V is not bounded, this method does not provide an exact simulation. On the contrary, if we can find an upper bound such that $\sup_{t\in T} V(t) < \tau$ a.s. then choosing τ as threshold we know for sure that $X^{(N_\tau)} = X$ on T, hence the method provides an exact simulation of X. The pseudocode of the algorithm can be seen in Algorithm 1.

Algorithm 1 Threshold method

procedure ThresholdMethod(Domain K, threshold τ)
  $X(t) \leftarrow 0$ for all $t \in K$
  Simulate $\Gamma \sim \text{Exp}(1)$
  while $\tau/\Gamma \geq \inf_{t\in K} X(t)$ **do**
    Simulate $V \sim F_V$
    $X(t) \leftarrow \max\{\Gamma^{-1}V(t), X(t)\}$ for all $t \in K$
    Simulate $E \sim \text{Exp}(1)$
    $\Gamma \leftarrow \Gamma + E$
  end while
  return X
end procedure

The spectral process V in representation (3.4) is not unique. Hence, we are free to choose it in such a way that it is more suitable for simulations. Then, it might be useful to choose V bounded and easy to simulate.

Whether a spectral process V is easy to simulate depends on each case and therefore cannot be studied on an abstract setting. However, it is possible to provide some ideas as to find some spectral representations with a bounded spectral process. The main idea is to find a spectral process which is normalized with respect to a norm.

Consider a process X with representation (3.4) and a norm $\|\cdot\|$ on $\mathbb{R}^T_+$. We want to find a way to change the representation to

$$X \overset{\text{d}}{=} \sup_{i\geq 1} \Gamma_i^{-1} V_i^{\|\cdot\|}(t)\,,$$

where $V^{\|\cdot\|}$ is a spectral process such that $\left\|V^{\|\cdot\|}\right\| = \theta_{\|\cdot\|}$ constant a.s. First, we show that the value $\theta_{\|\cdot\|}$ does not depend on the representation. Let T be a compact set, then.

$$\begin{aligned}\lim_{n\to\infty} n\mathbb{P}(\|X\| \geq n) &= \lim_{n\to\infty} \frac{\mathbb{P}(\|X\| \geq n)}{\mathbb{P}(\Gamma_1^{-1} \geq n)} \\ &= \lim_{n\to\infty} \frac{\mathbb{P}(\Gamma_1^{-1}\|V_1\| \geq n)}{\mathbb{P}(\Gamma_1^{-1} \geq n)} = \mathbb{E}\left[\,\|V\|\,\right]\end{aligned} \tag{3.9}$$

where in the last equality we used Breiman's lemma.[6]

This proves that $\mathbb{E}[\,\|V\|\,]$ does not depend on the representation. Moreover, since $\left\|V^{\|\cdot\|}\right\| = \theta_{\|\cdot\|}$ a.s., we have also that $\mathbb{E}[\,\|V\|\,] = \mathbb{E}\left[\,\left\|V^{\|\cdot\|}\right\|\,\right] = \theta_{\|\cdot\|}$. Then, $\theta_{\|\cdot\|}$ does not depend on $V^{\|\cdot\|}$.

More generally, we can extend this proof for a 1-homogeneous functional l on the space of the spectral process. Then, we can extend our reasoning to, for instance, functionals of the form $l(v) = v(t)$, which appear naturally in the context of regular variation. In the following we see how we can find an explicit form for the distribution of an l-normalized spectral process.

Proposition 3.3.1.1. *Let $X = \sup_{i\geq 1} \Gamma_i^{-1} V_i$ where $V_i \overset{iid}{\sim} F_V$ and consider a stochastic process Y with distribution $F_Y \ll F_V$[7] and an F_V-a.s. positive 1-homogeneous functional $l : C_0 \to \mathbb{R}_+$. Then,*

$$X \overset{d}{=} \sup_{i\geq 1} \Gamma_i^{-1} \mathbb{E}[\,l(Y)\,] \frac{Y}{l(Y)}$$

if and only if $F_Y(\mathrm{d}v) = \frac{l(v)}{\mathbb{E}[l(v)]} F_V(\mathrm{d}v)$.

Proof. From Proposition 3.2.1.9, we only need to show that for each integer $m \geq 1$ and $t_1, \ldots, t_m \in T$ and $x_1, \ldots, x_m$ we have

$$\mathbb{E}_{F_V}\left[\max_{i=1,\ldots,m} \frac{V(t_i)}{x_i}\right] = \mathbb{E}_{F_Y}\left[\max_{i=1,\ldots,m} \frac{\mathbb{E}[\,l(Y)\,]}{l(Y)} \frac{Y(t_i)}{x_i}\right].$$

We wrote $\mathbb{E}[\,l(Y)\,]$ not specifying the measure since, as we showed in (3.9), it does not depend on the representation, but only on the process X. We have

$$\begin{aligned}
\mathbb{E}_{F_Y}\left[\max_{i=1,\ldots,m} \frac{\mathbb{E}[\,l(Y)\,]}{l(Y)} \frac{Y(t_i)}{x_i}\right] &= \mathbb{E}[\,l(V)\,] \int_{v\in C_0} \max_{i=1,\ldots,m} \frac{1}{l(v)} \frac{v(t_j)}{x_j} F_Y(\mathrm{d}v) \\
&= \mathbb{E}[\,l(V)\,] \int_{v\in C_0} \max_{i=1,\ldots,m} \frac{1}{l(v)} \frac{v(t_j)}{x_j} \frac{\mathrm{d}F_Y}{\mathrm{d}F_V}(v) F_V(\mathrm{d}v) \\
&= \int_{v\in C_0} \max_{i=1,\ldots,m} \frac{v(t_j)}{x_j} F_V(\mathrm{d}v),
\end{aligned}$$

where the last equality is true for all integers $m \geq 1$, $t_1, \ldots, t_m \in T$ and $x_1, \ldots, x_m$ if and only if $\frac{\mathrm{d}F_Y}{\mathrm{d}F_V}(v) = \frac{l(v)}{\mathbb{E}[l(V)]}$. □

[6] Breiman's Lemma informally says that for a regularly varying random variable $X > 0$ with index $\alpha > 0$ and a random variable $Y > 0$ with finite α-th moment we have $\mathbb{P}(XY > x) \sim \mathbb{P}(X > x)\mathbb{E}[\,Y^\alpha\,]$ as $x \to \infty$.

[7] By $\ll$ we mean absolute continuity of all the finite dimensional distributions.

Observation 3.3.1.2. From Proposition 3.3.1.1 we can directly find the distribution of $V^l = \mathbb{E}\left[\, l(Y)\,\right]\frac{Y}{l(Y)}$ as

$$\mathbb{P}(V^l \in A) = \mathbb{E}_{F_V}\left[\frac{l(V)}{\mathbb{E}\left[\, l(V)\,\right]}\mathbf{1}_{\left\{\mathbb{E}[\,l(V)\,]\frac{V}{l(V)}\in A\right\}}\right].$$

Now we consider two examples of spectral normalization: sup-normalization and sum-normalization.

Sup-normalized threshold method

This kind of normalization is the more natural one. In fact, given the fact that we are looking for a bounded spectral process, this normalization let us know exactly the sup and therefore an optimal threshold for the threshold method. More precisely: given a max-stable process X on a compact set K with representation (3.4), we want to find the distribution of the sup-normalized spectral process $V^{\|\cdot\|_\infty}$ defined as

$$V^{\|\cdot\|_\infty} = \frac{\theta_{\|\cdot\|_\infty}}{\|Y\|_\infty}Y$$

where $\|Y\|_\infty \coloneqq \sup_{t\in K} Y(t)$ and Y is a stochastic process with the distribution given by Proposition 3.3.1.1 with $l(\cdot) = \|\cdot\|_\infty$. Then, one can use the threshold method using $\tau = \theta_{\|\cdot\|_\infty}$.

