LÉVY-DRIVEN QUEUES

K. DȨBICKI AND M. MANDJES

ABSTRACT. This survey addresses the class of queues with Lévy input, which covers the classical M/G/1 queue and reflected Brownian motion as special cases. First the stationary behavior is treated, with special attention to the case of the input process having one-sided jumps (i.e., spectrally one-sided Lévy processes). Then various transient metrics are focused on (such as the transient distribution, the busy period, and the workload correlation function). Distinguishing between light-tailed and heavy-tailed input, we give an account of results on the tail of the workload distribution; in addition we present the main asymptotic results for the various transient quantities. We then extend our basic model to various more advanced queueing systems: queues with a finite buffer, queues in which the current buffer level affects the characteristics of the Lévy input ('feedback'), and polling type of models. The last part of the survey considers networks of queues: starting with the tandem queue, we subsequently describe the stationary behavior of a general class of Lévy-driven queueing networks. At the methodological level, a variety of techniques has been used, such as transform-based techniques, martingales, rate-conservation arguments, change-of-measure, importance sampling, and large deviations.

KEYWORDS. Queues ⋆ Lévy processes ⋆ Laplace transforms ⋆ transient behavior ⋆ tail asymptotics ⋆ networks

Date: December 1, 2010.

2010 *Mathematics Subject Classification.* Primary 60K25, 60G51; Secondary 90B05.

The research of the first author was supported by MNiSW Research Grant N N201 394137 (2009-2011). The authors thank the Isaac Newton Institute, Cambridge, UK, for hospitality.

CONTENTS

1. Introduction 3
2. Lévy processes, Lévy driven queues 5
2.1. Infinitely divisible distributions, Lévy processes 5
2.2. Spectrally one-sided Lévy processes 5
2.3. α-stable Lévy motions 7
2.4. Lévy-driven queues 8
3. Steady-state workload 9
3.1. Spectrally-positive case 9
3.2. Spectrally-negative case 12
4. Transient workload 12
4.1. Transform of transient distribution 13
4.2. Busy period 14
4.3. Correlation function 16
4.4. Infimum over given time interval 18
5. Workload asymptotics 18
5.1. Light-tailed regime 19
5.2. Intermediate regime 20
5.3. Heavy-tailed regime 20
6. Transient asymptotics, rare-event simulation 22
6.1. Asymptotics of transient metrics 22
6.2. Asymptotics of joint transient distribution 23
6.3. Rare-event simulation, importance sampling 24
7. Variants of the standard queue 27
7.1. Finite-buffer queues 27
7.2. Models with feedback 28
7.3. Vacation and polling models 28
7.4. Models with Markov-additive input 29
8. Lévy-driven tandem queues 31
8.1. A representation of the downstream workload distribution 31
8.2. Steady-state workload of the downstream queue 32
8.3. Downstream workload asymptotics 34
8.4. Bivariate distribution 35
9. Networks 36
9.1. Definition, multi-dimensional Skorokhod problem 37
9.2. Lévy-driven tree networks 37
9.3. Representation for the stationary workload 38
10. Concluding remarks 40
References 41

1. INTRODUCTION

The class of Lévy processes consists of all stochastic processes with stationary independent increments, and as such it covers well-studied processes such as for instance Brownian motions and Poisson processes. In this sense Lévy processes can be seen as the genuine continuous-time counterpart of the random walk $S_n := \sum_{i=1}^{n} \xi_i$, with independent and identically distributed ξ_i.
Lévy processes play an increasingly important role in various application domains, ranging from finance to biology. A brief account of the history of Lévy processes (initially simply known as 'processes with stationary and independent increments') and its application fields is given in [7]. In finance, Lévy processes are being used intensively to analyze various phenomena; they are for instance suitable when studying credit risk, or for option pricing purposes [34], but play a pivotal role in insurance mathematics as well [10]. An attractive feature of Lévy processes, particularly with applications in finance in mind, is that this class is rich in terms of possible path structures: it is perhaps the simplest class of processes that allows sample paths to have continuous parts interspersed with jumps at random epochs.
Another important application domain lies in operations research. According to the functional central limit theorem, under mild conditions on the distribution of the increments, a scaled version of discrete-time random walks converges to a Brownian motion. In line with this one can argue that, under a suitable scaling, there is convergence of 'classical' GI/G/1 queueing systems (with discrete customers) to a queue with Brownian input [96], also often referred to as *reflected Brownian motion*.
A more specific example, in which the limiting process is *not* necessarily Brownian motion, relates to the performance analysis of resources in communication networks. In the mid-nineties it was observed that the sizes of the documents transferred over the internet obey heavy-tailed distributions. This entails that under particular conditions, see [75, 94] and [96, Ch. 4], the aggregate of traffic generated by many users converges to fractional Brownian motion, but under other conditions there is convergence to (a specific class of) Lévy processes. In the latter regime, the performance of the network element can be evaluated by analyzing a queue fed by Lévy input.
The above considerations underscore the importance of analyzing queues with Lévy input (or: *Lévy-driven queues*). It should be noticed, though, that it is not *a priori* clear what should be understood by such a queue: for instance in case the Lévy process under consideration is Brownian motion, the input process is not increasing, and therefore the corresponding queue cannot be seen as a storage system in the classical sense. Relying on a description of the queue as the solution of a so-called Skorokhod problem [96], however, a formal definition of a Lévy-driven queue can be given (in fact, any stochastic process can serve as input). It is stressed that queues of the 'classical' M/G/1 type (that is, Poisson arrivals, generally distributed jobs, one server) fit in the framework of Lévy-driven queues.

Having defined Lévy-driven queues, all questions that have been studied for classical queues now have their Lévy counterpart. A first question relates to the distribution of the steady-state workload of the queue, but also issues regarding the transient distribution, the busy period, the queue's correlation structure can be assessed. In addition, just as in the world of 'classical' queues, one can think of a variety of variants of the standard Lévy-driven queue: queues with a finite buffer, queues whose input characteristics are affected by the current workload level ('feedback'), queues with vacations and service interruptions, and Lévy-driven polling models. Finally, under specific conditions on the Lévy processes involved, one can let the output of a queue serve as the input for a next queue, and in this way we arrive at a notion of Lévy-driven queueing networks.

The objective of this survey is to give an account of the literature on Lévy-driven queues. In addition, we also intend to give an impression of the wide set of techniques that has been developed over the past decades. In this survey techniques that are highlighted include transform-based techniques, martingales, rate-conservation arguments, change-of-measure, importance sampling, and large deviations.
A few words on additional recommended literature. In the first place there are the textbooks by Bertoin [26], Kyprianou [67], and Sato [90], which provide a fairly general account of the theory of Lévy processes. All of these have a specific focus, though, on *fluctuation theory*, which can be understood as the theory that describes the extreme values that are attained by the Lévy process under consideration, and which is a topic that is intimately related to Lévy-driven queues. We also mention the book by Applebaum [6], that concentrates more on stochastic calculus. The book chapters [8, Ch. IX] and [83, Ch. 4] present elements of Lévy-driven queues.

The survey is organized as follows. Section 2 formalizes the notion of Lévy-driven queues; it is argued how in general queues can be defined without assuming that the input process is nondecreasing. We also introduce the special class of *spectrally one-sided* Lévy inputs, that is, Lévy processes with either only positive jumps or only negative jumps; we will extensively rely on this notion throughout the survey. In Section 3 we characterize the steady-state workload Q. For spectrally-positive input this is done through its Laplace transform, which is a result that dates back to [98] and which is commonly referred to as the 'generalized Pollaczek-Khinchine formula', whereas the spectrally-negative case can be dealt with explicitly.
Then, in Section 4, four metrics are analyzed that relate to the transient workload. In the first place we characterize (in terms of transforms) the distribution of the workload Q_t at time t, conditional on $Q_0 = x$. We then consider the busy period distribution: how long does it take for the queue to idle? A variety of techniques is used to analyze the correlation between Q_0 and Q_t, assuming the queue is in stationarity at time 0; specifically, it covers the structural result that the workload correlation function is positive, decreasing and convex. The last part of this section addresses the distribution of the lowest value attained by the workload process in an interval of given length.
Where the full distribution of Q was uniquely characterized in Section 3, Section 5 considers its tail asymptotics. Distinguishing between Lévy processes with light-tailed and heavy-tailed features, functions $f(\cdot)$ are identified such that $\mathbb{P}(Q > u)/f(u) \to 1$ as $u \to \infty$. In Section 6 we present asymptotics related to the transient metrics that we defined earlier. In addition, we point out how importance-sampling based simulation are of great help when estimating rare-event probabilities (and small covariances).
Where the previous section considered the standard Lévy driven queue, Section 7 presents results on several variants: queues with a finite buffer, queues in which the current buffer level affects the characteristics of the Lévy input ('feedback'), and vacation and polling type of models. Then, Section 8 presents results on Lévy-driven tandem queues: the output of the 'upstream queue' serves as input for the 'downstream queue'. For this model the joint steady state workload is determined, and various special cases are considered. Section 9 is devoted to networks, with particular focus

on feed-forward tree structures. In Section 10, that concludes this survey, a brief discussion of the state of the art is given.

2. LÉVY PROCESSES, LÉVY DRIVEN QUEUES

In classical queueing systems, there is the notion of customers (or work) arriving at a server that is working in a certain predefined manner. As Lévy processes are to be understood as processes with stationary independent increments (covering for instance Brownian motion), it is not immediately clear how a queue with Lévy input should be defined.

In this section we introduce the notion of Lévy-driven queues, by first providing an explicit definition of Lévy processes, and then extending the classical definition of a queue to a notion that can be used for general input processes as well (i.e., any real-valued stochastic process can serve as input). For more background, we refer to [6, 8, 67, 90].

2.1. **Infinitely divisible distributions, Lévy processes.** We say that a continuous-time process $(X_t)_t$ is a Lévy process if it has stationary and independent increments, with $X_0 = 0$ and càdlàg sample paths a.s. (*cádlág* meaning 'continuous from right, limits from left'). This definition implies that X_t is, for any t, *infinitely divisible*: we have the distributional equality, with $X_t^{(i)}$ i.i.d. copies of X_t:

$$X_t \stackrel{\mathrm{d}}{=} \sum_{i=1}^{n} X_{t/n}^{(i)},$$

for any $n \in \mathbb{N}$. Informally, each Lévy process can be associated with an infinitely divisible distribution, and vice versa. One can alternatively say that, for any value of t, $\log \mathbb{E}e^{sX_t} = t \log \mathbb{E}e^{sX_1}$, where $s \in \mathbb{C}$. It is possible to characterize Lévy processes more specifically: the so-called *Lévy exponent* $\log \mathbb{E}e^{sX_1}$ is necessarily of the form

$$\log \mathbb{E}e^{sX_1} = sd + \frac{1}{2}s^2\sigma^2 + \int_{-\infty}^{\infty} (e^{sx} - 1 - sx1_{[0,1)}(|x|))\Pi(\mathrm{d}x), \tag{2.1}$$

where $d \in \mathbb{R}$, $\sigma \geq 0$ and the spectral measure $\Pi(\cdot)$ satisfies

$$\int_{\mathbb{R}\setminus\{0\}} \min\{x^2, 1\}\Pi(\mathrm{d}x) < \infty.$$

The triplet (d, σ^2, Π) is commonly referred to as the *characteristic triplet*, as it uniquely defines the Lévy process.

It is immediately seen that this class of processes contains e.g. the Poisson process ($\log \mathbb{E}e^{sX_1} = \lambda(e^s - 1)$ for $\lambda > 0$) and Brownian motion ($\log \mathbb{E}e^{sX_1} = \frac{1}{2}\sigma^2 s^2$ for $\sigma^2 > 0$) as special cases; later on we mention various other examples. For obvious reasons, we call the first parameter of the characteristic triplet, d, the deterministic drift, whereas the term $\frac{1}{2}s^2\sigma^2$ is often referred to as the Brownian term. The third term in (2.1) corresponds to the jumps of the Lévy process; if these are only in the upward (downward) direction, then the integral should be over $[0, \infty)$ ($(-\infty, 0]$, respectively).

2.2. **Spectrally one-sided Lévy processes.** Let $(X_t)_{t\geq 0}$ be a Lévy process, with 'mean drift' $\mathbb{E}X_1 < 0$. In this survey, we specifically focus on two cases.

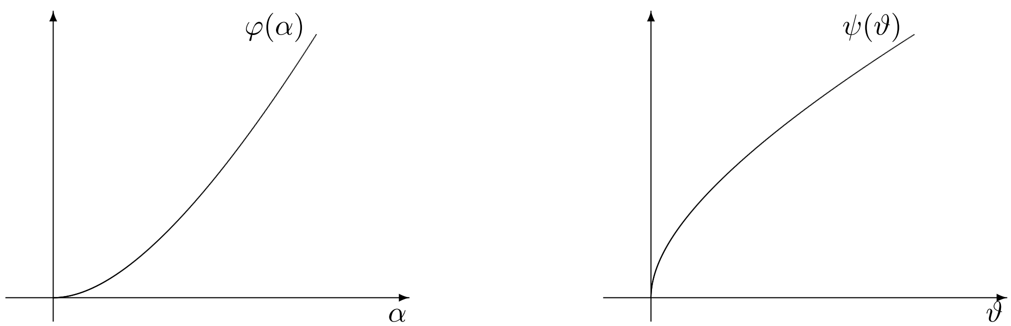

FIGURE 1. Spectrally-positive case.

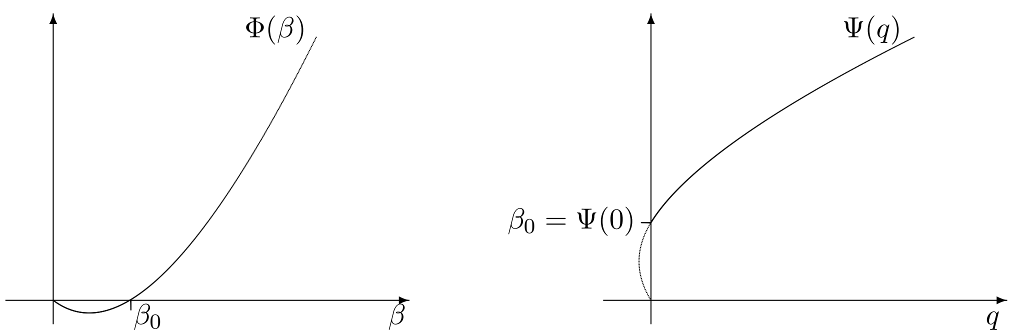

FIGURE 2. Spectrally-negative case.

(A) $(X_t)_{t\geq 0}$ has no negative jumps, or is *spectrally positive*; we write $X \in \mathscr{S}_+$. Then the *Laplace exponent* is given by the function $\varphi(\cdot) : [0,\infty) \mapsto [0,\infty)$, defined through $\varphi(\alpha) := \log \mathbb{E}e^{-\alpha X_1}$. It is a straightforward consequence of Hölder's inequality that $\varphi(\cdot)$ is convex on $[0,\infty)$; due to the assumption $\mathbb{E}X_1 < 0$, and observing that $\varphi(\cdot)$ has slope $\varphi'(0) = -\mathbb{E}X_1$ in the origin, we conclude that $\varphi(\cdot)$ is increasing on $[0,\infty)$. Therefore the inverse $\psi(\cdot)$ of $\varphi(\cdot)$ is well-defined on $[0,\infty)$. In the sequel we also require that X_t is not a *subordinator*, i.e., a monotone process; thus X_1 has probability mass on the positive half-line, which implies that $\lim_{\alpha\to-\infty} \varphi(\alpha) = \infty$.

(B) $(X_t)_{t\geq 0}$ has no positive jumps, or is *spectrally negative*; we write $X \in \mathscr{S}_-$. Now we define $\Phi(\beta) := \log \mathbb{E}e^{\beta X_1}$, which is well-defined for any $\beta \geq 0$. Again ruling out that X_t is a subordinator (and recalling that $\Phi'(0) = \mathbb{E}X_1 < 0$), we see that $\Phi(\beta)$ is *no* bijection on $[0,\infty)$; we define the *right* inverse through $\Psi(q) := \sup\{\beta \geq 0 : \Phi(\beta) = q)$. Realize that $\beta_0 := \Psi(0) > 0$.

Important examples of such Lévy processes are the following. (1) *Brownian motion with drift*, being actually both spectrally positive and negative (as its sample paths are continuous a.s.). We write

$X \in \mathbb{B}\mathrm{m}(\mu, \sigma^2)$ when $\varphi(\alpha) = -\alpha\mu + \frac{1}{2}\alpha^2\sigma^2$. The mean drift of this process is μ, which was assumed to be negative. (2) *Compound Poisson with drift*, which is spectrally positive. Non-negative jobs arrive according to a Poisson process of rate λ; the jobs $B_1, B_2, \ldots$ are i.i.d. samples from a distribution with Laplace transform $b(\alpha) := \mathbb{E}e^{-\alpha B}$; the storage system is continuously depleted at a rate r. We write $X \in \mathbb{C}\mathrm{P}(r, \lambda, b(\cdot))$; it can be verified that $\varphi(\alpha) = r\alpha - \lambda + \lambda b(\alpha)$. The mean drift of this process is $\mathbb{E}X_1 = \lambda\,\mathbb{E}B - r$, which we assume to be negative. Clearly, if the depletion rate r would be positive, and the jobs would be i.i.d. samples from a non-positive distribution (that is, the jumps are downward), then the resulting process is spectrally negative.

We conclude this subsection with two lemmas. First consider $X \in \mathscr{S}_+$, and let $\tau(x) := \inf\{t \geq 0 : X_t \leq -x\}$. Observe that $e^{-\varphi(\alpha)t}\,e^{-\alpha X_t}$ is a mean-1 martingale [97]. Noticing that $X_{\tau(x)} = -x$ (which is a direct consequence of the process not having jumps in the downward direction), and, assuming that $\mathbb{E}X_1 < 0$, it holds that $\tau(x) < \infty$ almost surely. Then 'optional sampling' [97, Ch. A14] implies the following property, which will be useful later on; observe that it entails that $\tau(x)$ is an (increasing) Lévy process.