This method works well in theory, but it has several problems in practice. Two are the main problem that are difficult to overcome in order to have exact simulations:

- In general, there is no analytic expression for the coefficient $\theta_{\|\cdot\|_\infty}$ (which is known in the literature as *extremal coefficient*). There exist some estimation procedures for $\theta_{\|\cdot\|_\infty}$, but there is no chance to find the exact value.
- There exists no generic simple algorithm for exact sampling from Y.

Sum-normalized threshold method

Consider a finite number of points $t_1, \ldots, t_m \in T$ and the l^1-norm in $\mathbb{R}^m_+$

$$\|\mathbf{x}\|_1 \coloneqq \sum_{i=1}^{m} x_i.$$

Given a max-stable process X on T with representation (3.4) we want to simulate its restriction to $t_1, \ldots, t_m$. The sum-normalized spectral process on $t_1, \ldots, t_m$ is

$$V^{\|\cdot\|_1} = \frac{\theta_{\|\cdot\|_1}}{\|Y\|_1}Y$$

where the distribution of Y is given by the relation in Proposition 3.3.1.1 with $l(\cdot) = \|\cdot\|_1$. Since in $\mathbb{R}_+^m$ we have $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1$, if we put $\tau = \theta_{\|\cdot\|_1}$ we are sure that

$$\max_{i=1,\dots,m} V^{\|\cdot\|_1}(t_i) \leq \tau.$$

Hence, we can use this threshold for exact simulations. This normalization is found to be the most useful for simulations. The first reason is that it is usually easy to compute $\theta_{\|\cdot\|_1}$. In fact

$$\theta_{\|\cdot\|_1} = \mathbb{E}\left[V(t_1) + \cdots + V(t_m)\right] = \mathbb{E}\left[V(t_1)\right] + \cdots + \mathbb{E}\left[V(t_m)\right]$$

If we assume X strictly stationary, then $\mathbb{E}\left[V(t_1)\right] = \cdots = \mathbb{E}\left[V(t_m)\right]$. Moreover, assuming that the marginals are unit Fréchet we need to have $\mathbb{E}\left[V(t_1)\right] = 1$, then $\theta_{\|\cdot\|_1} = m$.

The second reason is that it is easy to simulate the process Y. In fact, let us call Y^{l_t} the process such that $\mathbb{E}\left[Y^{l_t}(t)\right] \frac{Y^{l_t}}{Y^{l_t}(t)}$ is the spectral process for X normalized under the functional l_t defined by $l_t(v) = v(t)$. We want to prove that $Y \overset{\mathrm{d}}{=} Y^{l_T}$ where T is a random variable sampled uniformly from $\{t_1, \dots, t_m\}$. To show this, from Proposition 3.3.1.1 we have

$$\begin{aligned}
F_Y(\mathrm{d}v) =& \frac{1}{m}\left(\sum_{i=1}^m v(t_i)\right) F_V(\mathrm{d}v) \\
=& \frac{1}{m}\left(\sum_{i=1}^m v(t_i) F_V(\mathrm{d}v)\right) \\
=& \frac{1}{m}\left(\sum_{i=1}^m F_{Y^{l_{t_i}}}(\mathrm{d}v)\right) \\
=& \sum_{i=1}^m \mathbb{P}\left(Y^{l_T} \in \mathrm{d}v \mid T = t_i\right) \mathbb{P}(T = t_i) = \mathbb{P}\left(Y^{l_T} \in \mathrm{d}v\right).
\end{aligned}$$

Then, we only need to be able to simulate from $Y^{l_{t_i}}$ for simulating $V^{\|\cdot\|_1}$. If we are directly interested in simulating $V^{\|\cdot\|_1}$, instead of sampling Y^{l_T} we can equivalently sample directly from

$$V^{l_T} := \frac{Y^{l_T}}{Y^{l_T}(T)}.$$

In fact,

$$V^{\|\cdot\|_1} = \frac{m}{\|Y\|_1} Y = \frac{m}{\|Y^{l_T}\|_1} Y^{l_T} = \frac{m Y^{l_T}(T)}{\|Y^{l_T}\|_1} \frac{Y^{l_T}}{Y^{l_T}(T)} = \frac{m}{\|V^{l_T}\|_1} V^{l_T}.$$

We will use this method in our simulations.

3.3.2 Extremal functions method

This method is completely different in its concept from the threshold method. In order to explain it, we need to define what are called *extremal functions* in this context.

Definition 3.3.2.1 (Extremal function)**.** Consider a max-stable process X with representation (3.4) on a compact set K. We denote $\phi_i = \Gamma_i^{-1} V_i$ for $i = 1, 2, \dots$ Then ϕ_i is called K-*extremal* if there exists $t \in K$ such that $\phi_i(t) = X(t)$. We denote by Φ_K^+ the set of extremal functions in K and by ϕ_t^+ the a.s. unique extremal function in $t \in K$

The idea of the extremal functions method is to directly sample extremal functions on K. Moreover, it is possible to show that for a stochastically continuous max-stable process X on a compact set K, Φ_K^+ is finite, then there is a chance to find an algorithm that provide an exact sample after an a.s. finite number of iterations.

The problem now is, given a set of points $t_1, \dots, t_m$, finding a way of sampling $\phi_{t_i}^+$ for $i = 1, \dots, m$. The following theorem (Theorem 1 in [DEO16]) gives us a way of iteratively sampling $\phi_{t_i}^+$. In the following we denote by P_t the distribution of the l_t-normalized spectral process V^{l_t} and we call X_n the process defined by

$$X_n(t) := \max_{1 \le i \le n} \phi_{t_i}^+(t)$$

Theorem 3.3.2.2. *The distribution of $(\phi_{t_i}^+)_{1 \le i \le m}$ is given by the following sequential procedure. The initial distribution for the extremal function $\phi_{t_1}^+$ has the same distribution as $\Gamma_1^{-1} V^{t_1}$. For $n = 2, \dots, m$ the conditional distribution of $\phi_{t_n}^+$ given $(\phi_{t_i}^+)_{1 \le i \le n-1}$ is equal to the distribution of*

$$\tilde{\phi}_{t_n}^+ = \begin{cases} \operatorname{argmax}_{\phi \in \tilde{\Phi}_n} \phi(t_n)\,, & \tilde{\Phi}_n \neq \emptyset \\ \operatorname{argmax}_{\phi \in \Phi^+_{\{t_1, \dots, t_{n-1}\}}} \phi(t_n)\,, & \tilde{\Phi}_n = \emptyset \end{cases}$$

where $\tilde{\Phi}_n$ is the set of the points of a Poisson random measure $N_{\tilde{\Phi}_n}$ with intensity

$$\mathbf{1}_{\{f(t_i) \le X_{n-1}(t_i),\, i=1,\dots,n-1\}}\, \mathbf{1}_{\{f(t_n) > X_{n-1}(t_n)\}}\, \mu(\mathrm{d}f)$$

where μ is the mean measure of the Poisson random measure

$$N = \sum_{i=1}^{\infty} \varepsilon_{\Gamma_i^{-1} V_i}.$$

In order to make this theorem useful for simulations we add two remarks.

- Given Proposition 3.3.1.1 we can equivalently sample from N using any other spectral process which satisfies the conditions of the proposition. In particular at the n-th iteration a good candidate for the spectral process is V^{t_n}.

- We can sample from $N_{\tilde{\Phi}_n}$ simply by rejecting from N the points f that do not satisfy the conditions $f(t_i) \leq X_{n-1}(t_i)$ for all $i = 1, \dots, n-1$ and $f(t_n) > X_{n-1}(t_n)$. Moreover, since N has a.s. infinitely many points, if we sample using the spectral process V^{t_n} we can use the condition $f(t_n) > X_{n-1}(t_n)$ as a stopping criterion. In fact, in this case $f(t_n) = \Gamma_i V_i^{t_n}(t_n) = \Gamma_i$. Hence, as soon as $\Gamma_i < X_{n-1}(t_n)$ the second condition cannot be satisfied anymore and we can stop generating points from N.

Algorithm 2 Extremal functions Method

procedure ExtremalFunctionsMethod(Domain $K = \{t_1, \dots, t_m\}$)
  $X(t_i) \leftarrow 0$ for all $i \in 1, \dots, m$
  for i =1,...,m **do**
    Simulate $\Gamma \sim \text{Exp}(1)$
    while $\Gamma^{-1} > X(t_i)$ **do**
      Simulate $V^{t_i} \sim F_{V^{t_i}}$
      if $k = 1$ or $\Gamma^{-1} V^{t_i}(t_k) \leq X(t_k)$ for all $k = 1, \dots, i-1$ **then**
        $X(t) \leftarrow \max\{\Gamma^{-1} V^{t_i}(t_i), X(t_i)\}$
      end if
      Simulate $E \sim \text{Exp}(1)$
      $\Gamma \leftarrow \Gamma + E$
    end while
  end for
  return X
end procedure

3.3.3 Probability threshold method

This method is a variation on the threshold method, where instead of fixing a threshold on the value of the spectral process we fix a confidence value such that the approximate process is an exact simulation. To use this method we need to find an expression of the probability that the truncated process $X^{(j)}$ defined in (3.8) is equal to X at the points of the simulation $t_1, \dots, t_m$.

Proposition 3.3.3.1. *Let X be a max-stable process with representation (3.4) and $X^{(j)}$ the truncated process defined in (3.8). Then,*

$$\begin{aligned}&\mathbb{P}\left(X(t_1) = X^{(j)}(t_1), \dots, X(t_m) = X^{(j)}(t_m) \mid X^{(j)}, \Gamma_j\right)\\ &= \exp\left\{-\mathbb{E}\left[\left(\max_{i=1,\dots,m} \frac{V(t_i)}{X^{(j)}(t_i)} - \Gamma_j\right)\mathbf{1}_{\left\{\max_{i=1,\dots,m} \frac{V(t_i)}{X^{(j)}(t_i)} \geq \Gamma_j\right\}} \;\middle|\; X^{(j)}, \Gamma_j\right]\right\} \quad (3.10)\end{aligned}$$

Observation 3.3.3.2. If our interest is to compute at a single time t the probability $\mathbb{P}\left(X(t) = X^{(j)}(t) \mid X^{(j)}\right)$, equation (3.10) becomes

$$\begin{aligned}&\mathbb{P}\left(X(t) = X^{(j)}(t) \mid X^{(j)}\right)\\ &\quad = \exp\left\{-\frac{1}{X^{(j)}(t)}\mathbb{E}\left[\left(V(t) - X^{(j)}(t)\,\Gamma_j\right)\mathbf{1}_{\left\{V(t) > X^{(j)}(t)\Gamma_j\right\}} \;\middle|\; X^{(j)}, \Gamma_j\right]\right\}. \quad (3.11)\end{aligned}$$

Where the expected value can be seen as the fair price of a European call option with strike price $S = X^{(j)}(t)\,\Gamma_j$. We will see that in the case of Brown-Resnick process we have an analytic formula to compute this value.

Proof (Proposition 3.3.3.1). Define the point process N_j as

$$N_j = \sum_{i>j} \varepsilon_{\Gamma_i^{-1} V_i}.$$

N_j is a PRM$(\nu \circ h^{-1})$ where $\nu = \text{Leb} \times F_V$ is the mean measure of the M_j, a Poisson random measure on $[0, \infty) \times \mathcal{C}_0$ defined as

$$M_j = \sum_{i>j} \varepsilon_{(\Gamma_i - \Gamma_j, V_i)}$$

and $h : [0, \infty) \times \mathcal{C}_0 \to \mathcal{C}_0$ is the function defined by

$$h(\gamma, v) := (\gamma + \Gamma_j)^{-1} v.$$

Given Γ_j and $X^{(j)}$ we have that the event

$$\{X(t_1) = X^{(j)}(t_1), \dots, X(t_m) = X^{(j)}(t_1)\}$$

is equivalent to

$$N_j\left(\left\{f \in \mathcal{C}_0 \mid \max_{i=1,\dots,m} \frac{f(t_i)}{X^{(j)}(t_i)} > 1\right\}\right) = 0.$$

Then, we want to compute its mean measure

$$\begin{aligned}
&\nu\left(h^{-1}\left(\left\{f\in\mathcal{C}_0 \mid \max_{i=1,\dots,m}\frac{f(t_i)}{X^{(j)}(t_i)}>1\right\}\right)\right)\\
=&\nu\left(\left\{(v,\gamma)\in\mathcal{C}_0\times\mathbb{R}_+ \mid \max_{i=1,\dots,m}\frac{v(t_i)}{(\Gamma_j+\gamma)X^{(j)}(t_i)}>1\right\}\right)\\
=&\int_{\mathcal{C}_0}\int_0^\infty \mathbf{1}_{\left\{\max_{i=1,\dots,m}\frac{v(t_i)}{(\Gamma_j+\gamma)X^{(j)}(t_i)}>1\right\}}\,\mathrm{d}\gamma F_V(\mathrm{d}v)\\
=&\int_{\mathcal{C}_0}\int_0^{\left(\max_{i=1,\dots,m}\frac{v(t_i)}{X^{(j)}(t_i)}-\Gamma_j\right)_+}\mathrm{d}\gamma F_V(\mathrm{d}v)\\
=&\int_{\mathcal{C}_0}\left(\max_{i=1,\dots,m}\frac{v(t_i)}{X^{(j)}(t_i)}-\Gamma_j\right)_+F_V(\mathrm{d}v)=\mathbb{E}\left[\left(\max_{i=1,\dots,m}\frac{V(t_i)}{X^{(j)}(t_i)}-\Gamma_j\right)_+\right],
\end{aligned}$$

where $V\sim F_V$ is independent of $V_1,\dots,V_j$. Then, the result follows from the fact that $\mathbb{P}(N_j(A)=0)=\exp\{-\nu(h^{-1}(A))\}$. □

The idea of the method is the following: fix a value $p\in[0,1]$, then iteratively compute $X^{(j)}$ on $t_1,\dots,t_m$ until

$$\mathbb{P}\left(X(t_1)=X^{(j)}(t_1),\dots,X(t_m)=X^{(j)}(t_m)\mid X^{(j)},\Gamma_j\right)>p.$$

The main problem with this method is that the value (3.10) is in general not known analytically and an estimation might be computationally expensive. We can think of some possible ways to overcome this difficulty.

- We can approximate (3.10) as

$$\begin{aligned}
&\mathbb{P}\left(X(t_1)=X^{(j)}(t_1),\dots,X(t_m)=X^{(j)}(t_m)\mid X^{(j)},\Gamma_j\right)\\
&\qquad\sim\prod_{i=1}^m\mathbb{P}\left(X(t_i)=X^{(j)}(t_i)\mid X^{(j)},\Gamma_j\right)
\end{aligned}$$

  which in most cases is much smaller than the original probability we want to compute, since if two points t and s are close to each other, then the events $X(t)=X^{(j)}(t)$ and $X(s)=X^{(j)}(s)$ are positively correlated.

- As a reference we can use the value $\mathbb{P}\left(X(t)=X^{(j)}(t)\mid X^{(j)},\Gamma_j\right)$ at time t. This value is in general much higher than what we would like to compute, therefore we need a higher threshold to have the same behavior. This is the most efficient solution. A possible choice for the time might be taking $t=\operatorname{argmin}_{s\in t_1,\dots,t_m}X^{(j)}(s)$

- We can use something in between the two preceding solutions: we compute $\mathbb{P}\left(X(t_i) = X^{(j)}(t_i) \mid X^{(j)}, \Gamma_j\right)$ onto a subset $I \subset \{1, \dots, m\}$, and then approximate

$$\mathbb{P}\left(X(t_1) = X^{(j)}(t_1), \dots, X(t_m) = X^{(j)}(t_m) \mid X^{(j)}, \Gamma_j\right) \sim \prod_{i \in I} \mathbb{P}\left(X(t_i) = X^{(j)}(t_i) \mid X^{(j)}, \Gamma_j\right).$$

The pseudocode can be found in Algorithm 3, where we have denoted the probability in (3.10) (or its approximation) by $P(X, \Gamma)$.