Lemma 1. *Let $X \in \mathscr{S}_+$, and $\mathbb{E}X_1 < 0$. For $\vartheta \geq 0$, $x > 0$,*

$$\mathbb{E}e^{-\vartheta\tau(x)} = e^{-\psi(\vartheta)x}.$$

There is no immediate counterpart of this lemma for $X \in \mathscr{S}_-$, in that no explicit expression for $\mathbb{E}e^{-\vartheta\tau(x)}$ can be given in the spectrally-negative case. It *is* possible though to uniquely characterize the distribution of $\tau(x)$ through a so-called *double transform*.

Lemma 2. *Let $X \in \mathscr{S}_-$, and $\mathbb{E}X_1 < 0$. For $q \geq 0$, $x > 0$, $\beta > 0$,*

$$\int_0^\infty e^{-\beta x}\mathbb{E}e^{-q\tau(x)}\mathrm{d}x = \frac{1}{\beta}\left(1 - \frac{q}{\Psi(q)}\frac{\Psi(q) - \beta}{q - \Phi(\beta)}\right).$$

The proof of the above lemma relies on part (ii) of [67, Exercise 6.7], which states that

$$\int_0^\infty e^{-\beta x}\mathbb{E}e^{-q\tau(x)}\mathrm{d}x = \frac{\hat{\kappa}(q,\beta) - \hat{\kappa}(q,0)}{\beta\hat{\kappa}(q,\beta)};$$

here $\hat{\kappa}(q,\beta)$ relates to the transform of the so-called *descending ladder process*, and is given, in this spectrally-negative case, by $\hat{\kappa}(q,\beta) = (q - \Phi(\beta))/(\Psi(q) - \beta)$.

2.3. **α-stable Lévy motions.** This subsection focuses on a subclass of Lévy processes that has attracted substantial attention in the literature: α-stable Lévy motions. This class of models is particularly suitable when modelling various sorts of heavy-tailed phenomena [89].

To introduce these processes, we first define the class of stable distributions. We here follow the exposition in [48], but various other parameterizations are possible [95]. We say that a random variable Y has a stable distribution if for any $a, b > 0$ there exist $c > 0$ and $d \in \mathbb{R}$ such that

$$aY' + bY'' \stackrel{\mathrm{d}}{=} cY + d,$$

where Y' and Y'' are independent copies of Y. It turns out [28, Thm. 8.3.2] that the characteristic function of Y can be written in the form

$$\log\mathbb{E}e^{\mathrm{i}\theta Y} = \begin{cases} -\sigma^\alpha|\theta|^\alpha(1 - \mathrm{i}\beta\mathrm{sign}(\theta)\tan(\pi\alpha/2)) + \mathrm{i}m\theta & \alpha \neq 1; \\ -\sigma|\theta|(1 + \mathrm{i}\beta\pi/2\mathrm{sign}(\theta)\log|\theta|) + \mathrm{i}m\theta & \alpha = 1, \end{cases}$$

where $\alpha \in (0,2]$, $\beta \in [-1,1]$, $\sigma \in [0,\infty)$, $m \in \mathbb{R}$, and $\mathrm{sign}(x) = 1_{(0,\infty)}(x) - 1_{(-\infty,0)}(x)$. We write that Y is distributed $S_\alpha(\sigma,\beta,m)$.

Let us consider the meaning of the parameters in more detail.

- The parameter α is commonly referred to as the *index of stability*. Later we will observe that α is directly related to the 'heaviness' of the tail distribution. In particular, if $\alpha \in (0,1]$, then $\mathbb{E}|Y| = \infty$ (for $\alpha = 1$ we have the Cauchy distribution). For $\alpha = 2$ we obtain the Normal distribution.
- The parameter β is known as the *skewness*. The extreme cases are $\beta = 1$, corresponding to a *totally skewed to the right* distribution, and $\beta = -1$, which corresponds to a *totally skewed to the left* distribution. For $\alpha < 1$, $m = 0$ and $\beta = 1$ ($\beta = -1$, respectively), the support of the distribution is the positive (negative) half-line, but this is no longer true for $\alpha \geq 1$. The choice of $\beta = 0$ and $m = 0$ leads to a symmetric distribution.
- For obvious reasons, σ is called the *scale parameter*.
- For $\alpha \in (1,2]$, we have that $\mathbb{E}Y = m$. This explains why m is called the *shift parameter*.

The following useful property, describing the distribution's tail asymptotics, can be found in e.g. [89, p. 16]. The Gamma function $\Gamma(\cdot)$ is defined in the usual way.

Proposition 1. *Let $Y \stackrel{\mathrm{d}}{=} S_\alpha(\sigma,\beta,m)$ with $\beta \in (-1,1]$. Then, as $u \to \infty$,*

$$\mathbb{P}(Y > u)u^\alpha \to C_{\alpha,\sigma}\left(\frac{1+\beta}{2}\right),$$

where

$$C_{\alpha,\sigma} := \begin{cases} \sigma^\alpha(1-\alpha)/\left(\Gamma(2-\alpha)\cos(\pi\alpha/2)\right) & \alpha \neq 1; \\ 2\sigma/\pi & \alpha = 1. \end{cases}$$

Having defined stable distribution, we can now introduce α-stable Lévy motions, as follows. We say that $(X_t)_t$ is an α-stable Lévy motion if $(X_t)_t$ has stationary independent increments such that

$$X_t \stackrel{\mathrm{d}}{=} S_\alpha(t^{1/\alpha},\beta,m);$$

we write $X \in \mathbb{S}(\alpha,\beta,m)$. From the above we conclude that if $\beta = \pm 1$, then $X \in \mathscr{S}_\pm$.

One could say that α-stable Lévy motions are *self-similar*: picking $m = 0$, and writing $(X_t^{(\alpha)})_t$ to stress the dependence on α, one has that

$$\left(X_{Mt}^{(\alpha)}\right)_t \stackrel{\mathrm{d}}{=} \left(M^{1/\alpha}X_t^{(\alpha)}\right)_t$$

(unless $\alpha = 1$, $\beta \neq 0$). In other words: when zooming in, one essentially sees the same pattern, given that one adjusts the axes in a suitable fashion.

2.4. **Lévy-driven queues.** Having defined Lévy processes, we now introduce the notion of queues with Lévy input (or: *Lévy-driven queues*). Notice, however, that these definitions are by no means restricted to the Lévy framework; based on the formalism defined below, one can define for *any* real-valued stochastic process the corresponding workload process. We provide two types of characterizations.

In the first approach, we define the Lévy-driven queue as the continuous-time counterpart of the classical discrete-time queue. In discrete time, a queue can be described through the well-known Lindley recursion: we have

$$Q_{n+1} = \max\{Q_n + Y_n, 0\}.$$

Iterating this recursion, we obtain $Q_{n+1} = \max\{Q_{n-1} + Y_{n-1} + Y_n, Y_n, 0\}$. With $X_n := \sum_{i=0}^n Y_i$, this eventually leads to, with $Q_0 = x$,

$$Q_n = X_n + \max\left\{x, \max_{0 \le i \le n} -X_i\right\}.$$

Then a queue in continuous-time can be defined by just taking the continuous-time counterpart of the above, so that we obtain

$$Q_t = X_t + \max\{x, L_t\},\ t \ge 0,$$

with

$$L_t := \sup_{0 \le s \le t} -X_s = -\inf_{0 \le s \le t} X_s;$$

this increasing process L_t is often referred to as *local time*. Assuming the queue has been running from $-\infty$, one can alternatively write $Q_t = \sup_{s \le t}(X_t - X_s)$; in case of input processes $(X_t)_t$ with stationary increments (as is the case in our Lévy context) it needs to be assumed that $\mathbb{E}X_1 < 0$ in order to ensure stability (which we do throughout this survey). If the input process X_t is reversible (which is true in the Lévy case), then we have the following distributional equality for the stationary workload Q, commonly attributed to Reich [84]:

$$Q \stackrel{\mathrm{d}}{=} \sup_{t \ge 0} X_t. \tag{2.2}$$

An alternative is to define the Lévy-driven queue as the solution of a so-called Skorokhod problem; then one commonly says that $(Q_t)_t$ is *the reflection of* $(X_t)_t$ *at 0*. This is done as follows. Let $(L_t^\star)_t$ be a nondecreasing right-continuous process such that the following two requirements are fulfilled:

(A) $(Q_t)_t$, given by $Q_0 = x$ and $Q_t = X_t + L_t^\star$, is non-negative for all $t \ge 0$;

(B) $L_t^\star$ can only increase when $Q_t = 0$, that is

$$\int_0^T Q_t \mathrm{d}L_t^\star = 0, \quad \text{for all } T > 0.$$

Observe that it is natural to impose these conditions on a queueing process (in the M/G/1 context, the process $L_t^\star$ directly relates to the queue's idle time, and then requirement B essentially says that this idle time only increases when the buffer is empty).

Then it can be proven that the only process satisfying these conditions is $L_t^\star = \max\{x, L_t\}$, so that $Q_t = X_t + \max\{x, L_t\}$ for $t \ge 0$, where L_t is defined as above. Conclude that the definition found in this way coincides with the one obtained when taking the continuous counterpart of the discrete-time definition.

3. Steady-state workload

In this section we analyze the distribution of the stationary workload Q. We distinguish between the spectrally-positive and spectrally negative case.

3.1. **Spectrally-positive case.** We first consider the special case of compound Poisson input and constant depletion rate r; assume $\lambda \mathbb{E}B < r$. For any $x > 0$, a rate conservation argument yields that the density $f_Q(\cdot)$ of the steady-state workload satisfies

$$rf_Q(x) = \lambda\left(\int_{(0,x)} f_Q(y)\mathbb{P}(B > x - y)\mathrm{d}y + p_0\mathbb{P}(B > x)\right),$$

with $p_0 := \mathbb{P}(Q = 0)$; the left-hand (right-hand) side represents the 'probability flux' into (out of) the set $(0, x]$. Hence

$$\begin{aligned}\bar\kappa(\alpha) &:= \int_{(0,\infty)} e^{-\alpha x} f_Q(x)\mathrm{d}x \\ &= \frac{1}{r}\int_{(0,\infty)} e^{-\alpha x}\lambda\left(\int_{(0,x)} f_Q(y)\mathbb{P}(B > x-y)\mathrm{d}y + p_0\mathbb{P}(B > x)\right)\mathrm{d}x,\end{aligned}$$

which after elementary calculus (interchange the integrals, integration by parts) reduces to

$$r\bar\kappa(\alpha) = \lambda\left(\bar\kappa(\alpha) + p_0\right)\frac{1 - b(\alpha)}{\alpha}.$$

Realizing that $\kappa(\alpha) := \mathbb{E}e^{-\alpha Q} = p_0 + \bar\kappa(\alpha)$ and $\kappa(\alpha) \to 1$ as $\alpha \downarrow 0$, we conclude that $p_0 = (1 - \lambda\,\mathbb{E}B/r)$, so that we arrive at the following theorem, usually referred to as the Pollaczek-Khintchine formula [64, 81].

Theorem 1. *Let $X \in \mathbb{C}\mathrm{P}(r, \lambda, b(\cdot))$. For $\alpha \geq 0$,*

$$\kappa(\alpha) := \mathbb{E}e^{-\alpha Q} = \frac{r\alpha p_0}{r\alpha - \lambda(1 - b(\alpha))} = \frac{\alpha(r - \lambda\,\mathbb{E}B)}{r\alpha - \lambda(1 - b(\alpha))}.$$

Remark 1. Let $B_1^{\mathrm{res}}, B_2^{\mathrm{res}}, \ldots$ be i.i.d. samples from the residual lifetime distribution of B, that is

$$\mathbb{P}(B^{\mathrm{res}} \leq x) = \frac{1}{\mathbb{E}B}\int_0^x \mathbb{P}(B > y)\mathrm{d}y.$$

Realizing that $b^{\mathrm{res}}(\alpha) := \mathbb{E}e^{-\alpha B^{\mathrm{res}}} = (1 - b(\alpha))/(\alpha\,\mathbb{E}B)$, Thm. 1 can alternatively be written as

$$\kappa(\alpha) = \left(1 - \frac{\lambda\,\mathbb{E}B}{r}\right)\sum_{n=0}^{\infty}\left(\frac{\lambda\,\mathbb{E}B}{r}\right)^n\left(b^{\mathrm{res}}(\alpha)\right)^n.$$

As a consequence, with $\varrho := \lambda\,\mathbb{E}B/r$,

$$\mathbb{P}(Q \leq x) = \mathbb{P}\left(\sum_{n=1}^{N} B_n^{\mathrm{res}} \leq x\right), \tag{3.1}$$

where $\mathbb{P}(N = n) = (1 - \varrho)\varrho^n$. This means that the steady-state workload Q can be interpreted as a geometric number of residuals of the job size B. ♢

Now focus on the spectrally positive case. Our goal is to find an expression for $\kappa(\alpha) = \mathbb{E}e^{-\alpha Q}$ for any $X \in \mathscr{S}_+$, by approximating $\varphi(\alpha)$ by a sequence $\varphi_n(\alpha)$ that correspond to compound Poisson processes, then apply Thm. 1 for these compound Poisson processes, and finally take $n \to \infty$.
In the spectrally-positive case we have, for a certain d, $\sigma^2 \geq 0$, and measure $\Pi_\varphi(\cdot)$ such that $\int_{(0,\infty)} \min\{1, x^2\}\Pi_\varphi(\mathrm{d}x) < \infty$, that the Laplace exponent reads

$$\varphi(\alpha) = \alpha d + \frac{1}{2}\alpha^2\sigma^2 + \int_{(0,\infty)} (e^{-\alpha x} - 1 + \alpha x\,1_{(0,1)}(x))\Pi_\varphi(\mathrm{d}x).$$

Now define, for a sequence ε_n such that $\varepsilon_n \to 0$ as $n \to \infty$,

$$\varphi_n(\alpha) := \left(d + \int_{\varepsilon_n}^1 x\Pi_\varphi(\mathrm{d}x) + \frac{\sigma^2}{\varepsilon_n}\right)\alpha + \frac{\sigma^2}{\varepsilon_n^2}\left(e^{-\alpha\varepsilon_n} - 1\right) + \int_{\varepsilon_n}^{\infty}(e^{-\alpha x} - 1)\Pi_\varphi(\mathrm{d}x).$$

It is easily verified that $\varphi_n(s) \to \varphi(s)$ as $n \to \infty$, whereas, for all $n \in \mathbb{N}$, $\varphi_n'(0) = \varphi'(0)$, but, more importantly, $\varphi_n(\alpha)$ is *the Laplace exponent of a compound Poisson process*. This is seen as follows. The drift term of this compound Poisson process is

$$d_n := d + \int_{\varepsilon_n}^1 x\Pi_\varphi(\mathrm{d}x) + \frac{\sigma^2}{\varepsilon_n} > 0.$$

Then, the term $\sigma^2/\varepsilon_n^2 \cdot (e^{-\alpha\varepsilon_n} - 1)$ can be interpreted as the contribution of a Poisson stream (arrival rate $\lambda_{1,n} := \sigma^2/\varepsilon_n^2$) of jobs of deterministic size $\beta_{1,n} := \varepsilon_n$. Finally,

$$\int_{\varepsilon_n}^\infty (e^{-\alpha x} - 1)\Pi_\varphi(\mathrm{d}x) = \Pi_\varphi([\varepsilon_n, \infty)) \int_{\varepsilon_n}^\infty (e^{-\alpha x} - 1) \frac{\Pi_\varphi(\mathrm{d}x)}{\Pi_\varphi([\varepsilon_n, \infty))},$$

which is the contribution of a Poisson stream (arrival rate $\lambda_{2,n} := \Pi_\varphi([\varepsilon_n, \infty))$) of jobs, whose sizes are i.i.d. samples from a 'truncated distribution' with density $\Pi_\varphi(\mathrm{d}x)/\Pi_\varphi([\varepsilon_n, \infty))$, for $x \geq \varepsilon_n$, and mean

$$\beta_{2,n} := \int_{\varepsilon_n}^\infty x \frac{\Pi_\varphi(\mathrm{d}x)}{\Pi_\varphi([\varepsilon_n, \infty))}.$$

Let Q_n be the steady state workload of the queue fed by a compound Poisson process with Laplace exponent $\varphi_n(\alpha)$. Due to $\varphi_n(\alpha) \to \varphi(\alpha)$ it is conceivable that $\mathbb{E}e^{-\alpha Q_n} \to \mathbb{E}e^{-\alpha Q}$. From Thm. 1,

$$\begin{aligned} \mathbb{E}e^{-\alpha Q_n} &= \alpha(d_n - \lambda_{1,n}\beta_{1,n} - \lambda_{2,n}\beta_{2,n}) \Big/ \left(d_n\alpha - \frac{\sigma^2}{\varepsilon_n^2}\left(1 - e^{-\alpha\varepsilon_n}\right) - \int_{\varepsilon_n}^\infty (1 - e^{-\alpha x})\Pi_\varphi(\mathrm{d}x) \right) \\ &\to \frac{\alpha\varphi'(0)}{\varphi(\alpha)} \quad \text{as } n \to \infty; \end{aligned}$$

the convergence follows from straightforward algebra. In other words, under the proviso that we can prove that $\mathbb{E}e^{-\alpha Q_n} \to \mathbb{E}e^{-\alpha Q}$, we have established the following result. Thm. 2 is often attributed to Zolotarev [98].

Theorem 2. *Let $X \in \mathscr{S}_+$. For $\alpha \geq 0$,*

$$\kappa(\alpha) = \mathbb{E}e^{-\alpha Q} = \frac{\alpha\varphi'(0)}{\varphi(\alpha)}.$$

The convergence $\mathbb{E}e^{-\alpha Q_n} \to \mathbb{E}e^{-\alpha Q}$ is a technical issue that lies beyond the scope of this survey. Thm. 2 provides us with the Laplace transform of the random variable under consideration, but it is noticed that there are powerful techniques to numerically invert these transforms. Besides the classical reference [3], we wish to draw attention on novel ideas developed by den Iseger, reported on in [42].