Algorithm 3 Probability Threshold Method

procedure ProbabilityThresholdMethod(Domain K, threshold p)
  $X(t) \leftarrow 0$ for all $t \in K$
  Simulate $\Gamma \sim \text{Exp}(1)$
  repeat
    Simulate $V \sim F_V$
    $X(t) \leftarrow \max\{\Gamma^{-1}V(t), X(t)\}$ for all $t \in K$
    Simulate $E \sim \text{Exp}(1)$
    $\Gamma \leftarrow \Gamma + E$
  until $P(X, \Gamma) > p$
  return X
end procedure

3.4 Simulation of the Brown-Resnick process

In this section we apply the 3 methods described in the preceding section to the Brown-Resnick process (Example 3.2.0.5). We focus on the Brown-Resnick process with unit Fréchet marginals since we can obtain all the other choices of marginals (Fréchet with $\alpha \neq 1$, Gumbel or Weibull) through simple transformations. An example of the same sample of a Brown-Resnick process with three possible marginals is in Figure 3.1.[8]

The main problem with the simulation of a Brown-Resnick process

$$X(t) = \sup_{i \geq 1} \Gamma_i^{-1} e^{W_i(t) - \frac{1}{2}\sigma^2(t)}$$

[8]In the following, we sometimes plot the Brown-Resnick process with Gumbel marginals instead of the one with Fréchet marginals, since the heavy tail of the Fréchet distribution make in some cases the figures less clear.

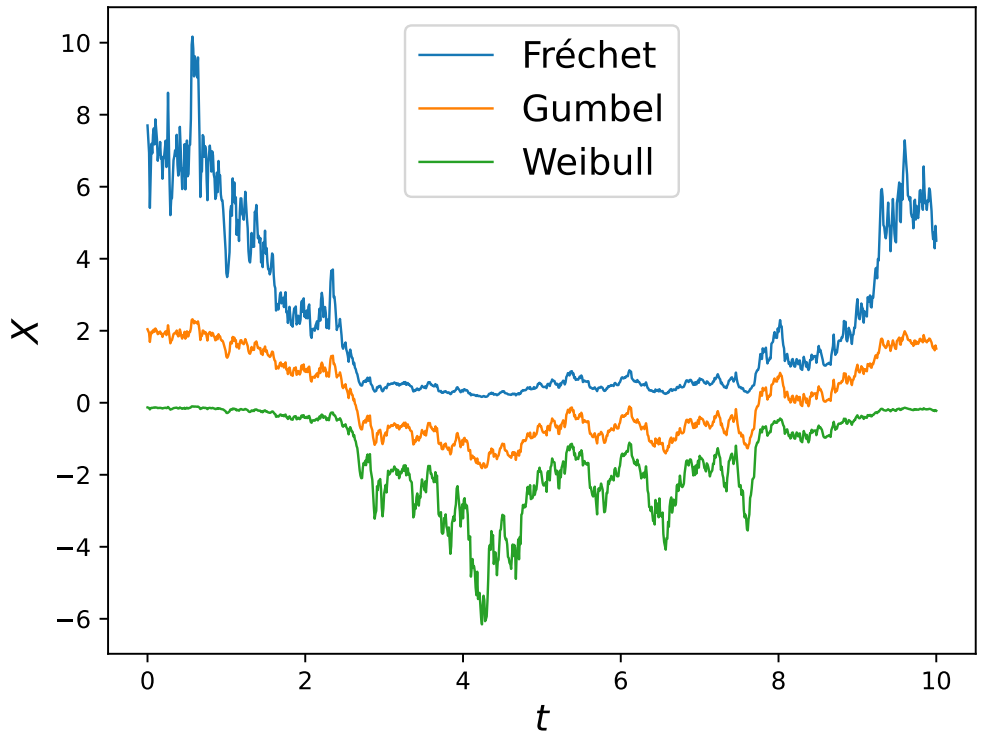


Figure 3.1: Transformations of the same sample path of a Brown-Resnick process in order to have three different marginal distributions. The blue line corresponds to Φ_1 marginals, the orange line corresponds to Λ marginals and the green line corresponds to Ψ_1 marginals.

is that, even if the spectral process

$$V = e^{W(t)-\frac{1}{2}\sigma^2(t)}$$

has $\mathbb{E}[V(t)] = 1$ for all $t \in T$, we have that $V(t) \to 0$ a.s. as $t \to \infty$. We prove that for $H = \frac{1}{2}$ using the law of the iterated logarithm (Theorem 2.2.7.1). Consider $t \in \mathbb{N}$ and write $V(t) = e^{-\frac{t}{2}} e^{\sum_{s=1}^{t}(W(s)-W(s-1))}$. Since for each $s = 1, \dots, t$, $W(s) - W(s-1) \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$, we can apply the law of iterated logarithm and say that

$$|W(t)| \sim \sqrt{2t \log\log t} \text{ a.s.} \quad \text{as } t \to \infty.$$

Hence, $W(t) - \frac{t}{2} \to -\infty$ a.s. and $V(t) \to 0$ a.s. In practice, this tells us that for big t it is not efficient to use V as spectral process in the simulations. We can also see this from formula (3.11) applied in the context of the Brown-Resnick process.

Proposition 3.4.0.1. *Let X be a Brown-Resnick process with W^H fractional Brownian motion with Hurst exponent $H \in (0,1)$, then*

$$\mathbb{E}\left[\left(V(t) - X^{(j)}(t)\,\Gamma_j\right)_+ \;\middle|\; X^{(j)}, \Gamma_j \right] = F_{\mathcal{N}}(d_1^{(H)}) - X^{(j)}(t)\,\Gamma_j F_{\mathcal{N}}(d_2^{(H)}) \tag{3.12}$$

where $V = e^{W^H(t)-\frac{1}{2}\sigma^2(t)}$ is independent of $W_1^H, \dots, W_j^H$, $F_{\mathcal{N}}$ is the cumulative distribution function of a standard normal random variable and

$$d_1^{(H)} = \frac{-\log(X^{(j)}(t)\,\Gamma_j) + \frac{1}{2}t^{2H}}{t^H}$$
$$d_2^{(H)} = \frac{-\log(X^{(j)}(t)\,\Gamma_j) - \frac{1}{2}t^{2H}}{t^H}$$

Proof. The proof only require to compute explicitly $\mathbb{E}\left[\left(V(t) - X^{(j)}(t)\,\Gamma_j\right)_+ \right]$ for general $X^{(j)}(t)$, Γ_j deterministic. In the following we write $K \coloneqq X^{(j)}(t)\,\Gamma_j$.

$$\begin{aligned}
\mathbb{E}\left[(V(t) - K)_+ \right] =& \mathbb{E}\left[\left(e^{W^H(t)-\frac{1}{2}\sigma^2(t)} - K \right)_+ \right] \\
=& \frac{1}{(2\pi\sigma^2(t))^{\frac{1}{2}}} \int_{-\infty}^{+\infty} \left(e^{w-\frac{1}{2}\sigma^2(t)} - K \right)_+ e^{-\frac{w^2}{2\sigma^2(t)}}\, \mathrm{d}w \\
=& \frac{1}{(2\pi\sigma^2(t))^{\frac{1}{2}}} \int_{\log K + \frac{1}{2}\sigma^2(t)}^{+\infty} \left(e^{-\frac{1}{2}(\frac{w}{\sigma(t)} - \sigma(t))^2} - K e^{-\frac{w^2}{2\sigma^2(t)}} \right) \mathrm{d}w \\
=& \frac{1}{\sqrt{2\pi}} \int_{\frac{\log K - \frac{1}{2}\sigma^2(t)}{\sigma(t)}}^{\infty} e^{-\frac{1}{2}z^2}\, \mathrm{d}z - \frac{K}{\sqrt{2\pi}} \int_{\frac{\log K + \frac{1}{2}\sigma^2(t)}{\sigma(t)}}^{\infty} e^{-\frac{1}{2}z^2}\, \mathrm{d}z \\
=& F_{\mathcal{N}}\left(\frac{-\log K + \frac{1}{2}\sigma^2(t)}{\sigma(t)} \right) - K F_{\mathcal{N}}\left(\frac{-\log K - \frac{1}{2}\sigma^2(t)}{\sigma(t)} \right),
\end{aligned}$$

where $\sigma(t) = |t|^H$ in the case of fractional Brownian motion.