Alternative proofs of Thm. 2 rely on martingale techniques, most notably the celebrated *Kella-Whitt martingale* [62]; see also [67, Section 4.4] and [8, Section IX.3]. With

$$L_t(x) := \max\{0, L_t - x\} = \max\left\{0, -\inf_{0\leq s\leq t} X_s - x\right\},$$

it was shown using stochastic integration theory that, for $X \in \mathscr{S}_+$,

$$K_t := \varphi(\alpha)\int_0^t e^{-\alpha Q_s}\mathrm{d}s + e^{-\alpha x} - e^{-\alpha Q_t} - \alpha L_t(x)$$

is a martingale. Now consider this martingale, and assume that the queue is in stationarity at time 0. Then, stopping the martingale at time 1 results in the identity

$$0 = \mathbb{E}K_1 = \varphi(\alpha)\mathbb{E}e^{-\alpha Q} + \mathbb{E}e^{-\alpha Q} - \mathbb{E}e^{-\alpha Q} - \alpha\mathbb{E}L_1(Q),$$

so that

$$\mathbb{E}e^{-\alpha Q} = \frac{\alpha\,\mathbb{E}L_1(Q)}{\varphi(\alpha)}.$$

Now realizing that $\mathbb{E}e^{-\alpha Q} \to 1$ as $\alpha \downarrow 0$, we retrieve Thm. 2. In passing, we have shown that, in stationarity, the 'mean amount of local time per time unit' equals $\varphi'(0)$.

Example 1. Suppose $X \in \mathbb{B}\mathrm{m}(\mu, \sigma^2)$ for some $\mu < 0$. Then, with $\nu := -2\mu/\sigma^2 > 0$,

$$\mathbb{E}e^{-\alpha Q} = \frac{\alpha\varphi'(0)}{\varphi(\alpha)} = \frac{\nu}{\nu+\alpha}.$$

We conclude that the steady-state workload in a Brownian queue has an exponential distribution with mean $1/\nu$. Observe that the workload of this queue has no point mass at 0 (which could be expected, due to the non-differentiability of the sample paths). $\diamondsuit$

Example 2. Consider the case of $X \in \mathbb{S}(\alpha, 1, -r)$ with $\alpha \in (1, 2)$ and $r > 0$. Then, using that

$$\varphi(s) = rs + \frac{1}{\cos\left(\pi(\alpha/2-1)\right)}s^{\alpha},$$

one can invert the transform of Thm. 2 to obtain [48]

$$\mathbb{P}\left(Q > u\right) = \sum_{n=0}^{\infty} \frac{\left(-r\cos\left(\pi(\alpha/2-1)\right)\right)^n}{\Gamma(1+(\alpha-1)n)} u^{(\alpha-1)n}.$$

It is concluded that Q has a so-called Mittag-Leffler distribution. $\diamondsuit$

Thm. 2 reveals all moments of the steady-state queue Q, and in particular its mean and variance:

$$\mu := \mathbb{E}Q = -\frac{\mathrm{d}}{\mathrm{d}\alpha}\left.\frac{\alpha\varphi'(0)}{\varphi(\alpha)}\right|_{\alpha\downarrow 0} = \frac{\varphi''(0)}{2\varphi'(0)}, \tag{3.2}$$

and similarly

$$v := \mathbb{V}\mathrm{ar}Q = \frac{1}{4}\left(\frac{\varphi''(0)}{\varphi'(0)}\right)^2 - \frac{1}{3}\frac{\varphi'''(0)}{\varphi'(0)}. \tag{3.3}$$

3.2. **Spectrally-negative case.** For spectrally negative input, the reasoning is substantially simpler. First observe that $\mathbb{E}e^{\beta_0 X_t}$ is a martingale, with $\beta_0 := \Psi(0) > 0$. 'Optional sampling' [97, Ch. A14] thus gives, for any positive x,

$$\mathbb{P}(\exists t \geq 0 : X_t > x)e^{\beta_0 x} = 1,$$

using that, due to the fact that there are no jumps in the upward direction, given a certain level $x > 0$ is reached, it is attained with equality. As Q is distributed as the supremum over $t \geq 0$ of X_t ('Reich's identity', see Eqn. (2.2)), we obtain the following result.

Theorem 3. *Let $X \in \mathscr{S}_-$. Then Q is exponentially distributed with mean $1/\beta_0$.*

4. Transient workload

This section focuses on various transient metrics. In terms of Laplace transforms we subsequently address (i) the transient distribution, (ii) the busy period, (iii) the correlation of the workload, and (iv) the smallest value attained by the workload process in an interval of given length.

4.1. **Transform of transient distribution.** In this section our objective is to analyze (in terms of transforms) the distribution of Q_t, for some $t>0$, conditional on $Q_0=x$. As before, we distinguish between $X\in\mathscr{S}_+$ and $X\in\mathscr{S}_-$.

Spectrally positive case. Let $\zeta(x):=\mathbb{E}_x e^{-\alpha Q_T}$, where T is exponentially distributed with mean $1/\vartheta>0$, independently of the Lévy process under consideration. We first focus on $X\in\mathbb{CP}(r,\lambda,b(\cdot))$; for ease we normalize time such that $r=1$. Starting at zero, one should distinguish (i) between the clock T expires before the first jump of the compound Poisson process, and vice versa, and (ii) whether or not the buffer has become empty. One thus obtains

$$\begin{aligned}\zeta(x) &= \int_0^\infty\int_0^x \lambda e^{-(\lambda+\vartheta)y}\zeta(x-y+z)\mathrm{d}y\mathrm{d}\mathbb{P}(B\le z)+\frac{\vartheta}{\vartheta+\lambda-\alpha}(e^{-\alpha x}-e^{-(\lambda+\vartheta)x})\\ &\quad+\frac{\lambda}{\vartheta+\lambda}e^{-(\lambda+\vartheta)x}\int_0^\infty\zeta(z)\mathrm{d}\mathbb{P}(B\le z)+\frac{\vartheta}{\vartheta+\lambda}e^{-(\lambda+\vartheta)x}.\end{aligned}$$

It is a lengthy though elementary verification to check that $\zeta(x)=Ke^{-kx}+Le^{-\ell x}$ satisfies this equation, when picking $k=\alpha$, $\ell=\psi(\vartheta)$, $K=\vartheta/(\vartheta-\varphi(\alpha))$ and $L=-K\alpha/\psi(\vartheta)$. As before, we can then approximate any $X\in\mathscr{S}_+$ by a compound Poisson process, yielding the following result.

Theorem 4. *Let $X\in\mathscr{S}_+$, and let T be exponentially distributed with mean $1/\vartheta$, independently of X. For $\alpha\ge 0$, $x\ge 0$,*

$$\mathbb{E}_x e^{-\alpha Q_T}=\vartheta\int_0^\infty e^{-\vartheta t}\mathbb{E}_x e^{-\alpha Q_t}=\frac{\vartheta}{\vartheta-\varphi(\alpha)}\left(e^{-\alpha x}-\frac{\alpha}{\psi(\vartheta)}e^{-\psi(\vartheta)x}\right).$$

Thm. 4 can alternatively be derived by a level-crossing argument — see e.g. [23] —, or by application of the Kella-Whitt martingale — see e.g. [8, Thm. IX.3.10] and [60]. We now detail the latter approach. Let T be exponentially distributed with mean $1/\vartheta$; then we have

$$0=\mathbb{E}K_T=\varphi(\alpha)\int_0^\infty\int_0^t\vartheta e^{-\vartheta t}e^{-\alpha Q_s}\mathrm{d}s\mathrm{d}t-e^{-\alpha x}-\mathbb{E}_x e^{-\alpha Q_T}-\alpha\mathbb{E}L_T(x).$$

The first term of the right-hand side can alternatively be written as

$$\varphi(\alpha)\int_0^\infty\int_s^\infty\vartheta e^{-\vartheta t}e^{-\alpha Q_s}\mathrm{d}t\mathrm{d}s=\frac{\varphi(\alpha)}{\vartheta}\,\mathbb{E}_x e^{-\alpha Q_T}.$$

Now $\mathbb{E}_x e^{-\alpha Q_T}$ can be solved, and we obtain an expression in which the unknown term $\mathbb{E}L_T(x)$ appears in the numerator, and in which the denominator equals $\vartheta-\varphi(\alpha)$. Then use the fact that the root of the denominator (i.e., $\alpha=\psi(\vartheta)$) should be a root of the numerator as well (otherwise the transform equals ∞). This enables us to solve $\mathbb{E}L_T(x)$, and finally we obtain the result of Thm. 4.

The special case of $X\in\mathbb{B}\mathrm{m}(\mu,\sigma^2)$ can be solved explicitly. It turns out that [51, p. 49]

$$\mathbb{P}(Q_t\le y\,|\,Q_0=x)=1-\Phi_{\mathrm{N}}\left(\frac{-y+x+\mu t}{\sigma\sqrt{t}}\right)-e^{2\mu y/\sigma^2}\Phi_{\mathrm{N}}\left(\frac{-y-x-\mu t}{\sigma\sqrt{t}}\right),$$

with $\Phi_{\mathrm{N}}(\cdot)$ denoting the distribution function of a standard Normal random variable.

Spectrally negative case. Following the setup of [67, Chapter 8], we first introduce, for spectrally negative Lévy processes, families of functions $W^{(q)}(\cdot)$ and $Z^{(q)}(\cdot)$ as follows. Let $W^{(q)}(x)$ be a strictly increasing and continuous function whose Laplace transform satisfies

$$\int_0^\infty e^{-\beta x}W^{(q)}(x)\mathrm{d}x=\frac{1}{\Phi(\beta)-q},\qquad \beta>\Psi(q). \tag{4.1}$$

In addition,

$$Z^{(q)}(x) := 1 + q\int_0^x W^{(q)}(y)\mathrm{d}y. \tag{4.2}$$

$W^{(q)}(\cdot)$ and $Z^{(q)}(\cdot)$ are usually referred to as the *q-scale functions*. Then [80, Eqn. (19)] gives, with some abuse of notation, the following transform (with respect to q) of the density of Q_t, given that $Q_0 = x$:

$$\int_0^\infty e^{-qt}\mathbb{P}_x(Q_t = y)\mathrm{d}t = e^{-\Psi(q)y}\frac{\Psi(q)}{q}Z^{(q)}(x) - W^{(q)}(x-y).$$

It is now a matter of straightforward calculus to show that the previous display leads to, with T denoting an exponential random variable with mean q^{-1},

$$\int_0^\infty e^{-\beta x}\mathbb{E}_x e^{-\alpha Q_T}\mathrm{d}x = I_1(\alpha,\beta,q) - I_2(\alpha,\beta,q);$$

where the integrals $I_1(\alpha,\beta,q)$ and $I_2(\alpha,\beta,q)$ are given by

$$\begin{aligned} I_1(\alpha,\beta,q) &:= \int_0^\infty\int_0^\infty qe^{-\beta x}e^{-\alpha y}e^{-\Psi(q)y}\frac{\Psi(q)}{q}Z^{(q)}(x)\mathrm{d}x\mathrm{d}y,\\ I_2(\alpha,\beta,q) &:= \int_0^\infty\int_0^\infty qe^{-\beta x}e^{-\alpha y}W^{(q)}(x-y)\mathrm{d}x\mathrm{d}y. \end{aligned}$$

We now compute $I_1(\alpha,\beta,q)$ and $I_2(\alpha,\beta,q)$ explicitly. Using (4.1) and (4.2), we obtain

$$\begin{aligned} I_1(\alpha,\beta,q) &= \frac{\Psi(q)}{\Psi(q)+\alpha}\int_0^\infty e^{-\beta x}Z^{(q)}(x)\mathrm{d}x\\ &= \frac{\Psi(q)}{\Psi(q)+\alpha}\left(\frac{1}{\beta}+\int_0^\infty\int_y^\infty qW^{(q)}(y)e^{-\beta x}\mathrm{d}x\mathrm{d}y\right) = \frac{\Psi(q)}{\Psi(q)+\alpha}\frac{1}{\beta}\left(1+\frac{q}{\Phi(\beta)-q}\right). \end{aligned}$$

Likewise,

$$I_2(\alpha,\beta,q) = \int_0^\infty qe^{-(\alpha+\beta)y}\frac{1}{\Phi(\beta)-q}\mathrm{d}y = \frac{q}{\alpha+\beta}\frac{1}{\Phi(\beta)-q}.$$

This leads to the following result.

Theorem 5. *Let $X\in\mathscr{S}_-$, and let T be exponentially distributed with mean $1/q$, independently of X. For $\alpha ge 0$ and $\beta > 0$,*

$$\int_0^\infty e^{-\beta x}\mathbb{E}_x e^{-\alpha Q_T}\mathrm{d}x = \frac{1}{\beta}\left(\frac{\Psi(q)}{\Psi(q)+\alpha}+\frac{q}{\Phi(\beta)-q}\frac{\Psi(q)-\beta}{\Psi(q)+\alpha}\frac{\alpha}{\alpha+\beta}\right).$$

4.2. **Busy period.** In this section we address the busy period in a Lévy-driven queue. We let τ denote the busy-period duration, starting from steady-state at time 0: $\tau := \inf\{t\geq 0 : Q_t = 0\}$, where Q_0 is distributed according to the stationary distribution. Throughout this section we write $p(t) := \mathbb{P}(\tau > t)$; we derive the Laplace transform of $p(\cdot)$.

Spectrally positive case. Let us start by considering the spectrally-positive case. We have, with $\tau(x) := \inf\{t\geq 0 : X_t = -x\}$, due to Lemma 1,

$$\begin{aligned} \int_0^\infty e^{-\vartheta t}p(t)\mathrm{d}t &= \int_0^\infty\left(\int_0^\infty e^{-\vartheta t}\mathbb{P}(\tau(x)>t)\mathrm{d}t\right)\mathrm{d}\mathbb{P}(Q_0<x)\\ &= \frac{1}{\vartheta}\int_0^\infty\left(1-e^{-\psi(\vartheta)x}\right)\mathrm{d}\mathbb{P}(Q_0<x) = \frac{1}{\vartheta}\left(1-\kappa(\psi(\vartheta))\right), \end{aligned}$$

where we recall that $\kappa(\alpha)$ was defined as $\mathbb{E}e^{\alpha Q}$, with Q the stationary workload. Application of Thm. 2 now leads to the following result.

Proposition 2. *Let $X \in \mathscr{S}_+$. For $\vartheta \geq 0$,*

$$\int_0^\infty e^{-\vartheta t} p(t)\mathrm{d}t = \frac{1}{\vartheta} - \varphi'(0)\frac{\psi(\vartheta)}{\vartheta^2}.$$

In the special case that $X \in \mathbb{C}\mathrm{P}(r, \lambda, b(\cdot))$, the notion of a busy period starting in 0 is well-defined. We denote this random variable by τ^0; let $\pi(\vartheta) := \mathbb{E}e^{-\vartheta\tau^0}$ be the corresponding Laplace transform, which is known to satisfy the fixed-point equation $\pi(\vartheta) = b(\vartheta + \lambda - \lambda\pi(\vartheta))$, after having renormalized time such that $r = 1$. Recall that the Laplace exponent is $\varphi(\alpha) = \alpha - \lambda + \lambda b(\alpha)$. Therefore

$$0 = b(\vartheta + \lambda - \lambda\pi(\vartheta)) - \pi(\vartheta) = \frac{1}{\lambda}\varphi(\vartheta + \lambda - \lambda\pi(\vartheta)) - \frac{\vartheta}{\lambda},$$

and hence $\varphi(\vartheta + \lambda - \lambda\pi(\vartheta)) = \vartheta$. Apply $\psi(\cdot)$ to both sides, and we obtain the following result.

Proposition 3. *Let $X \in \mathbb{C}\mathrm{P}(1, \lambda, b(\cdot))$. For $\vartheta \geq 0$,*

$$\pi(\vartheta) = \frac{\lambda + \vartheta}{\lambda} - \frac{1}{\lambda}\psi(\vartheta).$$

In fact, more refined results can be found [74]. Consider for instance

$$L(\vartheta; \alpha, \bar{\alpha}) := \int_0^\infty e^{-\vartheta t}\mathbb{E}\left[e^{-\alpha Q_0 - \bar{\alpha} Q_t}; \tau > t\right]\mathrm{d}t = \int_0^\infty e^{-(\alpha+\bar{\alpha})x}\mathbb{E}\left[\int_0^{\tau(x)} e^{-\bar{\alpha}X_t - \vartheta t}\mathrm{d}t\right]\mathrm{d}\mathbb{P}(Q_0 \leq x).$$

Now observe that for any Lévy process $(Z_t)_t$ and δ for which the expressions are well-defined, we have that

$$M_s := e^{-\delta Z_s} - 1 - \left[\log \mathbb{E}e^{-\delta Z_1}\right] \cdot \int_0^s e^{-\delta Z_t}\mathrm{d}t$$

is a martingale. Now pick $\delta = \bar{\alpha}$ and $Z_t = X_t + (\vartheta/\bar{\alpha})t$, and use 'optional sampling' to obtain

$$\mathbb{E}\left[\int_0^{\tau(x)} e^{-\bar{\alpha}X_t - \vartheta t}\mathrm{d}t\right] = \frac{1 - e^{(\bar{\alpha} - \psi(\vartheta))x}}{\vartheta - \varphi(\bar{\alpha})};$$

here it is used that $X_{\tau(x)} = -x$. Combining the above, we end up with

$$L(\vartheta; \alpha, \bar{\alpha}) = \frac{\varphi'(0)}{\vartheta + \varphi(\alpha)}\left(\frac{\alpha + \bar{\alpha}}{\varphi(\alpha + \bar{\alpha})} - \frac{\alpha + \psi(\vartheta)}{\varphi(\alpha + \psi(\vartheta))}\right).$$

It is actually also possible to compute the *joint* transform of the stationary workload and busy period:

$$\mathbb{E}e^{-\alpha Q - \theta\tau} = \int_0^\infty e^{-\alpha x}\mathbb{E}e^{-\vartheta\tau(x)}\mathrm{d}\mathbb{P}(Q \leq x) = \int_0^\infty e^{-\alpha x}e^{-\psi(\vartheta)(x)}\mathrm{d}\mathbb{P}(Q \leq x) = \kappa(\alpha + \psi(\vartheta)),$$

with $\kappa(\cdot)$ as given in Thm. 2.

Spectrally negative case. The spectrally-negative case can be dealt with similarly. First recall that

$$\int_0^\infty e^{-qt}\mathbb{P}(\tau > t)\mathrm{d}t = \frac{1}{q}\left(1 - \mathbb{E}e^{-q\tau}\right).$$

Then Lemma 2, in conjunction with Thm. 3, yields

$$\mathbb{E}e^{-q\tau} = \int_0^\infty \beta_0 e^{-\beta_0 x}\mathbb{E}e^{-q\tau(x)}\mathrm{d}x = 1 - \frac{q}{\Psi(q)}\frac{\Psi(q) - \beta_0}{q - \Phi(\beta_0)}.$$

Using that $\Phi(\beta_0) = 0$, we find that $\mathbb{E}e^{-q\tau} = \Psi(0)/\Psi(q)$, and in addition the following result is obtained.

Proposition 4. *Let $X \in \mathscr{S}_-$. For $q \geq 0$,*

$$\int_0^\infty e^{-qt} p(t)\mathrm{d}t = \frac{1}{q}\left(1 - \frac{\Psi(0)}{\Psi(q)}\right).$$

Similarly to what we did above for $X \in \mathscr{S}_+$, we have for $X \in \mathscr{S}_-$

$$L(q;\beta,\bar\beta) := \int_0^\infty e^{-qt}\mathbb{E}\left[e^{-\beta Q_0 - \bar\beta Q_t}; \tau > t\right]\mathrm{d}t = \int_0^\infty \beta_0 e^{-(\beta_0+\beta+\bar\beta)x}\mathbb{E}\left[\int_0^{\tau(x)} e^{-\bar\beta X_t - qt}\mathrm{d}t\right]\mathrm{d}x,$$

using that Q_0 is exponentially distributed with mean $1/\beta_0$. As in the spectrally-positive case,

$$\mathbb{E}\left[\int_0^{\tau(x)} e^{-\bar\beta X_t - qt}\mathrm{d}t\right] = \frac{1 - e^{\bar\beta X_{\tau(x)} - q\tau(x)}}{q - \Phi(-\bar\beta)}.$$

Applying the so-called second factorization identity — see e.g. [67] — we obtain that

$$L(q;\beta,\bar\beta) = \frac{\beta_0}{\beta_0+\beta+\bar\beta}\frac{\Psi(q)-\beta_0-\beta}{\Psi(q)+\bar\beta}\frac{1}{q-\Phi(\beta+\beta_0)}.$$

The joint transform of the stationary workload and busy period follows directly from Lemma 2 and Thm. 3:

$$\mathbb{E}e^{-\beta Q - q\tau} = \beta_0\int_0^\infty e^{-(\beta+\beta_0)x}\mathbb{E}e^{-\vartheta\tau(x)}\mathrm{d}x = \frac{\beta_0}{\beta+\beta_0}\left(1 - \frac{q}{\Psi(q)}\frac{\Psi(q)-\beta-\beta_0}{q-\Phi(\beta+\beta_0)}\right).$$

4.3. **Correlation function.** Thms. 4 and 5 enable us to find explicitly the Laplace transform $\hat r(\cdot)$ corresponding to the correlation of the workload process:

$$r(t) := \mathbb{C}\mathrm{orr}(Q_0, Q_t) = \frac{\mathbb{C}\mathrm{ov}(Q_0,Q_t)}{\sqrt{\mathbb{V}\mathrm{ar}Q_0\cdot\mathbb{V}\mathrm{ar}Q_t}} = \frac{\mathbb{E}(Q_0Q_t) - (\mathbb{E}Q_0)^2}{\mathbb{V}\mathrm{ar}Q_0},$$

as we show now. Here it is assumed that the system is in steady-state at time 0.

Spectrally positive case. In this case, Q_0 obeys the distribution featured in Thm. 2. Let T have an exponential distribution with mean $1/\vartheta$. First realize that

$$\mathbb{E}(e^{-\alpha Q_T} \mid Q_0 = q) = \int_0^\infty \vartheta e^{-\vartheta t}\mathbb{E}(e^{-\alpha Q_t} \mid Q_0 = q)\mathrm{d}t.$$

By differentiation with respect to α and subsequently letting $\alpha \downarrow 0$, we obtain

$$\int_0^\infty \vartheta e^{-\vartheta t}\mathbb{E}(Q_t \mid Q_0 = q)\mathrm{d}t = -\frac{\varphi'(0)}{\vartheta} + q + \frac{e^{-\psi(\vartheta)q}}{\psi(\vartheta)}. \tag{4.3}$$

Concentrate on the Laplace transform $\gamma(\vartheta)$ of $\mathbb{C}\mathrm{ov}(Q_0,Q_t)$. Straightforward calculus reveals that

$$\begin{aligned}\gamma(\vartheta) &:= \int_0^\infty \mathbb{C}\mathrm{ov}(Q_0,Q_t)e^{-\vartheta t}\mathrm{d}t = \int_0^\infty(\mathbb{E}(Q_0Q_t) - \mu^2)e^{-\vartheta t}\mathrm{d}t\\ &= \int_0^\infty\int_0^\infty q\cdot\mathbb{E}(Q_t \mid Q_0 = q)\cdot e^{-\vartheta t}\mathrm{d}\mathbb{P}(Q_0\leq q)\mathrm{d}t - \frac{\mu^2}{\vartheta};\end{aligned}$$

it is assumed that the queue is in stationarity at time 0 (and hence it is in stationarity at time t as well). By invoking (4.3) we find that the expression in the previous display equals

$$\begin{aligned}&\int_0^\infty \frac{q}{\vartheta}\left(-\frac{\varphi'(0)}{\vartheta} + q + \frac{e^{-\psi(\vartheta)q}}{\psi(\vartheta)}\right)\mathrm{d}\mathbb{P}(Q_0\leq q) - \frac{\mu^2}{\vartheta}\\ &= -\frac{\mu\varphi'(0)}{\vartheta^2} + \frac{v}{\vartheta} + \frac{1}{\vartheta\psi(\vartheta)}\mathbb{E}(Q_0e^{-\psi(\vartheta)Q_0}).\end{aligned} \tag{4.4}$$

From Thm. 2 we obtain by differentiating

$$\mathbb{E}(Q_0e^{-\alpha Q_0}) = \varphi'(0)\left(-\frac{1}{\varphi(\alpha)} + \alpha\frac{\varphi'(\alpha)}{(\varphi(\alpha))^2}\right).$$

Inserting this relation, in addition to (3.2), into (4.4) we obtain the Laplace transform of $\mathbb{C}\mathrm{ov}(Q_0, Q_t)$:

$$\gamma(\vartheta) = -\frac{\varphi''(0)}{2\vartheta^2} + \frac{v}{\vartheta} + \frac{\varphi'(0)}{\vartheta^2}\left(\frac{1}{\vartheta\psi'(\vartheta)} - \frac{1}{\psi(\vartheta)}\right).$$

This trivially also provides us with the Laplace transform of $\mathbb{C}\mathrm{orr}(Q_0, Q_t)$, as stated in the following theorem, which is due to [46] . When specializing to compound Poisson input, we retrieve [23, Eqn. (6.2)].

Theorem 6. *Let $X \in \mathscr{S}_+$. For $\vartheta \geq 0$, and v as in* (3.3)*,*

$$\hat{r}(\vartheta) := \int_0^\infty r(t)\, e^{-\vartheta t}\mathrm{d}t = \frac{\gamma(\vartheta)}{v} = \frac{1}{\vartheta} - \frac{\varphi''(0)}{2v\vartheta^2} + \frac{\varphi'(0)}{v\vartheta^2}\left[\frac{1}{\vartheta\psi'(\vartheta)} - \frac{1}{\psi(\vartheta)}\right]. \tag{4.5}$$

Remark 2. Using Thm. 2, it is readily verified that the result in Thm. 6 can be simplified to

$$\hat{r}(\vartheta) = \frac{1}{\vartheta} - \frac{1}{v}\left(\frac{\varphi''(0)}{2\vartheta^2} + \frac{\kappa'(\psi(\vartheta))}{\vartheta\psi(\vartheta)}\right),$$

with $\kappa(\alpha)$, as before, denoting $\mathbb{E}e^{-\alpha Q}$. ♢

Example 3. Consider the situation that $(X_t)_{t\geq 0}$ corresponds to standard Brownian motion decreased by a linear drift (say of rate 1, so $X \in \mathbb{B}\mathrm{m}(-1, 1)$). In other words: the Laplace exponent of the Lévy process is given by $\varphi(\alpha) = \alpha + \frac{1}{2}\alpha^2$, and its inverse is $\psi(\vartheta) = -1 + \sqrt{1 + 2\vartheta}$. Now consider the workload process $(Q_t)_{t\geq 0}$ and its correlation function. The above theory yields that the Laplace transform of $r(\cdot)$ is given by

$$\hat{r}(\vartheta) = \frac{1}{\vartheta} - \frac{2}{\vartheta^2} + \frac{2}{\vartheta^3}\left(\sqrt{1 + 2\vartheta} - 1\right).$$

It turns out to be possible to explicitly invert $\hat{r}(\cdot)$:

$$r(t) = 2(1 - 2t - t^2)\left(1 - \Phi_{\mathrm{N}}(\sqrt{t})\right) + 2\sqrt{t}(1 + t)\phi_{\mathrm{N}}(\sqrt{t}), \tag{4.6}$$

with $\Phi_{\mathrm{N}}(\cdot)$ (resp. $\phi_{\mathrm{N}}(\cdot)$) the standard Normal distribution (resp. density). Eqn. (4.6) is in agreement with the results in [1] and [73, Section 12.1]. ♢

Spectrally negative case. The analysis of the spectrally-negative case is similar; it is based on the facts (i) that we have the double transform of Q_t through Thm. 5, and (ii) that Q_0 is exponentially distributed (as we know from Thm. 3). We refer to [50].

Theorem 7. *Let $X \in \mathscr{S}_-$. For $q \geq 0$,*

$$\hat{r}(q) := \int_0^\infty r(t)\, e^{-qt}\mathrm{d}t = \frac{1}{q} + \frac{\beta_0^2}{q^2}\Phi'(\beta_0)\left(\frac{1}{\Psi(q)} - \frac{1}{\beta_0}\right).$$

Relying on the theory of completely monotone functions, in particular on the fact that completely monotone functions can be regarded as Laplace transforms of nonnegative random variables, various structural properties can be proven; [24]. The following result was established for compound Poisson input in [78], and generalized to $X \in \mathscr{S}_+$ in [46], whereas the case of $X \in \mathscr{S}_-$ was dealt with in [50].

Proposition 5. *Let $X \in \mathscr{S}_+$ or $X \in \mathscr{S}_-$. Then $r(\cdot)$ is positive, decreasing, and convex.*

In the proof for $X \in \mathscr{S}_+$, it plays a crucial role that $-\psi(\cdot)$ is the Laplace exponent of an *increasing* Lévy process, as follows from Lemma 1. This essentially means that this Lévy process does not have a Brownian component, and it entails that $\psi'(\cdot)$ is completely monotone. Using this property, it is then a matter of delicate manipulation with Laplace transforms to prove that the Laplace transform of $r''(t)$ (the second derivative of the correlation function) is completely monotone, and therefore $r(\cdot)$ is convex. Likewise, it is concluded that $r(\cdot)$ is decreasing and positive. Details are found in [46, Section 3].

The proof for $X \in \mathscr{S}_-$ works quite similarly. There it is a crucial to note that $\Psi(0)/\Psi(q)$ is completely monotone, which follows from the fact that $\mathbb{E}e^{-q\tau} = \Psi(0)/\Psi(q)$. The full proof is given in [50, Section 2].

4.4. **Infimum over given time interval.** In this subsection, which is based on [40], we consider the distribution of the random variable $M_t := \inf_{s\in[0,t]} Q_s$, assuming the workload is in stationarity at time 0. Again, we find explicit expressions for Laplace transforms for the spectrally one-sided situation. Observe that $M_t > u$ corresponds to $Q_0 + \inf_{s\in[0,t]} X_s > u$. Hence

$$\begin{aligned}
&\int_0^\infty e^{-\vartheta t}\int_0^\infty e^{-\alpha u}\mathbb{P}(M_t > u)\mathrm{d}u\mathrm{d}t \\
&= \int_0^\infty e^{-\vartheta t}\int_0^\infty e^{-\alpha u}\int_u^\infty \mathbb{P}\left(\inf_{s\in[0,t]} X_s > u - q\right)\mathrm{d}\mathbb{P}(Q_0 \le q)\mathrm{d}u\mathrm{d}t \\
&= \int_0^\infty \int_0^q e^{-\alpha u}\int_0^\infty e^{-\vartheta t}\mathbb{P}(\tau(q-u) > t)\mathrm{d}t\mathrm{d}u\mathrm{d}\mathbb{P}(Q_0 \le q).
\end{aligned}$$

The inner integral is the transform of the tail probability $\mathbb{P}(\tau(q-u) > t)$, so that integration by parts yields

$$\int_0^\infty \int_0^q e^{-\alpha u}\frac{1}{\vartheta}\left(1 - \mathbb{E}e^{-\vartheta\tau(q-u)}\right)\mathrm{d}u\mathrm{d}\mathbb{P}(Q_0 \le q). \tag{4.7}$$

Now we have to distinguish between $X \in \mathscr{S}_+$ and $\mathscr{S}_-$. In the former case we can use Lemma 1 to evaluate the inner integral; then we have to perform a bit of straightforward calculus, and we have to apply Thm. 2. We obtain the following result.

Proposition 6. *Let $X \in \mathscr{S}_+$. For $\alpha, \vartheta \ge 0$,*

$$\int_0^\infty e^{-\vartheta t}\int_0^\infty e^{-\alpha u}\mathbb{P}(M_t > u)\mathrm{d}u\mathrm{d}t = \frac{1}{\vartheta}\left(\frac{1}{\alpha} - \frac{\varphi'(0)}{\varphi(\alpha)}\right) - \frac{\varphi'(0)}{(\alpha - \psi(\vartheta))\vartheta}\left(\frac{\psi(\vartheta)}{\vartheta} - \frac{\alpha}{\varphi(\alpha)}\right).$$

In the latter case, i.e., $X \in \mathscr{S}_-$, we recall from Thm. 3 that Q_0 has an exponential distribution with parameter β_0. Interchanging the order of integration in (4.7), and applying the second factorization identity [67], we obtain the following result.

Proposition 7. *Let $X \in \mathscr{S}_-$. For $\beta, q \ge 0$,*

$$\int_0^\infty e^{-qt}\int_0^\infty e^{-\beta u}\mathbb{P}(M_t > u)\mathrm{d}u\mathrm{d}t = \frac{1}{\beta + \beta_0}\frac{\Psi(q)}{\Psi(q) + \beta_0}.$$

5. Workload asymptotics

The goal of this section is to characterize $\mathbb{P}(Q > u)$ for u large. We distinguish between three classes of Lévy processes: those with a light upper tail, an intermediate upper tail, and a heavy upper tail. It should be born in mind that $\mathbb{P}(Q > u)$ equals $\mathbb{P}(\exists t \ge 0 : X_t > u) = \mathbb{P}(\sigma(u) < \infty)$, where $\sigma(u)$ is defined as the hitting time of level u, i.e., $\inf\{t : X_t \ge u\}$.

5.1. **Light-tailed regime.** We define by $\mathscr{L}$ the class of Lévy processes such that there is an $\omega > 0$ such that $\mathbb{E}e^{\omega X_1} = 1$ and $\mathbb{E}X_1 e^{\omega X_1} < \infty$. For ease we start by focusing on $X \in \mathbb{CP}(1, \lambda, b(\cdot))$, and consider the more general case of $X \in \mathscr{L}$ later.

Assume $X \in \mathbb{CP}(1, \lambda, b(\cdot))$, with $\rho := \lambda\,\mathbb{E}B < 1$. Then ω solves $\lambda + \omega = \lambda b(-\omega)$. Referring to the original probability measure by $\mathbb{P}$, we introduce an alternative measure $\mathbb{Q}$ that is characterized as $\mathbb{CP}(1, \lambda + \omega, \bar{b}(\cdot))$, where $\bar{b}(\alpha) := b(\alpha - \omega)/b(-\omega)$. From the definition of ω and the convexity of $\mathbb{E}e^{\omega X_1}$, it can be concluded (in self-evident notation) that

$$(\lambda + \omega)\,\mathbb{E}_{\mathbb{Q}}B = (\lambda + \omega)\left(-\frac{b'(-\omega)}{b(-\omega)}\right) = -\lambda b'(-\omega) =: \rho_{\mathbb{Q}} > 1, \tag{5.1}$$

so that under $\mathbb{Q}$ the queue is *unstable*. In other words: under $\mathbb{Q}$ we have that $\sigma(u) < \infty$ almost surely, for any $u > 0$. Using this fact, in conjunction with $\mathbb{E}e^{\omega X_t} = 1$ for all $t \geq 0$, it is a standard that

$$\mathbb{P}(Q > u) = \mathbb{E}_{\mathbb{Q}}e^{-\omega X_{\sigma(u)}}. \tag{5.2}$$

Now realize that $X_{\sigma(u)} = u + R_u$, where R_u is the *overshoot* over level u. Let L_n be the n-th *ladder height*, i.e., the difference between the n-th and $(n-1)$-st record; these random variables are positive and i.i.d., and, due to (5.1), nondefective. Renewal theory now yields that R_u converges to a limiting random variable R, where

$$\mathbb{Q}(R \leq v) = \frac{1}{\mathbb{E}_{\mathbb{Q}}L}\int_0^v (1 - \mathbb{Q}(L \leq y))\mathrm{d}y,$$

with L denoting a ladder height. Due to the definition of $\mathbb{Q}$, we have

$$\mathrm{d}\mathbb{Q}(L \leq y) = e^{\omega y}\mathrm{d}\mathbb{P}(L \leq y) = e^{\omega y}\lambda\mathbb{P}(B > y)\mathrm{d}y;$$

it follows from the definition of ω that this density integrates to 1. Combining the above, we obtain that, as $u \to \infty$,

$$\mathbb{P}(Q > u)e^{\omega u} \to \frac{1}{\mathbb{E}_{\mathbb{Q}}L}\int_0^\infty e^{-\omega y}(1 - \mathbb{Q}(L \leq y))\mathrm{d}y.$$

Straightforward calculus now yields the classical Cramér-Lundburg asymptotics.