□

From the preceding proposition we see that

$$\mathbb{E}\left[\left(V(t) - X^{(j)}(t)\,\Gamma_j\right) \mathbf{1}_{\left\{V(t) > X^{(j)}(t)\Gamma_j\right\}} \;\middle|\; X^{(j)}, \Gamma_j \right] \to 1$$

as $t \to \infty$ for fixed $X^{(j)}$ and Γ_j, hence we get from (3.11)

$$\mathbb{P}(X(t) = X^{(j)}(t) \mid X^{(j)}(t)) \sim e^{-\frac{1}{X^{(j)}(t)}}, \quad \text{as } t \to \infty,$$

but, since for any finite j, $\lim_{t\to\infty} X^{(j)}(t) = 0$ a.s. because $V_j \to 0$ a.s., we have

$$\lim_{t\to\infty} \mathbb{P}(X(t) = X^{(j)}(t)) = 0.$$

This means that for t big enough we need a lot of iterations to have enough confidence on the precision of the approximated simulation of the value $X(t)$. We can see this behavior in Figure 3.2 where we presented $\log(X^{(j)})$ (or equivalently $X^{(j)}$ for a Brown-Resnick process with Gumbel marginals). We notice that even with 10^6 iterations the process starts having a descending direction for very small t.

See the different behaviors of the Brown-Resnick process when the Gaussian process is a fractional Brownian motion for different values of the Hurst parameter H. In Figure 3.3 we see that for smaller H the sample paths are much rougher, while for higher H we can see smoother sample paths and the creations of "mountains" of high values.

3.4.1 Exact simulation methods for the Brown-Resnick process

In this subsection we provide a comparison between two methods for exact simulations of the Brown-Resnick process: the sum-normalized threshold stopping method introduced in [DM15] and extremal functions method introduced in [DEO16].

In figure 3.4, we can see two examples of sample paths simulated with the two different methods, in which we highlighted in black the extremal functions that contributed to the sample paths. The approaches of the two methods are completely different, in the threshold method the first processes $\Gamma_i V_i$ that are simulated are the ones that have the highest probability to be extremal functions. On the contrary, in the extremal functions method we start simulating processes from above, and we accept them only when they are small enough. In practice, extremal functions tend to appear later in the simulation compared to the sum-normalized method. For this reason approximate procedures using variations of these methods tends to be much more efficient for the sum-normalized threshold method, as shown in [OS22].

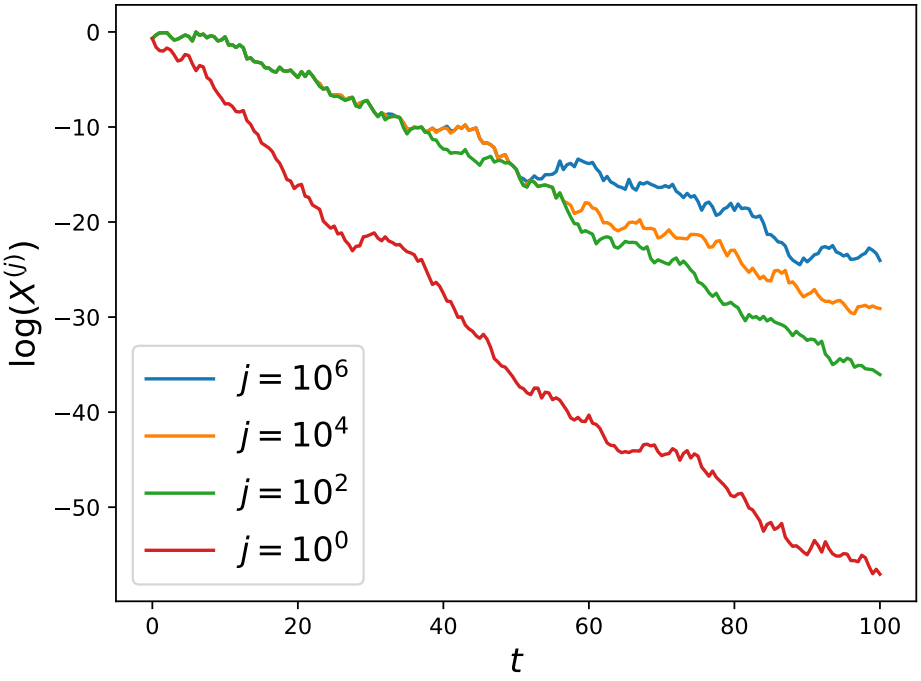


Figure 3.2: Naive simulation of a Brown-Resnick process with Gumbel marginal for different values of the truncation value j. Here the Gaussian process W is a standard Brownian motion

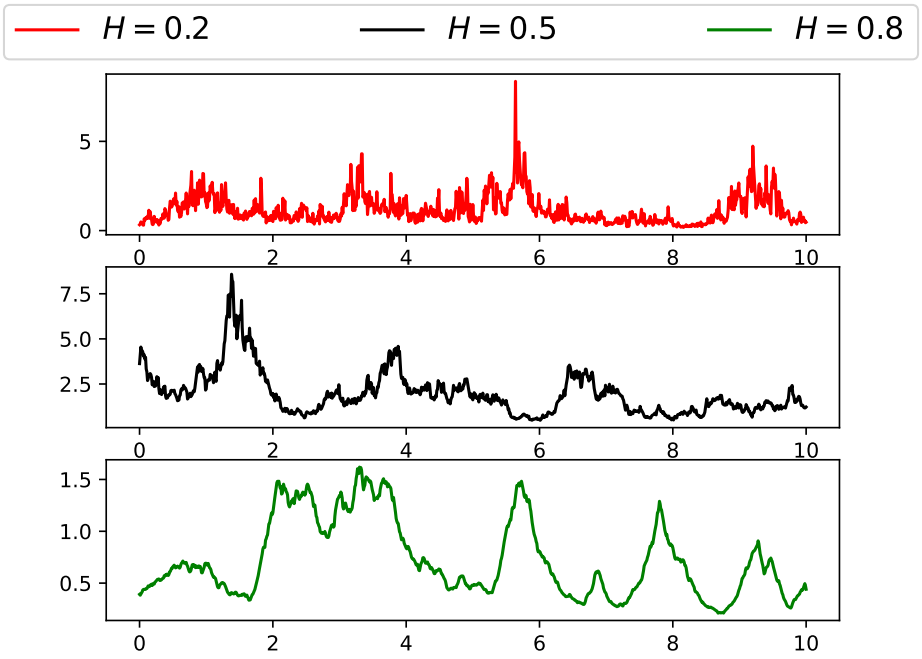
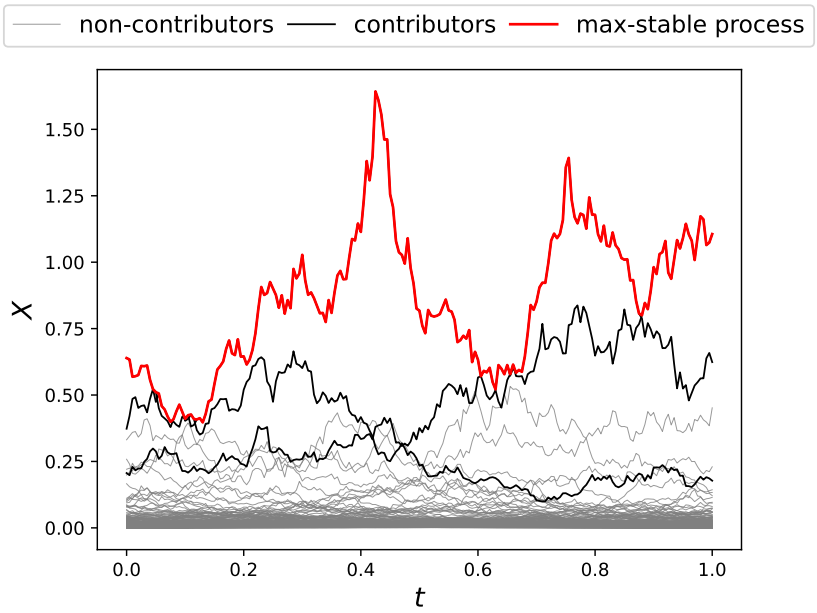