Theorem 8. *Let $X \in \mathbb{CP}(1, \lambda, b(\cdot)) \cap \mathscr{L}$. Then, as $u \to \infty$,*

$$\mathbb{P}(Q > u)e^{\omega u} \to \frac{1 - \rho}{\rho_{\mathbb{Q}} - 1}.$$

In passing, we also proved that, for all $u \geq 0$, $\mathbb{P}(Q > u) \leq e^{-\omega u}$ (realize that $R_u \geq 0$). In [39] it is argued that this uniform bound applies for all $X \in \mathscr{L}$, i.e., not just for compound Poisson; the proof relies on a change-of-measure argument.

Corollary 1. *Let $X \in \mathscr{L}$. Then $\mathbb{P}(Q > u) \leq e^{-\omega u}$.*

A next step is to consider asymptotics for more general Lévy processes in $\mathscr{L}$: is it for instance possible to extend Thm. 8 to $\mathscr{S}_+$? To this end, realize that we have the Laplace transform of Q, viz. $\alpha\varphi'(0)/\varphi(\alpha)$. Then we would like to use this transform to obtain the tail asymptotics. One such an approach is through application of the so-called Heaviside principle, as advocated in e.g. [3]. To apply this, first note that

$$\int_0^\infty e^{-\alpha x}\mathbb{P}(Q > x)\mathrm{d}x = \frac{1}{\alpha} - \frac{\varphi'(0)}{\varphi(\alpha)}.$$

Now observe that when $X \in \mathscr{L}$, $\varphi(\cdot)$ has a pole in $-\omega$, and

$$\lim_{\alpha \downarrow -\omega} \int_0^\infty e^{-\alpha x} \mathbb{P}(Q > x) \mathrm{d}x = \frac{\varphi'(0)}{-\varphi'(-\omega)} > 0;$$

note that we assumed that the denominator of the last expression is finite (due to $\mathbb{E} X_1 e^{\omega X_1} < \infty$). Now the Heaviside principle yields that, as $u \to \infty$,

$$\mathbb{P}(Q > u) e^{\omega u} \to \frac{\varphi'(0)}{-\varphi'(-\omega)};$$

it is readily checked that for the compound Poisson case this expression agrees with that of Thm. 8. It is noted, however, that the Heaviside principle, although well established in the literature and frequently used [35], lacks full mathematical rigor.
The most general result is due to Bertoin and Doney [25]: there tail asymptotics for $\mathbb{P}(Q > u)$ are derived for the full class $\mathscr{L}$. These are of the form $Ce^{-\omega u}$, where ω solves $\mathbb{E}e^{\omega X_1} = 1$, but with some rather involved expression for C. A nice alternative proof of this result, relying on an embedding approach, was given in [43].

5.2. **Intermediate regime.** Define

$$\omega := \sup\{\delta \geq 0 : \mathbb{E}e^{\delta X_1} < \infty\}.$$

We say that $X \in \mathscr{I}$ if $\omega \in (0, \infty)$ and $\mathbb{E}e^{\omega X_1} < 1$; this basically means that at $\delta = \omega$, the moment generating function $\mathbb{E}e^{\delta X_1}$ jumps from a value strictly smaller than 1 to ∞.
Interestingly, again the change-of-measure technique can be used to find a uniform upper bound. Defining $M(\delta) := \mathbb{E}e^{\delta X_1}$, we could identify with $\mathbb{Q}(\vartheta)$ the Lévy process that obeys

$$\mathbb{E}_{\mathbb{Q}(\vartheta)} e^{\delta X_1} = \frac{M(\delta + \vartheta)}{M(\vartheta)}.$$

As before, we obtain the inequality, for all $\vartheta < \omega$,

$$\mathbb{P}(Q > u) = \mathbb{E}_{\mathbb{Q}(u)} \left(e^{-\vartheta X_{\sigma(u)}} \cdot (M(\vartheta))^{\sigma(u)} \right) \leq e^{-\vartheta u}.$$

We obtain the following bound.

Corollary 2. *Let $X \in \mathscr{I}$. Then $\mathbb{P}(Q > u) \leq e^{-\omega u}$.*

The following exact asymptotics were derived in [43]; see also [65]. Interestingly, they show that for $X \in \mathscr{I}$ the tail distribution of Q is asymptotically proportional to that of X_1.

Proposition 8. *Let $X \in \mathscr{I}$. Then, as $u \to \infty$,*

$$\frac{\mathbb{P}(Q > u)}{\mathbb{P}(X_1 > u)} \to \frac{\mathbb{E}e^{\omega Q}}{M(\omega) \log M(\omega)}.$$

5.3. **Heavy-tailed regime.** In this section we consider Lévy processes for which $\mathbb{E}e^{\delta X_1} = \infty$ for all $\delta > 0$. An important subclass of these are the regularly varying Lévy processes $\mathscr{R}$. Considering the class of compound Poisson inputs, regular variation refers to the tail of the distribution of the jobs: it is assumed that for an index α and all $y > 0$, as $x \to \infty$,

$$\frac{\mathbb{P}(B > yx)}{\mathbb{P}(B > x)} \to y^{\alpha}.$$

We here sketch how one can find the tail asymptotics $\mathbb{P}(Q > u)$ for u large, following a recipe proposed in [99, pp. 36-39]. This recipe is based on the insight that in these heavy-tailed scenarios a large workload is (with overwhelming probability) due to a single big job. The approach therefore consists of a lower bound, in which the probability of this most likely scenario is evaluated, and an upper bound in which it is shown that the contributions of other scenarios (e.g. no big job, multiple big jobs) can be neglected. We here demonstrate how the lower bound is derived.

We consider $\mathbb{CP}(r,\lambda,b(\cdot))$, and we denote, as earlier, $\varrho := \lambda\,\mathbb{E}B$. First it is noted that due the the law of large numbers, we can find (for any $\delta,\varepsilon>0$) a $t_{\delta,\varepsilon}$ such that for all $t\geq t_{\delta,\varepsilon}$,

$$\mathbb{P}(X_t > (\varrho-\varepsilon)t) > 1-\delta.$$

It is noted that a sufficient condition for Q_0 exceeding u is that a job of size at least $u+(r-\varrho)t+\varepsilon t$ arrived at time $-t$, and that between $-t$ and 0 at least $(\varrho-\varepsilon)t$ arrived; notice that the former event is rare, as opposed to the latter. We obtain

$$\begin{aligned}\mathbb{P}(Q>u) &\geq \int_{t_{\delta,\varepsilon}}^{\infty}\lambda\mathbb{P}(B>u+(r-\varrho)t+\varepsilon t)\mathbb{P}(-X_{-t}>(\varrho-\varepsilon)t)\mathrm{d}t\\ &\geq (1-\delta)\int_{t_{\delta,\varepsilon}}^{\infty}\lambda\mathbb{P}(B>u+(1-\varrho)t+\varepsilon t)\mathrm{d}t\\ &= (1-\delta)\frac{\varrho}{r-\varrho+\varepsilon}\mathbb{P}(B^{\mathrm{res}}>u+t_{\delta,\varepsilon})\sim\frac{(1-\delta)\varrho}{r-\varrho+\varepsilon}\mathbb{P}(B^{\mathrm{res}}>u),\end{aligned}$$

where the last step is due to the definition of regular variation. Now let $\delta,\varepsilon\downarrow 0$. After having established the corresponding upper bound, the following theorem is obtained.

Theorem 9. *Let $X\in\mathbb{CP}(r,\lambda,b(\cdot))\cap\mathscr{R}$. Then, as $u\to\infty$,*

$$\mathbb{P}(Q>u)\sim\frac{\varrho}{r-\varrho}\mathbb{P}(B^{\mathrm{res}}>u).$$

There is an alternative approach though, that is helpful if the Laplace transform is available: Tauberian inversion. To this end, we first define the following notion.

Defintion 1. *We say that $f(x)\in\mathscr{R}_\delta(n,\eta)$, with $\delta\in(n,n+1)$, for $x\downarrow 0$, if*

$$f(x)=\sum_{i=0}^{n}\frac{f^{(i)}(0)}{i!}x^i+\eta x^\delta L(1/x),\quad x\downarrow 0,$$

for a slowly varying function $L(\cdot)$, i.e., $L(x)/L(tx)\to 1$ for $x\to\infty$, for any t.

Suppose now that $\varphi(\alpha)\in\mathscr{R}_\nu(n,\eta)$, it is readily checked that

$$\mathbb{E}e^{-\alpha Q}=\frac{\alpha\varphi'(0)}{\varphi(\alpha)}\in\mathscr{R}_{\nu-1}\left(n-1,\frac{\eta}{\varphi'(0)}\right).$$

The Tauberian theorem in Bingham, Goldie, and Teugels [28, Thm. 8.1.6] now yields the following result; see also [27].

Theorem 10. *Let $X\in\mathscr{S}_+\cap\mathscr{R}$, with $\varphi(\alpha)\in\mathscr{R}_\nu(n,\eta)$. Then, as $u\to\infty$,*

$$\mathbb{P}(Q>u)\sim\frac{(-1)^n}{\Gamma(2-\nu)}\cdot\left(\frac{\eta}{\varphi'(0)}\right)u^{1-\nu}L(u).$$

Example 4. Consider $X\in\mathbb{CP}(1,\lambda,b(\cdot))$. Suppose $\mathbb{P}(B>x)\sim x^{-\delta}L(x)$. From $\varphi(\alpha)=\alpha+\lambda b(\alpha)-\lambda$, it follows that $\varphi(\alpha)\in\mathscr{R}_\delta(n,\lambda\Gamma(1-\delta)(-1)^n)$ by applying 'Tauber'. Then the above theorem confirms the result of Thm. 9. ♢

Now define the class of heavy-tailed (or: *subexponential*) Lévy processes, as follows. To this end, we first introduce the notion of subexponential distribution functions, following the terminology of [43]. With $D(\cdot)$ being a distribution function on $[0,\infty)$ and $D^{\star 2}$ the convolution of D with itself, we say that D is subexponential if $1-D^{\star 2}(x)\sim 2(1-D(x))$ as $x\to\infty$. For a measure $\mu(\cdot)$ we say that it is subexponential if (i) $\mu([1,\infty))<\infty$, and (ii) $\mu([1,\cdot])/\mu([1,\infty))$ is subexponential. Then, for the spectral measure $\Pi(\cdot)$ of $(X_t)_t$, define

$$\Pi_I((x,\infty)) := \int_x^\infty \Pi((y,\infty))\mathrm{d}y.$$

We say that $X\in\mathscr{H}$ if $\Pi_I(\cdot)$ is a subexponential. The following result is found in [9]; a version with also local asymptotics was first presented in [43].

Theorem 11. *Let $X\in\mathscr{H}$. Then, as $u\to\infty$,*

$$\mathbb{P}(Q>u)\sim\frac{1}{-\mathbb{E}X_1}\int_u^\infty \mathbb{P}(X_1>x)\mathrm{d}x.$$

The class of α-stable Lévy motions belongs to $\mathscr{H}$. The following result [82] is an immediate consequence of Thm. 11, Prop. 1 and Karamata's theorem [28, Section 1.6]; recall that $m<0$.

Proposition 9. *Let $X\in\mathbb{S}(\alpha,\beta,m)$, with $\alpha\in(1,2)$ and $\beta\in(-1,1]$. Then, as $u\to\infty$,*

$$\mathbb{P}(Q>u)\sim\frac{1}{(-m)}\int_u^\infty x^{-\alpha}C_{\alpha,1}\left(\frac{1+\beta}{2}\right)\mathrm{d}x\sim\frac{1}{(-m)}\frac{1}{\alpha-1}u^{-\alpha+1}C_{\alpha,1}\left(\frac{1+\beta}{2}\right).$$

It is noted that there is a seeming incompatibility between the above symptotics and the corresponding result in [82] (which is Prop. 3.7 in [48]), but it is a matter of (straightforward but tedious) calculus to verify that both expressions match.

In [48] also the case is considered of Lévy input that is an aggregate of α-stable Lévy motion and compound Poisson with regularly varying jobs; it then turns out that the heaviest tail essentially dominates the asymptotics.

6. Transient asymptotics, rare-event simulation

In this section we first discuss the asymptotics of the transient metrics defined earlier, viz. the tail distribution of the busy period $\mathbb{P}(\tau>t)$, and the workload correlation function $r(t)$. Then we study, for various shapes of the function $T(u)$, probabilities of the type $\mathbb{P}(Q_0>pu, Q_{T(u)}>qu)$, for $p,q>0$. We finally consider the option of estimating small probabilities (and small correlations) through sophisticated simulation techniques.

6.1. **Asymptotics of transient metrics.** Let us first consider the light-tailed case. For $X\in\mathscr{L}$ we have that $\mathbb{E}e^{-sX_1}=1$ has a negative root, say $\omega<0$, which implies that $\mathbb{E}e^{-sX_1}$ has a minimizer somewhere between ω and 0. Relying on Heaviside heuristics, we now study the tail of $\mathbb{P}(\tau>t)$ and $r(t)$. We show the computations for $\mathbb{P}(\tau>t)$; those for $r(t)$ work similarly, and can be found in [46, 50].

We again start by considering the spectrally-positive case. As before, we assume that the equation $\varphi(\alpha)=0$ has a negative root. Observe that then Prop. 2 holds for any positive ϑ, but we can consider the analytic continuation up to the branching point $\vartheta^\star<0$ of $\psi(\cdot)$; let in the sequel $\zeta<0$ denote the minimizer of $\varphi(\cdot)$, so that $\varphi(\zeta)=\vartheta^\star<0$ (where it is noticed that $v_\varphi:=\varphi''(\zeta)>0$). Then the idea is to write, for $\vartheta \downarrow \vartheta^\star$, that $\psi(\vartheta) - \zeta \sim \sqrt{2/v_\varphi} \cdot \sqrt{\vartheta - \vartheta^\star}$. Hence, around $\vartheta^\star$, we have that, for some (irrelevant) constant κ,

$$\int_0^\infty e^{-\vartheta t} \mathbb{P}(\tau > t) \mathrm{d}t = \frac{1}{\vartheta} - \varphi'(0) \frac{\psi(\vartheta)}{\vartheta^2} \sim \kappa + A_\varphi \sqrt{\vartheta - \vartheta^\star}; \quad A_\varphi := -\frac{\varphi'(0)}{(\vartheta^\star)^2} \sqrt{\frac{2}{v_\varphi}} < 0,$$

and hence, applying 'Heaviside', we estimate the tail distribution of the busy period by

$$\mathbb{P}(\tau > t) \sim \frac{A_\varphi}{\Gamma(-\frac{1}{2})} \cdot \frac{e^{\vartheta^\star t}}{t\sqrt{t}}. \tag{6.1}$$

We now turn to the spectrally-negative case. Prop. 4 holds for any positive q, but we can consider the analytic continuation up to the branching point $q^\star < 0$ of $\Psi(\cdot)$. Let $\zeta > 0$ denote the minimizer of $\Phi(\cdot)$, so that $\Phi(\zeta) = q^\star < 0$. Similarly to the spectrally-negative case, we obtain, with $v_\Phi := \Phi''(\zeta) > 0$ and κ being some (irrelevant) number,

$$\int_0^\infty e^{-qt} \mathbb{P}(\tau > t) \mathrm{d}t = \frac{1}{q} \left(1 - \frac{\Psi(0)}{\Psi(q)} \right) \sim \kappa + A_\Phi \sqrt{q - q^\star}; \quad A_\Phi := \frac{\Psi(0)}{q^\star \zeta^2} \sqrt{\frac{2}{v_\varphi}} < 0,$$

and hence 'Heaviside' estimates the tail of the busy-period distribution by

$$\mathbb{P}(\tau > t) \sim \frac{A_\Phi}{\Gamma(-\frac{1}{2})} \cdot \frac{e^{q^\star t}}{t\sqrt{t}}. \tag{6.2}$$

For related results on the light-tailed case, we refer to e.g. [35, 79]. For the heavy-tailed case with compound Poisson input, we refer for results on the busy period to e.g. [14, 41], and for results on the correlation function to [46, Section 5].

6.2. **Asymptotics of joint transient distribution.** In [39] the focus is on probabilities of the type $\mathbb{P}(Q_0 > pu, Q_{T(u)} > qu)$, for $p, q > 0$. We summarize the main findings.

First conditions are identified under which the probability of interest is essentially dominated by the 'most demanding event', in the sense that it is asymptotically equivalent to $\mathbb{P}(Q > \max\{p, q\}u)$ for u large, where Q denotes the steady-state workload. These conditions turn out to reduce to $T(u)$ being sublinear (i.e., $T(u)/u \to 0$ as $u \to \infty$).