Figure 3.3: Simulations of a Brown-Resnick process with unit-Fréchet marginal for three different values of Hurst parameter H.

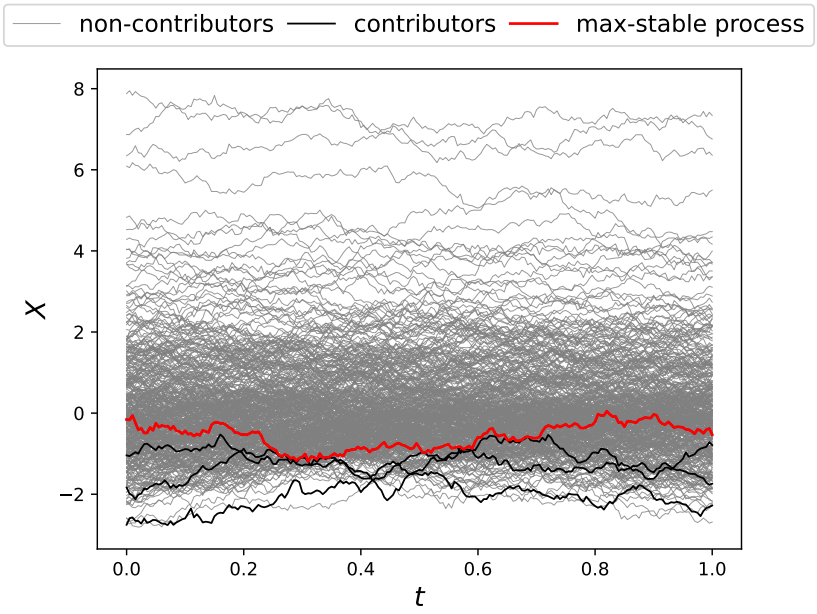
Then, we see a comparison in the efficiencies of the two methods for different values of m. In figure 3.5 we have compared the number of iterations needed to reach the stopping criteria for the two methods. We simulated sample paths on $m = 100$ and $m = 1000$ points. In the first case we simulated 1000 sample paths, while in the second, which is computationally more expensive, 100 paths. We can see that the results are consistent for the two values of m: the extreme functions method is more efficient while the sum-normalized threshold stopping method suffer both from having a higher median cost and being more subject to outliers with very high computational cost.

3.4.2 Probability threshold method for Brown-Resnick process

As we have seen in Proposition 3.4.0.1, given the peculiar structure of the Brown-Resnick process, it is possible to find an analytic expression for the probability 3.11 at a single point t. We can therefore implement the probability threshold method efficiently. Before doing the simulations, we provide an analytic estimate for the number of iterations needed to reach the probability threshold p at given time t using as spectral process $e^{W^H(t)-\frac{1}{2}\sigma^2(t)}$. As we have shown in (2.3) $\Gamma_j \sim \mathcal{O}(j)$

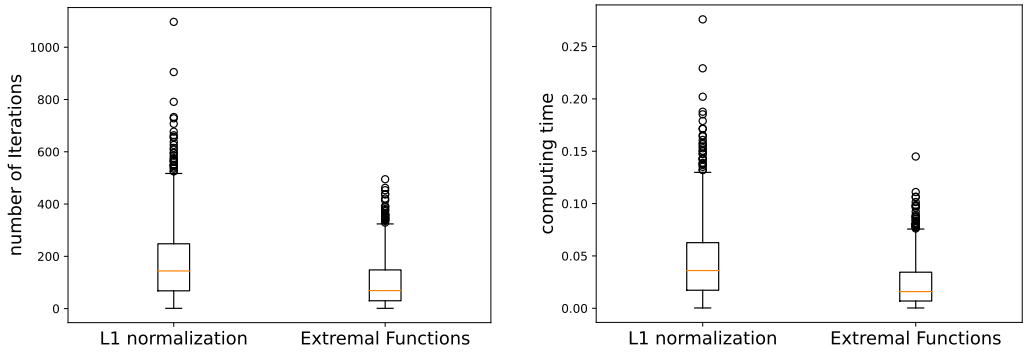


(a) Sum-normalized threshold stopping method



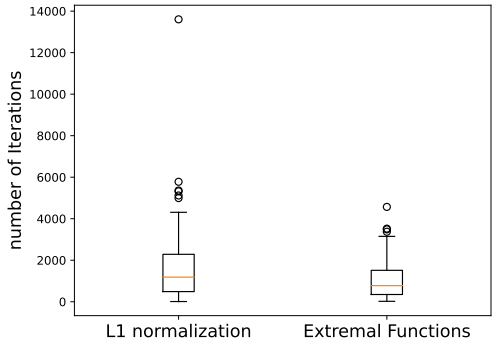
(b) Extremal function method

Figure 3.4: Realizations of two sample paths of a Brown-Resnick process using the sum-normalized threshold method and the extremal functions method. In order to make the plots more readable in the case of the extremal functions method we simulated a Brown-Resnick with Gumbel marginals. In red we see the simulation of X, in black we see the extremal functions (processes that contributed to the point-wise maxima), in grey all the other processes that have been simulated and that do not contributed to the max-stable process.

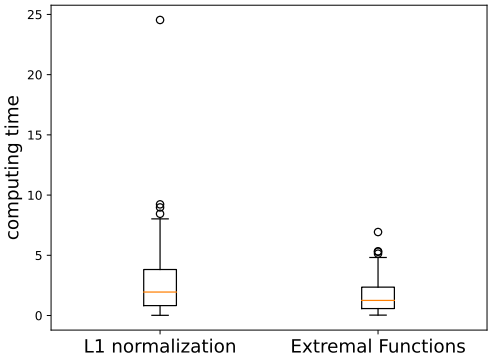


(a) Number of iterations, $m = 100$

(b) Computing time, $m = 100$



(c) Number of iterations, $m = 1000$



(d) Computing time, $m = 1000$

Figure 3.5: Box-plots comparing the efficiencies of the sum-normalized threshold method and extremal functions method. On the left-hand side we compare the number of iteration needed to reach the stopping condition. In the right column we see a comparison between their computation times. The first row uses 1000 simulations and $m = 100$ for each method. The second row uses 100 simulations and $m = 1000$ for each method.