Then a second condition is derived under which the probability of interest 'decouples', in that it is asymptotically equivalent to $\mathbb{P}(Q > pu)\mathbb{P}(Q > qu)$ for u large (meaning that their ratio tends to 1). Here a crucial role is played by Q^D, for $D > \mathbb{E}X_1$, which is distributed as $\sup_{t \geq 0} X_t - Dt$; as a result Q^D resembles the original queue Q, but the drain rate is adapted by D, due to (2.2). Then the decoupling condition is that for all $\eta > 0$, $D > \mathbb{E}X_1$,

$$\lim_{u \to \infty} \frac{\mathbb{P}(Q^D > \eta T(u))}{\mathbb{P}(Q > pu)\mathbb{P}(Q > qu)} = 0.$$

For various types of input considered in the literature this 'decoupling condition' reduces to requiring that $T(u)$ is superlinear (i.e., $T(u)/u \to \infty$ as $u \to \infty$). This is for instance the case if the tails of Q and Q^D decay exponentially, as is verified easily. The decoupling condition does *not* hold, however, for $X \in \mathscr{R}$. It is seen that, due to the fact that the tails of Q and Q^D decay in a regularly varying fashion, the 'decoupling condition' reduces to $T(u)/u^2 \to \infty$. The rationale behind the fact that we only have decoupling for $T(u)$ increasing superquadratically, is that for $T(u)$ increasing subquadratically with overwhelming probability it suffices to have a *single* big jump to cause overflow over pu at time 0, and over qu at time $T(u)$; 'decoupling', on the contrary, would correspond to *two* big jumps. These findings imply that for $X \in \mathscr{R}$ there is a third regime, viz. $T(u)$ increasing superlinearly but subquadratically; [39] also identifies the asymptotics for this case.

In [39] special attention is paid to the case $T(u) = Ru$ for some $R > 0$; for $X \in \mathscr{L}$ intuitively appealing asymptotics are derived, intensively relying on sample-path large deviations results [4]. The regimes obtained can be interpreted in terms of most likely paths to overflow. If R is small (that is, fulfilling an explicit criterion in terms of $p, q,$ and the characteristics of the Lévy process $(X_t)_t$), then one has asymptotics of the type $\mathbb{P}(Q > \max\{p,q\}u)$. If this condition does not apply, two cases are possible: for large R the most likely scenario is that the buffer drains, remains empty for a while, and starts building up relatively short before R (in this case the asymptotics look like $\mathbb{P}(Q > pu)\mathbb{P}(Q > qu)$), whereas for moderate R the buffer remains (most likely) nonempty between 0 and R. We thus obtain that there are (uniquely characterized) $\bar{R}$ and $\check{R}$ such that for all R smaller than $\bar{R}$, with ω solving $\mathbb{E}e^{\omega X_1} = 1$, as $u \to \infty$,

$$\log\left(\mathbb{P}(Q_0 > pu, Q_{Ru} > qu)\right) u^{-1} \to -\max\{p,q\}\omega,$$

for R between $\bar{R}$ and $\check{R}$,

$$\log\left(\mathbb{P}(Q_0 > pu, Q_{Ru} > qu)\right) u^{-1} \to -p\omega - R \cdot \sup_{\delta}\left(\delta\left(\frac{q-p}{R}\right) - \log \mathbb{E}e^{\delta X_1}\right),$$

and for R larger than $\check{R}$,

$$\log\left(\mathbb{P}(Q_0 > pu, Q_{Ru} > qu)\right) u^{-1} \to -(p+q)\omega.$$

6.3. **Rare-event simulation, importance sampling.** This subsection focuses on efficient computation of small tail probabilities using simulation techniques (based on importance sampling). Special attention is paid to estimating the workload correlation function $r(t)$ for t large.

Estimation of workload asymptotics. Let us consider $X \in \mathscr{L}$. Then ideas that date back to [92] can be applied to estimate $\mathbb{P}(Q > u)$ efficiently. It is well-known, see e.g. [73, Section 8.2], that the number of simulation runs needed to obtain an estimate with a given predefined precision (expressed in terms of the ratio of the width of the confidence interval and the estimate), is inversely proportional to the probability to be estimated. In the situation at hand this means that this number grows roughly exponentially in u, and as a result simulation experiments may take prohibitively long. Our objective now is to speed up the simulation.

Let, as before, ω solves $\mathbb{E}e^{\omega X_1} = 1$. The idea is now to not perform the simulation under the original measure $\mathbb{P}$, corresponding to the characteristic triplet (d, σ^2, Π), but under an alternative measure $\mathbb{Q}$ under which the event of interest occurs more frequently. After weighing the simulation output with an appropriate likelihood ratio, unbiasedness is recovered. This procedure is commonly referred [13, pp. 127-128] to as *importance sampling*.

This $\mathbb{Q}$ is an *exponentially twisted* version of $\mathbb{P}$, that is, $\mathbb{Q}$ is such that, in self-evident notation,

$$\mathbb{E}_{\mathbb{Q}}e^{\delta X_1} = \mathbb{E}e^{(\delta+\omega)X_1}.$$

It is now elementary to check that $\mathbb{Q}$ also corresponds to a Lévy process, with triplet

$$\left(d + \sigma^2\omega + \int_{-1}^{1} x(e^{\omega x} - 1)\Pi(\mathrm{d}x), \sigma^2, e^{\omega x}\Pi(\mathrm{d}x)\right).$$

Observe the methodological similarity with the derivation of the Cramér-Lundberg asymptotics in Section 5.1.

Recall that the convexity of $\mathbb{E}e^{\delta X_1}$ implies that $\mathbb{E}_{\mathbb{Q}}X_1 = \mathbb{E}X_1 e^{\omega X_1} > 0$, so that the random variable $\sigma(u) := \inf\{t : X_t \geq u\}$ becomes nondefective under $\mathbb{Q}$. It thus follows that

$$\mathbb{P}(Q > u) = \mathbb{E}_{\mathbb{Q}} e^{-\omega X_{\sigma(u)}}.$$

In other words: we should simulate under $\mathbb{Q}$ until $\sigma(u)$, record the value x_i of $e^{-\omega X_{\sigma(u)}}$ in each run i, perform N runs, and estimate $\mathbb{P}(Q > u)$ by $N^{-1}\sum_{i=1}^N x_i$. It is easy to check that this estimator is unbiased. In addition, due to the fact that each observation of $e^{-\omega X_{\sigma(u)}}$ is bounded by $e^{-\omega u}$, the estimator has excellent variance properties (in particular it has bounded relative error, see [13, p. 159]). Clearly, a prerequisite to apply this method is that one should be able to sample trajectories of Lévy processes; the state-of-the-art on this issue is presented in [13, Ch. XII].

For the case of heavy tails, we refer to [12, 16] and [13, Section VI.3]. Importantly, the above ideas for $X \in \mathscr{L}$ do not carry over to the heavy-tailed case, basically because (most likely) not several 'somewhat unlikely' events cause overflow, but rather a single big jump.

Estimation of busy-period asymptotics. We now aim at efficiently estimating $\mathbb{P}(\tau > t)$ for $X \in \mathscr{L}$. In this case [50] proposed the following alternative measure; for ease we concentrate on $X \in \mathscr{S}_+$, but $X \in \mathscr{S}_-$ can be dealt with similarly.

- Let, in the interval $(0, t]$, the Lévy process be twisted with $-\zeta = -\psi(\vartheta^\star) > 0$, as described above; $\vartheta^\star$ is as defined before.
- We in addition twist the workload at time 0, Q_0; we do so by a factor $\kappa \geq 0$, for which we identify a suitable value later on. This effectively means that we sample Q_0 from a distribution with Laplace transform $\mathbb{E}e^{-(\alpha-\kappa)Q_0}/\mathbb{E}e^{\kappa Q_0}$.

We denote from now on by $\mathbb{Q}_\kappa$ this new measure, consisting of twisting Q_0 (with κ) as well as a twisting $(X_s)_{s\in(0,t]}$ (with ζ).

In each run we simulate the process under $\mathbb{Q}_\kappa$ till time t, so that we can check whether $\tau > t$ or not. In this way, we perform n independent runs. Then the estimator, based on these n runs, reads $n^{-1}\sum_{i=1}^n L_i 1\{\tau_i > t\}$, where L_i is the likelihood ratio of run i. Let us write down this likelihood ratio more explicitly. First there is the contribution due to the twisted queue at time 0; using Thm. 2 we obtain

$$L_1 := e^{-\kappa Q_0} \cdot \mathbb{E}e^{\kappa Q_0} = e^{-\kappa Q_0} \cdot \frac{-\kappa\varphi'(0)}{\varphi(-\kappa)}.$$

Then there is the contribution due to the twisted Lévy process between 0 and t:

$$L_2 := e^{\psi(\vartheta^\star)X_t} \cdot \mathbb{E}e^{-\psi(\vartheta^\star)X_t} = e^{\psi(\vartheta^\star)X_t} \cdot e^{\vartheta^\star t}.$$

The 'total likelihood ratio' is thus $L := L_1 \times L_2$. It is standard that the resulting estimator is unbiased as $\mathbb{E}_{\mathbb{Q}_\kappa} L1\{\tau > t\}$ equals the probability of our interest.

As $\mathbb{V}\mathrm{ar}_{\mathbb{Q}_\kappa} L1\{\tau > t\} \geq 0$, we see that $\mathbb{E}_{\mathbb{Q}_\kappa} L^2 1\{\tau > t\} \geq (\mathbb{E}_{\mathbb{Q}_\kappa} L1\{\tau > t\})^2$. In this sense, we could call our change of measure logarithmically efficient if

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}_{\mathbb{Q}_\kappa} L^2 1\{\tau > t\} \leq \lim_{t\to\infty} \frac{1}{t} \log(\mathbb{E}_{\mathbb{Q}_\kappa} L1\{\tau > t\})^2 = 2\vartheta^\star.$$

Logarithmic efficiency essentially means that the number of replications needed to obtain an estimate with a certain fixed precision grows subexponentially in the 'rarity parameter' t, cf. [13, Ch. VI]. In [50] it is argued that $\kappa = 0$ does not necessarily lead to logarithmic efficiency, whereas choosing $\kappa = -\zeta$ is, in a certain sense, optimal.

Estimation of workload correlation function. We now consider the problem of estimating $r(t)$; following [50] we again restrict ourselves to $X \in \mathscr{L} \cap \mathscr{S}_+$ (were it is noted that the corresponding spectrally-negative case works similarly). Note that it suffices to estimate $c(t) := \mathbb{C}\mathrm{ov}(Q_0, Q_t)$, as $v = \mathbb{V}\mathrm{ar}\, Q$ is known. The naïve estimator of $c(t)$ is, in self-evident notation, and recalling that $\mathbb{E}Q$ is known,

$$c_n^{(\mathrm{NS})}(t) := \frac{1}{n} \sum_{i=1}^{n} Q_0^{(i)}\, Q_t^{(i)} - (\mathbb{E}Q)^2,$$

based on n independent runs. The variance of this estimator reads $(n^{-1}) \cdot \mathbb{V}\mathrm{ar}(Q_0 Q_t)$. Now note that, as $t \to \infty$,

$$\mathbb{V}\mathrm{ar}(Q_0 Q_t) = \mathbb{E}(Q_0^2 Q_t^2) - (\mathbb{E}Q_0 Q_t)^2 \to (\mathbb{E}Q^2)^2 - (\mathbb{E}Q)^4,$$

which is positive due to the fact that $\mathbb{E}Q^2 > (\mathbb{E}Q)^2$. Suppose our goal is to simulate until our estimate has a certain given relative precision ε (defined as the ratio between the width of the confidence interval and the estimate) and confidence α. The number of runs needed, say $n^{(\mathrm{NS})}(t)$, is roughly equal to the smallest n satisfying

$$2\delta_\alpha \frac{\sqrt{\mathbb{V}\mathrm{ar} c_n^{(\mathrm{NS})}(t)}}{c(t)} < \varepsilon,$$

for an appropriately chosen percentile of the standard Normal distribution δ_α. Now recall that in the situation at hand $c(t)$ decays roughly exponentially. We therefore obtain the following remarkable result for the naïve estimator: it says that the number of runs required blows up exponentially, but it is *quadratically* inversely proportional to $c(t)$, rather than just inversely proportional. This result underscores that efficient (simulation-based) computation of the workload correlation $c(t)$ poses fundamentally new questions (compared to the estimation of rare event probabilities), despite the fact that its decay resembles that of the busy-period asymptotics $p(t)$.

To overcome this problem, we now consider a coupling-based algorithm, that reduces the number of runs needed from quadratically inversely proportional to $c(t)$, to just inversely proportional. We write

$$c(t) = \mathbb{E}(Q_0 \cdot (Q_t - Q_t^\star)),$$

where both Q and $Q^\star$ are stationary versions of the workload, and $Q_t^\star$ is *independent* of Q_0. We construct such a coupling as follows: generate Q_0 and $Q_0^\star$ independently, sampled from the stationary distribution of the workload. Now use exactly the same incoming Lévy process X_t over $(0, t]$ to drive both $(Q_s)_{s\in(0,t]}$ and $(Q_s^\star)_{s\in(0,t]}$ from their two independently generated initial conditions. This makes Q_t and Q_0 correlated but $Q_t^\star$ and Q_0 independent. The new estimator becomes, in self-evident notation,

$$c_n^{(\mathrm{CS})}(t) := \frac{1}{n} \sum_{i=1}^{n} Q_0^{(i)} \left(Q_t^{(i)} - Q_t^{\star\,(i)} \right),$$

based on n independent runs. The key observation is that $Q_t^{(i)} = Q_t^{\star\,(i)}$ if in both systems the busy period (that started at time 0) has ended. Mainly due to this property, it is proven in [50] that the number of runs needed is roughly inversely proportional to $c(t)$. If this algorithm is augmented with importance sampling (very similarly to the way this was done in the algorithm to estimate $p(t)$ efficiently), one even obtains a logarithmically efficient algorithm [50, Section 4.3].

7. Variants of the standard queue

So far we consider the standard infinite-buffer queue with Lévy input. This section describes a number of variants of this standard model.

7.1. **Finite-buffer queues.** We now consider a Lévy-driven queue in which the workload cannot exceed level $K > 0$; again we call the corresponding process $(Q_t)_t$. A corresponding Skorokhod problem can be formulated, in which Q_t is expressed in terms of the local time at 0 (as before), but now also the local time at K plays a role. Assuming for ease that $Q_0 = 0$, we have that $Q_t = X_t + L_t - \bar{L}_t$, with L_t ($\bar{L}_t$) the local time at 0 (at K, respectively); popularly speaking, L_t only increases when $Q_t = 0$, whereas $\bar{L}_t$ only increases when $Q_t = K$. In [5, 66] it is shown that Q_t can be explicitly solved; [66] found

$$Q_t = X_t - \sup_{s\in[0,t]} \left(\max \left\{ \min \left\{ X_s - K, \inf_{u\in[0,t]} X_u \right\}, \inf_{u\in[s,t]} X_u \right\} \right),$$

whereas the alternative solution in [5] is slightly simpler, and reads

$$Q_t = \sup_{s\in[0,t]} \max \left\{ X_t - X_s, \inf_{u\in[s,t]} (K + X_t - X_u) \right\}.$$

It was proven for the infinite-buffer model that $\mathbb{E}(Q_t \,|\, Q_0 = 0)$ is increasing and concave in t [59, 61], but, interestingly, this conclusion remains valid also in the finite-buffer case [5].
The first part of the following result [71, 91] characterizes the steady-state workload Q in terms of a first-passage time; note that it is not needed now to require $\mathbb{E}X_1 < 0$. The second part, that can be found in e.g. [26, Thm. 8, p. 194], assumes spectrally-negative input, but realize that the spectrally-positive case can be dealt with analogously. Recall the (implicit) definition of the scale function $W^{(0)}(\cdot)$ from Eqn. (4.1); write $\pi_K(u) := \mathbb{P}(Q < u)$.

Proposition 10. *(i) For $u \in [0, K]$,*

$$1 - \pi_K(u) = \mathbb{P}(X_{\tau[u-K,u)} \le u),$$

where $\tau[u, v) := \inf\{t \ge 0 : X_t \notin [u, v)\}$, for $u \le 0 \le v$.
(ii) Let $X \in \mathscr{S}_-$. Then, for $u \in [0, K]$,

$$1 - \pi_K(u) = \frac{W^{(0)}(K - x)}{W^{(0)}(K)}.$$

As we know the transform of $W^{(0)}(\cdot)$, this result characterizes $\mathbb{P}(Q \ge u)$. For the case of Brownian input, it turns out that Q has a truncated exponential distribution, as is easily checked. In [45] scale functions are used to determine the busy-period distribution in a finite-buffer M/G/1 queue.
In models with a finite buffer , there is the notion of a *loss rate* ℓ^K, which we define by, in self-evident notation,

$$\ell^K := \mathbb{E}_{\pi_K} \bar{L}_1.$$

In [17] the following result was proven for general finite-buffer Lévy-driven queues.

Proposition 11. *If $\int_1^\infty y\Pi(\mathrm{d}y) = \infty$, then $\ell^K = \infty$, and otherwise*

$$\ell^K = \frac{\mathbb{E}X_1}{K} \int_0^K x\pi_k(\mathrm{d}x) + \frac{\sigma^2}{2K} + \frac{1}{2K} \int_0^K \int_{-\infty}^\infty k(x, y)\Pi(\mathrm{d}y)\pi_K(\mathrm{d}x),$$

where $k(x,y) := -(x^2+2xy)$ *for* $y \le -x$, $k(x,y) := y^2$ *for* $-x < y < K-x$, *and* $k(x,y) := 2y(K-x)-(K-x)^2$ *for* $y \ge K-x$.

For $X \in \mathscr{L}$, [17] also studies the asymptotics of ℓ^K for K large. These are of the form $Ce^{-\omega K}$, for some rather complicated C, and ω solving $\mathbb{E}e^{\omega X_1} = 1$.

7.2. **Models with feedback.** In the queues we studied so far, the input stream was not affected by the current level of the workload. In this section we *do* allow such dependencies. We start by considering a queue whose input is $\mathbb{CP}(r(x),\lambda(x),b(\cdot))$ when the current workload level is $x \ge 0$; note that the distribution of the jobs B does *not* depend on x. Mimicking the procedure outlined in Section 3.1, a rate conservation argument shows that the density $f_Q(\cdot)$ of the stationary workload obeys [20]

$$r(x)f_Q(x) = \int_{(0,x)} \lambda(x)f_Q(y)\mathbb{P}(B > x-y)\mathrm{d}y + \lambda(0)p_0\mathbb{P}(B > x),$$

with $p_0 := \mathbb{P}(Q=0)$. In the special case that the jobs have an exponential distribution with mean $1/\mu$, we obtain after multiplication with $e^{\mu x}$ the differential equation $g'(x) = g(x)\lambda(x)/r(x)$, with $g(x) := e^{\mu x}r(x)f_Q(x)$. For the case $p_0 > 0$ we obtain by an elementary separation of variables argument that

$$f_Q(x) = \frac{\lambda(0)p_0}{r(x)} \exp\left(\int_0^x \left(\frac{\lambda(y)}{r(y)} - \mu\right)\mathrm{d}y\right),$$

under appropriate integrability conditions; the case $p_0 = 0$ should be dealt with separately. Further details can be found in [20].

In [22] the focus is on a queue fed by a spectrally-positive Lévy process, where feedback information about the workload level may lead to adaptation of the Lévy exponent. Among other models, the paper addresses the class of models in which the workload can only be observed at Poisson instants; at these Poisson instants, the Lévy exponent may be adapted based on the amount of work present at that time. In [21] a somewhat related model is studied: the focus is on a Lévy-driven queue, where the Lévy exponent of the input process alternates between two different forms (depending on the evolution of the workload process in the past). A classical related paper is [32].

7.3. **Vacation and polling models.** In [30] the following Lévy-driven queue with server vacations is studied; it can be regarded as stochastic storage process alternatingly experiencing active and passive (vacation) periods.

During active periods, work is generated according to a Lévy process $X_{\mathrm{D}}(\cdot) \in \mathscr{S}_+$ with negative drift, until the workload reaches zero (i.e., the storage reservoir is empty). From then on, the storage level behaves according to a second Lévy process $X_{\mathrm{U}}(\cdot)$, which is assumed to be non-decreasing. As during this period work accumulates in the queue, it may be interpreted as a vacation; it lasts $aI + bV$, where I is a function of the length of the preceding active period, and V is an independent vacation time, and a and b are given nonnegative scalars. The case in which the workload is still zero after $aI+bV$, has to be treated separately: then the vacation period is extended until work is generated by $X_{\mathrm{U}}(\cdot)$. Subsequently a new active period starts; etc.

Consider the sequence of epochs right before an active period starts. The transform of the storage level at such an embedded epoch can be expressed in terms of the transform at the previous embedded epoch. As these transforms should be identical in equilibrium, we can thus obtain the transform of the stationary storage level at those embedded epochs [30, Section 3]. Relying on the

Kella-Whitt martingale, they can be translated into the transform of the workload at an arbitrary epoch; see [30, Section 4]. Interestingly, these vacation models can be related to so-called *polling models*, in which a single server visits multiple queues according to some predefined discipline.

The topic of Lévy-driven polling systems is explored in full detail in [29]. There the focus is on an N-queue polling model with switchover times. Each of the queues is fed by a nondecreasing Lévy process, which can be different during each of the consecutive periods within the server's cycle. The N-dimensional Lévy processes obtained in this fashion are described by their (joint) Laplace exponent, thus allowing for non-independent input streams. Again as a first step the steady-state distribution of the workload is determined at embedded epochs (which are now polling and switching instants); importantly the *joint* transform of all N workloads is found. As before, application of the Kella-Whitt martingale yields the steady-state distribution at an arbitrary epoch. The analysis heavily relies on the link between our polling system and so-called multitype Jiřina processes (continuous-state discrete-time branching processes). The results are so general that they cover the most important polling disciplines, like exhaustive and gated.

7.4. **Models with Markov-additive input.** Markov-additive processes (MAP s) date back to [33, 77], and can be seen as the Markov-modulated version of Lévy processes; we here concentrate on MAP s in continuous time. We now give the definition of a MAP; for ease we restrict ourselves to the spectrally positive case (which we call $\mathscr{S}_+^{\rm MAP}$), but more general cases can be introduced analogously.

A MAP is a bivariate Markovian process (X_t, J_t) that is defined as follows. Let $(J_t)_t$ be an irreducible continuous-time Markov chain with finite state space $E = \{1, \dots, N\}$, transition rate matrix $Q = (q_{ij})$ and a (unique) stationary distribution $\boldsymbol{\pi}$. For each state i that J_t can attain, let $(X_t^{(i)})_t$ be a Lévy process with Laplace exponent $\varphi_i(\alpha) = \log \mathbb{E} \exp(-\alpha X_1^{(i)})$. Letting T_n and T_{n+1} be two successive transition epochs of J_t, and given that J_t jumps from state i to state j at $t = T_n$, we define the additive process X_t in the time interval $[T_n, T_{n+1})$ through

$$X_t = X_{T_n-} + U_{ij}^n + [X_t^{(j)} - X_{T_n}^{(j)}],$$

where the $(U_{ij}^n)_n$ constitute a sequence of independent and identically distributed random variables with Laplace transform

$$b_{ij}(\alpha) = \mathbb{E} e^{-\alpha U_{ij}^1},$$

where $U_{ii}^1 \equiv 0$, describing the jumps at transition epochs. To make the MAP spectrally positive, it is required that $U_{ij}^1 \geq 0$ (for all $i, j \in \{1, \dots, N\}$) and that $X_t^{(i)}$ is allowed to have only positive jumps (for all $i \in \{1, \dots, N\}$).

Observe that the modulating Markov chain does not jump in $[t, t+h)$ with probability $1+q_{jj}h+o(h)$, given $J_t = j$ (recall that $q_{jj} < 0$), and jumps to k with probability $q_{jk}h + o(h)$. We therefore obtain, in self evident notation, with

$$\Xi_{ij}(\alpha, t) := \mathbb{E}_i(e^{-\alpha X_t}, J_t = j),$$

the following equation:

$$\begin{aligned}\Xi_{ij}(\alpha,t+h) &= (1+q_{jj}h)\Xi_{ij}(\alpha,t)\mathbb{E}e^{-\alpha X_h^{(j)}} + \sum_{k\neq j} q_{kj}h\cdot \Xi_{ik}(\alpha,t)b_{kj}(\alpha)+o(h)\\ &= (1+\varphi_i(\alpha))\Xi_{ij}(\alpha,t)+h\sum_{k=1}^{N}\Xi_{ik}(\alpha,t)q_{kj}b_{kj}(\alpha)+o(h).\end{aligned}$$

Subtracting $\Xi_{ij}(\alpha,t)$ from both sides and dividing by h, we obtain a system of linear differential equations. Its solution is given in the following proposition, which shows some sort of infinite-divisibility, but now at the matrix level. In this sense, the MAP can be regarded as a genuine matrix-counterpart of the Lévy process.

Proposition 12. *The matrix $(\Xi_{ij}(\alpha,t))_{ij}$ equals $e^{M(\alpha)t}$, where $M_{ij}(\alpha):=1_{\{i=j\}}\varphi_i(\alpha)+q_{ij}b_{ij}(\alpha)$.*

Just as in the Lévy case, we can now construct MAP-driven queues, which are stable under the assumption that

$$\mathbb{E}X_1=\sum_{i=1}^{N}\pi_i\mathbb{E}X_1^{(i)}+\sum_{i\neq j}\pi_i q_{ij}\mathbb{E}U_{ij}<0.$$

Having defined these, all issues we have addressed so far for the Lévy-driven queue (stationary distribution, transience, busy periods, tail probabilities, etc.) can be studied for the MAP-driven queue as well. We will not give an exhaustive overview of all results in this area here, as a vast body of literature is devoted to this topic; we rather restrict ourselves to a relatively short account of the main findings on the stationary distribution.

In [15] martingale methods are developed such to analyze, for $X\in\mathscr{S}_+^{\rm MAP}$, the joint distribution of the steady-state workload Q and the steady-state of the Markov chain J. Under the stability condition identified above,

$$\mathbb{E}(e^{-\alpha Q},J=j)=\left(\alpha\boldsymbol{\ell}(M(\alpha))^{-1}\right)_j,$$

where $\boldsymbol{\ell}$ is a row vector. It is interesting to compare the structure of this result with Thm. 2: observe that it is essentially its MAP-counterpart. The authors of [15] do not succeed in uniquely characterizing the vector $\boldsymbol{\ell}$; it can be seen that $\sum_i \ell_i=\mathbb{E}X_1$ though. We also refer to [53] for related results.

In [36] a method is developed to determine $\boldsymbol{\ell}$. In this approach, an important role is played by the first passage time process $\tau(x):=\inf\{t\geq 0: X(t)=-x\}$. It is readily seen that $J_{\tau(x)}$ is a time-homogeneous Markov process (as a function of x), say with generator Λ. The main finding of [54] is a way to identify this matrix, relying on the theory of Jordan chains. Then $\boldsymbol{\ell}$ can be expressed in terms of the invariant that is associated with Λ; in the proof of the key result a lemma on the number of zeros of the determinant of $M(\alpha)$ plays a crucial role [54]. A different approach is described in [44].

The case of $X\in\mathscr{S}_-^{\rm MAP}$ is also dealt with in [44, 54]. Then Q has a phase-type distribution, whose parameters again follow directly with the techniques developed in [54]; this can be viewed as the MAP-counterpart of the exponential distribution identified in Thm. 3. In that paper, also the case of doubly-reflected (i.e., finite buffer capacity) Markov modulated Brownian motion is dealt with. Other important papers are e.g. [31, 87, 68], and various parts of [83].

8. LÉVY-DRIVEN TANDEM QUEUES

In this section we analyze a system consisting of two *concatenated* Lévy-driven queues, a so-called *Lévy-driven tandem queue*. This model, being a natural extension of the one-node queuing system, can be regarded as a building-block for more complex network architectures, that will be concentrated on in Section 9.

We begin by giving an informal description of the tandem queue. Consider two single-node Lévy-driven queues. The output of the first (upstream) queue is fed into the second (downstream) queue. Let r_1 be the (positive, constant) service rate at the upstream node, and r_2 at the downstream node. In order to avoid the downstream node becoming degenerate, it is assumed throughout that $r_2 < r_1$. Suppose that a Lévy process J_t feeds into the first queue, with $\mathbb{E}J_1 < r_2$ to ensure stability. We assume that no additional work enters the second queue.

Notice, however, that the notion of 'output' is not always well defined: in a queue with compound Poisson input there is a logical concept of output, but in the Brownian case we lack this. This problem can be circumvented as follows. Let $Q^{(1)}, Q^{(2)}$ be the stationary workload at first and second node respectively, and let Q denote the *total* stationary workload present in stations 1 and 2 together. The stationary workload of the upstream queue can be defined in the usual way: with $X_t^{(1)} := J_t - r_1 t$, we have due to (2.2) that $Q^{(1)}$ is distributed as $\sup_{t\geq 0} X_t^{(1)}$. In addition, the total queue behaves as a single queue fed by J_t, but emptied at rate r_2 [18, 47, 88]. Then we can reconstruct $Q^{(2)}$ as the difference between the total workload, and the workload in the upstream queue, to obtain the distributional equality

$$(Q^{(1)}, Q^{(2)}) \stackrel{\mathrm{d}}{=} \left(\sup_{t\geq 0} X_t^{(1)}, \sup_{t\geq 0} X_t^{(2)} - \sup_{t\geq 0} X_t^{(1)} \right).$$

To make the notation more compact, in the sequel we let, for $S \subset \mathbb{R}$,

$$\bar{X}_S^{(i)} := \sup_{t\in S} X_t^{(i)}, \quad \bar{X}^{(i)} = \bar{X}^{(i)}_{[0,\infty)}.$$

8.1. **A representation of the downstream workload distribution.** In this section we focus on distributional properties of the downstream queue. Based on the above, we have that

$$\mathbb{P}(Q^{(2)} > u) = \mathbb{P}\left(\bar{X}^{(2)}_{[0,\infty)} - \bar{X}^{(1)}_{[0,\infty)} > u \right). \tag{8.1}$$

We note that despite this explicit formula, its direct applicability is limited, since $(X_t^{(1)})_{t\geq 0}$ and $(X_t^{(2)})_{t\geq 0}$ are highly dependant (e.g., note that $X_t^{(1)} - X_t^{(2)} = (r_2 - r_1)t$). However, under the assumption that J is Lévy , a considerably more tractable representation can be deduced. This is done as follows.

In the first place, it can be shown that we can 'shrink' the sets over which both suprema in (8.1) are taken, so that we obtain two adjacent intervals, as follows. For given $u > 0$, we define $t_u := u/(r_1 - r_2)$, the minimal time needed for the second queue to exceed level u, starting empty. Then a sample-path argument, which intensively uses the fact that $r_1 > r_2$, leads to [38]

$$\mathbb{P}(Q^{(2)} > u) = \mathbb{P}\left(\bar{X}^{(2)}_{[t_u,\infty)} - \bar{X}^{(1)}_{[0,t_u]} > u \right).$$

Using that $X_{t_u}^{(1)} - X_{t_u}^{(2)} = u$, we have

$$\bar{X}^{(2)}_{[t_u,\infty)} - \bar{X}^{(1)}_{[0,t_u]} = \left(\bar{X}^{(2)}_{[t_u,\infty)} - X_{t_u}^{(2)} \right) - \left(\bar{X}^{(1)}_{[0,t_u]} - X_{t_u}^{(1)} \right) + u.$$

In view of the stationarity and independence of the increments, this leads to the following representation [38].

Theorem 12. *For each $u > 0$, and $(\check{X}_t^{(1)})_{t\geq 0}, (\check{X}_t^{(2)})_{t\geq 0}$ independent copies of $(X_t^{(1)})_{t\geq 0}, (X_t^{(2)})_{t\geq 0}$ respectively,*

$$\mathbb{P}(Q^{(2)} > u) = \mathbb{P}\left(\sup_{t\in[0,\infty)} \check{X}_t^{(2)} > \sup_{t\in[0,t_u]} -\check{X}_t^{(1)}\right).$$

8.2. **Steady-state workload of the downstream queue.** Direct application of Thm. 12 to the class of spectrally one-sided input processes yields the Laplace transform $\mathbb{E}e^{-\beta Q^{(2)}}$. In the case of $J \in \mathscr{S}_+$, we also obtain a representation in the spirit of (3.1).

Spectrally positive case. Indeed, assume that $J_1 \in \mathscr{S}_+$. Let $\varphi_i(\alpha) = \log \mathbb{E}e^{-\alpha X_1^{(i)}}$ and $\psi_i(\cdot)$ the corresponding inverse. Also $\tau^{(1)}(x) := \inf\{t \geq 0 : X_t^{(1)} \leq -x\}$. Then, for each $x \geq 0$, using the notation of Thm. 12,

$$\mathbb{P}\left(\sup_{t\in[0,t_u]} -X_t^{(1)} < x\right) = \mathbb{P}\left(\tau^{(1)}(x) > t_u\right)$$

and $\mathbb{E}e^{-\vartheta\tau^{(1)}(x)} = e^{-x\psi_1^{-1}(\vartheta)}$, see Lemma 1. Obviously, $\sup_{t\in[0,\infty)} X_t^{(2)} \stackrel{\mathrm{d}}{=} Q$, as we saw above. Application of the above to Thm. 12, after a few elementary steps, leads to

$$\begin{aligned}
\int_0^\infty e^{-\alpha u}\mathbb{P}(Q^{(2)} > u)\mathrm{d}u &= \int_0^\infty e^{-\alpha u}\int_0^\infty \mathbb{P}\left(\tau^{(1)}(x) > t_u\right)\mathrm{d}\mathbb{P}(Q \leq x)\mathrm{d}u \\
&= (r_1 - r_2)\int_0^\infty\int_0^\infty e^{-\alpha(r_1-r_2)v}\mathbb{P}\left(\tau^{(1)}(x) > v\right)\mathrm{d}v\mathrm{d}\mathbb{P}(Q \leq x) \\
&= \frac{1}{\alpha}\left(1 - \int_0^\infty\int_0^\infty e^{-\alpha(r_1-r_2)v}\mathrm{d}\mathbb{P}\left(\tau^{(1)}(x) \leq v\right)\mathrm{d}\mathbb{P}(Q \leq x)\right) \\
&= \frac{1}{\alpha}\left(1 - \mathbb{E}e^{-\psi_1(\alpha(r_1-r_2))Q}\right).
\end{aligned}$$

As a consequence we obtain

$$\mathbb{E}e^{-\alpha Q^{(2)}} = \mathbb{E}e^{-\psi_1(\alpha(r_1-r_2))Q}, \tag{8.2}$$

which, combined with Thm. 2, gives the following result.

Theorem 13. *Let $J \in \mathscr{S}_+$. For $\alpha \geq 0$,*

$$\mathbb{E}e^{-\alpha Q^{(2)}} = \frac{-\mathbb{E}X_1^{(2)}}{r_1 - r_2}\frac{\psi_1(\alpha(r_1-r_2))}{\alpha - \psi_1(\alpha(r_1-r_2))}.$$

Now define $\bar{\tau}^{(1)}(x) := (r_1 - r_2)\tau^{(1)}(x)$. It follows from Lemma 1 that the process $(\bar{\tau}^{(1)}(x))_{x\geq 0}$ is an *increasing* Lévy process with $\mathbb{E}e^{-\vartheta\bar{\tau}^{(1)}(x)} = e^{-x\xi(\vartheta)}$, where $\xi(\vartheta) := \psi_1((r_1 - r_2)\vartheta)$. Thm. 13 can be written in the form [38]

$$\mathbb{E}e^{-\alpha Q^{(2)}} = (1 - \varrho)\sum_{i=1}^\infty \varrho^{i-1}\left(\ell_H(\alpha)\right)^i,$$

where $H(\cdot)$ is a distribution function such that $H(x) = 0$ for $x < 0$ and $\ell_H(\alpha) := \int_0^\infty e^{-\alpha v}\mathrm{d}H(v) = \xi(\alpha)/(\varrho\alpha)$, with

$$\varrho := \lim_{\alpha\downarrow 0}\frac{\xi(\alpha)}{\alpha} = \frac{r_1 - r_2}{-\mathbb{E}X_1^{(1)}},$$

cf. [93, Eq. (23)]. As a consequence we get the following counterpart of (3.1) for the downstream queue.