where j is the number of iterations needed to reach convergence. We only need to estimate the order of Γ_j. Consider $X^{(j)}$ fixed (we assume $X^{(j)} = 1$), we have

$$\begin{aligned}
p &\leq \exp\left\{-\left(F_{\mathcal{N}}\left(\frac{-\log\Gamma_j + \frac{1}{2}t^{2H}}{t^H}\right) - \Gamma_j F_{\mathcal{N}}\left(\frac{-\log\Gamma_j - \frac{1}{2}t^{2H}}{t^H}\right)\right)\right\} \\
&= \exp\left\{-F_{\mathcal{N}}\left(\frac{-\log\Gamma_j + \frac{1}{2}t^{2H}}{t^H}\right)\right\} \exp\left\{\Gamma_j F_{\mathcal{N}}\left(\frac{-\log\Gamma_j - \frac{1}{2}t^{2H}}{t^H}\right)\right\} \\
&\sim \exp\left\{\frac{e^{-\left(\frac{-\log\Gamma_j + \frac{1}{2}t^{2H}}{t^H}\right)^2}}{\frac{-\log\Gamma_j + \frac{1}{2}t^{2H}}{t^H}}\right\} \exp\left\{-\Gamma_j \frac{e^{-\left(\frac{-\log\Gamma_j - \frac{1}{2}t^{2H}}{t^H}\right)^2}}{\frac{-\log\Gamma_j - \frac{1}{2}t^{2H}}{t^H}}\right\} \\
&\sim \exp\left\{\frac{e^{-\left(\frac{-\log\Gamma_j + \frac{1}{2}t^{2H}}{t^H}\right)^2}}{\frac{-\log\Gamma_j + \frac{1}{2}t^{2H}}{t^H}}\right\} \leq \exp\left\{-e^{-\left(\frac{-\log\Gamma_j + \frac{1}{2}t^{2H}}{t^H}\right)^2}\right\}
\end{aligned}$$

Then, we have

$$-\log(-\log p) \leq \left(\frac{-\log\Gamma_j + \frac{1}{2}t^{2H}}{t^H}\right)^2,$$

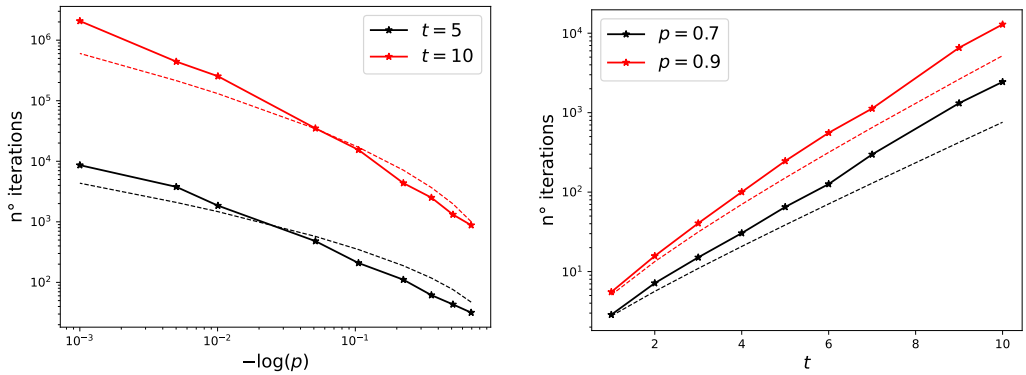
from which we get

$$\Gamma_j \geq e^{\frac{1}{2}t^{2H}}\, e^{t^H(-\log(-\log(p)))^{\frac{1}{2}}},$$

where in all this derivation we meant by $\leq$ and $\geq$, asymptotically $\leq$ and asymptotically $\geq$ for $\Gamma_j \to \infty$ (which happens when $p \to 1$ or $t \to \infty$). From this derivation we get

$$j \sim \mathcal{O}(e^{\frac{1}{2}t^{2H}}\, e^{t^H(-\log(-\log(p)))^{\frac{1}{2}}}).$$

This shows that simulating the Brown-Resnick process by using as spectral process the process $e^{W(t)-\frac{1}{2}\sigma^2(t)}$ is not efficient. In fact, for a standard Brownian motion and $t = 10$ we need $\mathcal{O}(10^5)$ iterations to get to a $p = 0.99$ if $X(t) = 1$. Comparing this with plots in Figure 3.5 makes clear that this is not an efficient method for the Brown-Resnick process even if t is relatively small. Using another spectral process is surely a better choice, but we need a way of estimating the probability (3.10) which might be done only through expensive Monte Carlo simulations making the method infeasible. In Figure 3.6 we can see the mean number of iterations depending on the value t and p computed on 1000 simulations. We see that the order of magnitude of the number of iterations is coherent with the expected theoretical order of magnitude.



(a) (b)

Figure 3.6: On the left-hand side there is a loglog plot of the number of iterations vs $-\log(p)$, for $p = 0.5, 0.6, 0.7, 0.8, 0.9, 0.95, 0.99, 0.995, 0.999$. The dotted lines represent the theoretically expected slopes, not the expected numbers of iterations. We can see that the slopes of the realized numbers of iterations are similar to the theoretical slopes, meaning that the results are consistent with the theory. On the right-hand side there is a semilogy plot of the number of iterations vs t. As before the dotted lines represent the theoretically expected slopes, not the expected numbers of iterations. Here we notice that the realized slopes tend to be steeper than the expected ones. This can be justified by the influence of samples in which $X^{(j)}(t) < X(t)$, in these cases the numbers of iterations needed to converge tend to exceed the expectations. Anyway, we can see an exponential grow in t as predicted.

List of abbreviations and symbols

Abbreviation or Symbol	**Explanation**
a.s.	almost sure, almost surely
$\mathcal{B}(E)$	the Borel σ-algebra of the set E
$\mathrm{Bin}(n,p)$	the binomial distribution with parameters (n,p)
$C_K^+(E)$	the space of non-negative continuous functions with compact support
$X \stackrel{\mathrm{d}}{=} Y$	equality in distribution, X has the same distribution as Y
$X_n \stackrel{\mathrm{d}}{\to} X$	X_n converges in distribution to X
$\mathbb{E}[X]$	mean of X
$\mathbb{E}_F[X]$	mean of X with respect to the distribution F
$\mathrm{Exp}(\lambda)$	exponential distribution with rate λ
$\Phi_\alpha(x)$	Fréchet distribution with index α
ϕ_X	characteristic function of X
Γ_n	often $\Gamma_n = E_1 + \cdots + E_n$ where $E_i \stackrel{\mathrm{iid}}{\sim} \mathrm{Exp}(1)$
$\mathrm{Gamma}(\alpha,\beta)$	gamma distribution with shape parameter α and rate β
H	often a max-stable distribution or Hurst exponent of a fractional Brownian motion
$\mathbf{1}_A$	indicator function of the set A
$\mathbf{1}_{\{X\in A\}}$	abbreviation for $\mathbf{1}_A(X)$
iff	if and only if
iid	independent and identically distributed
$(X_n)_{n\geq 1} \stackrel{\mathrm{iid}}{\sim} F$	$(X_n)_{n\geq 1}$ is a sequence of iid random variables with distribution F
$(J_n)_{n\geq 1}$	often a sequence of jumps
K	often a compact set
l	1-homogeneous functional
$\Lambda(x)$	Gumbel distribution
M, N	often a point process or a Poisson process
μ, ν	often mean measure of a Poisson process or a Lévy measure
MDA	maximum domain of attraction
$\mathbb{N}$	natural numbers

$\mathcal{N}(\mu, \sigma^2)$	normal distribution with mean μ and variance σ^2
$F_{\mathcal{N}}$	distribution of a standard normal random variable
$f = \mathrm{o}(g)$	for $x_0 \in \mathbb{R} \cup \infty$, $\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0$
$f = \mathcal{O}(g)$	for $x_0 \in \mathbb{R} \cup \infty$, $\limsup_{x \to x_0} \left\lvert \frac{f(x)}{g(x)} \right\rvert < \infty$
$\mathbb{P}$	probability measure
$X_n \xrightarrow{\mathbb{P}} X$	X_n converges in probability to X
$\mathrm{Pois}(\lambda)$	Poisson distribution with rate λ
PRM	Poisson random measure
$\Psi_\alpha(x)$	(negative) Weibull distribution
Ψ_N	Laplace functional of the point process N
$\mathcal{Q}$	Lévy measure
$\mathbb{R}$	real numbers
$\mathbb{R}_+$	$\mathbb{R} \cup [0, +\infty)$
$\mathbb{R}^d_{+,\mathbf{0}}$	$\mathbb{R}^d_+ \setminus \{\mathbf{0}\}$
$S_\alpha(\sigma, \beta, \mu)$	α-stable distribution with tail index α, scale parameter σ, skewness parameter β, location parameter μ
$\mathrm{Unif}(a, b)$	uniform distribution between a and b
U, U_n	uniform random variables between 0 and 1
V, V_n	often spectral process of a max-stable process
W, W_n	often Gaussian process with independent increments
$\mathbb{Z}$	integers
$X \sim F$	X is a random variable with distribution F
$f \sim g$	for $x_0 \in \mathbb{R} \cup \infty$, $\lim_{x \to x_0} \frac{f(x)}{g(x)} = 1$
$F^{\leftarrow}, \mu^{\leftarrow}$	generalized inverse or inverse Lévy measure

Bibliography

[AG07] S. Asmussen and P.W. Glynn. *Stochastic Simulation: Algorithms and Analysis*. Stochastic Modelling and Applied Probability. Springer New York, 2007. ISBN: 9780387690339. URL: https://books.google.dk/books?id=vMI2MdxchU0C.

[AR01] Søren Asmussen and Jan Rosiński. "Approximations of small jumps of Lévy processes with a view towards simulation". In: *Journal of Applied Probability* 38.2 (2001), pp. 482–493. DOI: 10.1239/jap/996986757.

[Bil95] P. Billingsley. *Probability and Measure*. Wiley Series in Probability and Statistics. Wiley, 1995. ISBN: 9780471007104. URL: https://books.google.dk/books?id=z39jQgAACAAJ.

[BM58] G. E. P. Box and Mervin E. Muller. "A Note on the Generation of Random Normal Deviates". In: *The Annals of Mathematical Statistics* 29.2 (1958), pp. 610–611. DOI: 10.1214/aoms/1177706645. URL: https://doi.org/10.1214/aoms/1177706645.

[Bon82] Lennart Bondesson. "On simulation from infinitely divisible distributions". In: *Advances in Applied Probability* 14 (Dec. 1982), pp. 855–869. DOI: 10.1017/S0001867800020851.

[BR77] Bruce M. Brown and Sidney I. Resnick. "Extreme Values of Independent Stochastic Processes". In: *Journal of Applied Probability* 14.4 (1977), pp. 732–739. ISSN: 00219002. URL: http://www.jstor.org/stable/3213346 (visited on 04/20/2023).

[Car+02] Peter Carr et al. "The Fine Structure of Asset Returns: An Empirical Investigation". In: *The Journal of Business* 75 (Feb. 2002). DOI: 10.1086/338705.

[CMS76] J. M. Chambers, C. L. Mallows, and B. W. Stuck. "A Method for Simulating Stable Random Variables". In: *Journal of the American Statistical Association* 71.354 (1976), pp. 340–344. DOI: 10.1080/01621459.1976.10480344. eprint: https://www.tandfonline.com/

doi/pdf/10.1080/01621459.1976.10480344. URL: https://www.tandfonline.com/doi/abs/10.1080/01621459.1976.10480344.

[DEO16] Clément Dombry, Sebastian Engelke, and Marco Oesting. "Exact simulation of max-stable processes". In: *Biometrika* 103.2 (May 2016), pp. 303–317. ISSN: 0006-3444. DOI: 10.1093/biomet/asw008. eprint: https://academic.oup.com/biomet/article-pdf/103/2/303/17460729/asw008.pdf. URL: https://doi.org/10.1093/biomet/asw008.

[DM15] A. Dieker and Thomas Mikosch. "Exact simulation of Brown-Resnick random fields at a finite number of locations". In: *Extremes* 18 (June 2015). DOI: 10.1007/s10687-015-0214-4.

[DR15] Clement Dombry and Mathieu Ribatet. "Functional regular variations, Pareto processes and peaks over threshold". In: *Statistics and Its Interface* 8 (Jan. 2015), pp. 9–17. DOI: 10.4310/SII.2015.v8.n1.a2.

[FT28] R. A. Fisher and L. H. C. Tippett. "Limiting forms of the frequency distribution of the largest or smallest member of a sample". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 24.2 (1928), pp. 180–190. DOI: 10.1017/S0305004100015681.

[Gla04] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Applications of mathematics : stochastic modelling and applied probability. Springer, 2004. ISBN: 9780387004518. URL: https://books.google.dk/books?id=e9GWUsQkPNMC.

[Haa84] L. De Haan. "A Spectral Representation for Max-stable Processes". In: *The Annals of Probability* 12.4 (1984), pp. 1194–1204. DOI: 10.1214/aop/1176993148. URL: https://doi.org/10.1214/aop/1176993148.

[Jac+05] Jean Jacod et al. "The approximate Euler method for Lévy driven stochastic differential equations". In: *Annales de l'IHP Probabilités et statistiques*. Vol. 41. 3. 2005, pp. 523–558.

[Kal02] O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer New York, 2002. ISBN: 9780387953137. URL: https://books.google.dk/books?id=L6fhXh13OyMC.

[KSH09] Zakhar Kabluchko, Martin Schlather, and Laurens de Haan. "Stationary max-stable fields associated to negative definite functions". In: *The Annals of Probability* 37.5 (2009), pp. 2042–2065. DOI: 10.1214/09-AOP455. URL: https://doi.org/10.1214/09-AOP455.

[Mik04] T. Mikosch. *Non-Life Insurance Mathematics: An Introduction With Stochastic Processes*. Non-life insurance mathematics: an introduction with stochastic processes v. 13. Springer, 2004. ISBN: 9783540406501. URL: https://books.google.dk/books?id=2n8IzUYrks4C.

[MN00] Thomas Mikosch and Rimas Norvaiša. "Stochastic integral equations without probability". In: *Bernoulli* (2000), pp. 401–434.

[OS22] Marco Oesting and Kirstin Strokorb. "A Comparative Tour through the Simulation Algorithms for Max-Stable Processes". In: *Statistical Science* 37.1 (2022), pp. 42–63. DOI: 10.1214/20-STS820. URL: https://doi.org/10.1214/20-STS820.

[Pet95] V.V. Petrov. *Limit Theorems of Probability Theory: Sequences of Independent Random Variables*. Oxford science publications. Clarendon Press, 1995. ISBN: 9780198534990. URL: https://books.google.dk/books?id=4LkdSaI4xXMC.

[Pro13] P. Protter. *Stochastic Integration and Differential Equations: A New Approach*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2013. ISBN: 9783662026199. URL: https://books.google.dk/books?id=fMruCAAAQBAJ.

[PT97] Philip Protter and Denis Talay. "The Euler scheme for Lévy driven stochastic differential equations". In: *The Annals of Probability* (1997), pp. 393–423.

[Res03] S. Resnick. *A Probability Path*. Modern Birkhäuser Classics. Birkhäuser Boston, 2003. ISBN: 9780817640552. URL: https://books.google.dk/books?id=T9-PMDSVDNsC.

[Res87] S.I. Resnick. *Extreme Values, Regular Variation, and Point Processes*. Applied probability : a series of the applied probability trust. Springer, 1987. ISBN: 9783540964810. URL: https://books.google.dk/books?id=DXi1QgAACAAJ.

[Ros01] Jan Rosiński. "Series Representations of Lévy Processes from the Perspective of Point Processes". In: *Lévy Processes: Theory and Applications*. Ed. by Ole E. Barndorff-Nielsen, Sidney I. Resnick, and Thomas Mikosch. Boston, MA: Birkhäuser Boston, 2001, pp. 401–415. ISBN: 978-1-4612-0197-7. DOI: 10.1007/978-1-4612-0197-7_18. URL: https://doi.org/10.1007/978-1-4612-0197-7_18.

[Sat99] K. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge studies in advanced mathematics. Cambridge University Press, 1999. ISBN: 9780521553025. URL: https://books.google.dk/books?id=tbZPLquJjSoC.