Proposition 13. *Let $J \in \mathscr{S}_+$. For $u \geq 0$,*

$$\mathbb{P}(Q^{(2)} \leq u) = (1-\varrho) \sum_{i=1}^{\infty} \varrho^{i-1} H^{\star i}(u).$$

Remark 3. The distribution $H(\cdot)$ has a natural representation in the language of the Lévy measure associated with $(\bar{\tau}^{(1)}(x))_{x\geq 0}$. As it is an increasing process, there is no Brownian term. In other words, we can let $(d, 0, \Pi)$ be the characteristic triplet corresponding to this Lévy process, so that

$$\xi(\alpha) = \alpha d + x \int_0^{\infty} e^{-\alpha x} \bar{\Pi}(\mathrm{d}x),$$

where $\bar{\Pi}(x) := \Pi((x,\infty))$ is the tail of the Lévy measure and

$$H(t) = \frac{d}{\varrho} + \frac{1}{\varrho} \int_0^t \bar{\Pi}(\mathrm{d}x)$$

for all $t \geq 0$. In addition $\varrho = d + \int_0^\infty \bar{\Pi}(\mathrm{d}x)$. $\diamondsuit$

Following [38], Thm. 12 enables us to find exact distribution function of the downstream workload for several input processes.

Example 5. Suppose $J \in \mathbb{B}\mathrm{m}(0,1)$. Then the density function of $\sup_{t\in[0,t_u]} -X_t^{(1)}$ equals

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}\left(\sup_{t\in[0,t_u]} -X_t^{(1)} \leq x\right) = \sqrt{\frac{2}{\pi t_u}} \exp\left(-\frac{(x-r_1 t_u)^2}{2t_u}\right) - 2r_1 e^{2r_1 x}\left(1-\Phi_{\mathrm{N}}\left(\frac{x+r_1 t_u}{\sqrt{t_u}}\right)\right),$$

see, e.g., [19]. Combining this with Example 1 and Thm. 12 yields, after standard calculus, for $u \geq 0$,

$$\mathbb{P}(Q^{(2)} > u) = \frac{r_1 - 2r_2}{r_1 - r_2} e^{-2r_2 u} \Phi_{\mathrm{N}}\left(\frac{r_1 - 2r_2}{\sqrt{r_1 - r_2}}\sqrt{u}\right) + \frac{r_1}{r_1 - r_2}\left(1 - \Phi_{\mathrm{N}}\left(\frac{r_1}{\sqrt{r_1-r_2}}\sqrt{u}\right)\right), \tag{8.3}$$

with, $\Phi_{\mathrm{N}}(\cdot)$, as before, the distribution function of a standard Normal random variable. $\diamondsuit$

A similar argument works also for the case of $J \in \mathbb{C}\mathrm{P}(0,\lambda,b(\cdot))$ (so $X^{(i)} \in \mathbb{C}\mathrm{P}(r_i,\lambda,b(\cdot))$ for $i = 1,2$). However, in this case, $\mathbb{P}(Q_2 > u)$ is expressed in terms of a convolution involving Q and $\sup_{t\in[0,t_u]} -X_t^{(1)}$, and of the corresponding distribution functions only series representations are available, see [38].

Spectrally negative case. Using the same line of reasoning for $X \in \mathscr{S}_-$ as for $X \in \mathscr{S}_+$, and using Th. 3, we obtain

$$\mathbb{E}e^{-\beta Q^{(2)}} = \beta_0 \int_0^{\infty} \mathbb{E}e^{-\beta(r_1-r_2)\tau^{(1)}(x)} e^{-\beta_0 x}\mathrm{d}x,$$

where β_0 solves $\beta_0 = \Psi_2(0) > 0$, with $\Psi_i(\cdot)$ denoting the right inverse of $\Phi_i(\cdot) := \log \mathbb{E}e^{\beta X_1^{(i)}}$ (see Section 3.2). Now invoking Lemma 2, we eventually obtain the following result.

Theorem 14. *Let $J \in \mathscr{S}_-$. For $\beta \geq 0$,*

$$\mathbb{E}e^{-\beta Q^{(2)}} = \frac{\beta_0 \beta(r_1-r_2) - \Psi_1(\beta(r_1-r_2))\Phi_1(\beta_0)}{\Psi_1(\beta(r_1-r_2))(\beta(r_1-r_2) - \Phi_1(\beta_0))}.$$

8.3. **Downstream workload asymptotics.** The aim of this section is to characterize $\mathbb{P}(Q_2 > u)$ as $u \to \infty$. We analyze two regimes: light- and heavy-tailed input (leaving out the intermediate regime).

Light-tailed regime. To get a feel for the general form of the asymptotics, we start by focusing on the special case of $J \in \mathbb{B}\mathrm{m}(0,1)$. The more general case of $J \in \mathscr{L}$ is considered later.
Suppose $J \in \mathbb{B}\mathrm{m}(0,1)$ and $r_1 > r_2$. Then, after some lengthy but standard calculus, formula (8.3) leads to the following asymptotics, as $u \to \infty$:

(i) if $r_1 > 2r_2$, then

$$\mathbb{P}(Q^{(2)} > u)e^{2r_2 u} \to \frac{r_1 - 2r_2}{r_1 - r_2};$$

(ii) if $r_1 = 2r_2$, then

$$\mathbb{P}(Q^{(2)} > u)\sqrt{u}e^{2r_2 u} \to \frac{1}{\sqrt{2\pi r_2}};$$

(iii) if $r_1 < 2r_2$, then

$$\mathbb{P}(Q^{(2)} > u)\left(\frac{u}{r_1 - r_2}\right)^{3/2} \exp\left(\frac{r_1^2}{2(r_1 - r_2)}u\right) \to \frac{1}{\sqrt{2\pi}}\frac{4r_2}{r_1^2(r_1 - 2r_2)^2}.$$

One can now distinguish between two situations. With $r_1^\star := 2r_2$, there is qualitatively different behavior for $r_1 \geq r_1^\star$ and $r_1 < r_1^\star$. In the former case, i.e., $r_1 \geq r_1^\star$, the most likely overflow scenario of the downstream queue is that, upon overflow, the upstream queue remains essentially empty. Thus the asymptotics in cases (i)-(ii) have roughly the same shape as those of $\mathbb{P}(Q > u) = e^{-2r_2 u}$. In the latter case, i.e., $r_1 < r_1^\star$, the most likely scenario is that J feeds into the first queue at a rate of about r_1 during t_u units of time.
The observed dichotomy extends to the more general class of light-tailed inputs. Assuming $J \in \mathscr{L} \cap \mathscr{S}_+$, following the setup given in [70], the asymptotics of the Laplace transform $\mathbb{E}e^{-\alpha Q_2}$ can be analyzed by the use of the Heaviside technique. Let $\bar{t}$ be (non-zero) root of $\varphi_1(\alpha) = (r_1 - r_2)\alpha$, $t_b := \inf_\alpha \varphi_1(\alpha)/(r_1 - r_2)$ and $\bar{\alpha} := \arg\inf \varphi_1(\alpha)$. Then 'Heaviside' gives that

(i) if $\varphi_1'(\bar{t}) > 0$, then, as $u \to \infty$,

$$\mathbb{P}(Q^{(2)} > u)e^{-\bar{t}u} \to \frac{-\mathbb{E}X_1^{(2)}\varphi_1'(\bar{t})}{(r_1 - r_2)(r_1 - r_2 - \varphi_1'(\bar{t}))};$$

(ii) if $\varphi_1'(\bar{t}) = 0$, then, as $u \to \infty$,

$$\mathbb{P}(Q^{(2)} > u)\sqrt{u}e^{-\bar{t}u} \to \frac{1}{\sqrt{2\pi}}\frac{-\mathbb{E}X_1^{(2)}}{r_1 - r_2}\sqrt{\frac{\varphi_1''(\bar{t})}{r_1 - r_2}};$$

(iii) if $\varphi_1'(\bar{t}) < 0$, then, as $u \to \infty$,

$$\mathbb{P}(Q^{(2)} > u)u^{3/2}e^{-t_b u} \to \frac{1}{\sqrt{2\pi}}\frac{-\mathbb{E}X_1^{(2)}}{(t_b - \bar{s})^2}\sqrt{\frac{1}{(r_1 - r_2)\varphi_1''(\bar{s})}}.$$

Heavy-tailed regime. We now study the asymptotics of the workload of the downstream queue in the case when $J \in \mathscr{S}_+ \cap \mathscr{R}$. Before we state the main result, we relate these asymptotics to those of $\mathbb{P}(Q > u)$. Following Section 5, assume that

$$\mathbb{P}(Q > u) = u^{1-\nu}L(u)(1 + o(1))$$

as $u \to \infty$, where $L(\cdot)$ is slowly varying at ∞, with $\nu \in (1,2)$. Then, by (8.2), Thm. 2 and 'Tauber'

$$\begin{aligned}\mathbb{E}e^{-\alpha Q^{(2)}} - 1 &= \mathbb{E}e^{-\psi_1(\alpha(r_1-r_2))Q} - 1 \\ &= \Gamma(2-\nu)(\psi_1(\alpha(r_1-r_2)))^{\nu-1} L\left(\frac{1}{\psi_1(\alpha(r_1-r_2))}\right)(1+o(1)) \\ &= \Gamma(2-\nu)\varrho^{\nu-1}\alpha^{\nu-1} L\left(\frac{1}{\alpha}\right)(1+o(1)),\end{aligned}$$

as $\alpha \to 0$, since $\lim_{\alpha\to 0}\psi_1(\alpha(r_1-r_2))/\alpha = \varrho$ with $\varrho = (r_2-r_1)/\mathbb{E}X_1^{(1)}$. Thus, again using 'Tauber', we obtain the asymptotics, as $u \to \infty$,

$$\mathbb{P}(Q^{(2)} > u) = \varrho^{\nu-1} u^{1-\nu} L(u)(1+o(1)).$$

The following theorem generalizes the above findings to the case of $X_1 \in \mathscr{S}_+ \cap \mathscr{R}$, with $\varphi_1(\alpha) \in \mathscr{R}_\nu(n,\eta)$; see [70, Thm. 4.7].

Theorem 15. *Let $X_1 \in \mathscr{S}_+ \cap \mathscr{R}$, with $\varphi_1(\alpha) \in \mathscr{R}_\nu(n,\eta)$. Then, as $u \to \infty$,*

$$\begin{aligned}\mathbb{P}(Q^{(2)} > u) &= \left(\frac{-\mathbb{E}X_1^{(1)}}{r_1-r_2}\right)^{1-\nu} \mathbb{P}(Q > u)(1+o(1)) = \\ &= \frac{(-1)^{n+1}}{\Gamma(2-\nu)} \frac{\eta}{-\mathbb{E}X_1^{(2)}} \left(\frac{-\mathbb{E}X_1^{(1)}}{r_1-r_2}\right)^{1-\nu} u^{1-\nu} L(u)(1+o(1)).\end{aligned}$$

Example 6. Consider the case of $J \in \mathbb{S}(\alpha,1,0)$, with $\alpha \in (1,2)$. Then Prop. 9 and Thm. 15 immediately give the asymptotics of $\mathbb{P}(Q^{(2)} > u)$, as $u \to \infty$. $\diamondsuit$

Example 7. Suppose $J \in \mathbb{CP}(0,\lambda,b(\cdot))$ and $\mathbb{P}(B > x) = x^{-\delta}L(x)$ with $\delta > 1$. The combination of Example 4 with Thm. 15 immediately implies

$$\mathbb{P}(Q^{(2)} > u) \sim \left(\frac{r_1 - \lambda\,\mathbb{E}B}{r_1-r_2}\right)^{1-\delta} \mathbb{P}(Q > u) \sim \frac{\lambda}{r_2 - \lambda\,\mathbb{E}B}\left(\frac{r_1 - \lambda\,\mathbb{E}B}{r_1-r_2}\right)^{1-\delta} \frac{1}{\delta-1} u^{1-\delta} L(u).$$

$\diamondsuit$

8.4. **Bivariate distribution.** We now analyze the joint distribution of $Q^{(1)}$ and $Q^{(2)}$. It turns out that in order to derive the associated bivariate Laplace transform, the notion of *splitting times* is particularly useful.

Recall that $(Q^{(1)}, Q^{(2)}) \stackrel{\rm d}{=} (\bar{X}^{(1)}, \bar{X}^{(2)} - \bar{X}^{(1)})$, so that

$$\mathbb{E}e^{-\alpha Q^{(1)} - \bar{\alpha} Q^{(2)}} = \mathbb{E}e^{-(\alpha-\bar{\alpha})\bar{X}_1 - \bar{\alpha}\bar{X}_2}.$$

Also, we let $G_i := \arg\sup_{t\geq 0} X_t^{(i)}$ be the (first) epoch that $(X_t^{(i)})_{t\geq 0}$ attains its maximum, for $i = 1,2$. Then

$$\mathbb{E}e^{-\alpha\bar{X}^{(1)} - \bar{\alpha}\bar{X}^{(2)}} = \mathbb{E}e^{-\alpha X_{G_1}^{(1)} - \bar{\alpha} X_{G_1}^{(2)}} e^{-\bar{\alpha}(\bar{X}^{(2)} - X_{G_1}^{(2)})}.$$

Now, using that $X_t^{(2)} - X_t^{(1)} = (r_1 - r_2)t > 0$, we have

$$-\alpha X_{G_1}^{(1)} - \beta X_{G_1}^{(2)} = -(\alpha+\bar{\alpha})\bar{X}^{(1)} - \bar{\alpha}(r_1-r_2)G_1$$

and

$$\bar{X}^{(2)} - X_{G_1}^{(2)} = \sup_{t\geq G_1} X_t^{(2)} - X_{G_1}^{(2)},$$

since $G_2 \geq G_1$ a.s. Then a crucial step is that $-\alpha X^{(1)}_{G_1} - \bar{\alpha} X^{(2)}_{G_1}$ and $\bar{X}^{(2)} - X^{(2)}_{G_1}$ are independent; see [26, Lemma VI.6] or [37, Lemma 2.2]. This straightforwardly leads to

$$\mathbb{E}e^{-\alpha\bar{X}^{(1)}-\bar{\alpha}\bar{X}^{(2)}} = \mathbb{E}e^{-(\alpha+\bar{\alpha})\bar{X}^{(1)}-\bar{\alpha}(r_1-r_2)G_1}\mathbb{E}e^{-\bar{\alpha}(\bar{X}^{(2)}-X^{(2)}_{G_1})}.$$

The factor $\mathbb{E}e^{-\bar{\alpha}(\bar{X}^{(2)}-X^{(2)}_{G_1})}$ can be computed upon choosing in the above equality $\alpha = 0$. Hence

$$\mathbb{E}e^{-\alpha\bar{X}^{(1)}-\bar{\alpha}\bar{X}^{(2)}} = \mathbb{E}e^{-(\alpha+\bar{\alpha})\bar{X}^{(1)}-\bar{\alpha}(r_1-r_2)G_1}\frac{\mathbb{E}e^{-\bar{\alpha}\bar{X}^{(2)}}}{\mathbb{E}e^{-\bar{\alpha}\bar{X}^{(1)}-\bar{\alpha}(r_1-r_2)G_1}}.$$

Combining the above with the known fact that for $X^{(i)} \in \mathscr{S}_+$,

$$\mathbb{E}e^{-\alpha G_i-\bar{\alpha}\bar{X}^{(i)}} = -\mathbb{E}X^{(i)}_1\frac{\psi_i(\alpha)-\bar{\alpha}}{\alpha-\varphi_i(\bar{\alpha})}.$$

for $\alpha, \bar{\alpha} \geq 0$, $(\alpha, \bar{\alpha}) \neq (0,0)$, $\bar{\alpha} \neq \psi_i(\alpha)$, $i = 1,2$ (see, e.g., [26, Thm. VII.4]) directly leads to the following result, see [37].

Theorem 16. *Let $J \in \mathscr{S}_+$. For $\alpha, \bar{\alpha} \geq 0$,*

$$\mathbb{E}e^{-\alpha Q^{(1)}-\bar{\alpha}Q^{(2)}} = \frac{-\mathbb{E}X^{(2)}_1\bar{\alpha}}{\bar{\alpha}-\psi_1((r_1-r_2)\bar{\alpha})}\frac{\psi_1((r_1-r_2)\bar{\alpha})-\alpha}{(r_1-r_2)\bar{\alpha}-\varphi_1(\alpha)}.$$

This idea can be generalized to considerably more complex network structures, including n-node tandem networks and networks with a *tree-type* structure (see Section 9). The corresponding spectrally negative case can be dealt with as well, cf. Section 8.2.

Having the formula for bivariate transform of the workload, one may try to use it to explicitly obtain the joint distribution of the workloads in steady-state. Due to the complexity of this task, it was solved only in few special cases; see [69, 72]. In the following proposition we consider the Brownian tandem case, see [69].

Proposition 14. *Let $J \in \mathbb{B}\mathrm{m}(0,1)$. For $u, v \geq 0$,*

$$\begin{aligned}
&\mathbb{P}(Q^{(1)} > u, Q^{(2)} > v) = \\
&= \frac{r_2}{r_1-r_2}\left(1-\Phi_{\mathrm{N}}\left(\frac{u+vr_1/(r_1-r_2)}{\sqrt{v/(r_1-r_2)}}\right)\right) + \left(1-\Phi_{\mathrm{N}}\left(\frac{-u+vr_1/(r_1-r_2)}{\sqrt{v/(r_1-r_2)}}\right)\right)e^{-2r_1u} \\
&\quad +\frac{r_1-2r_2}{r_1-r_2}\Phi_{\mathrm{N}}\left(\frac{-u+v(r_1-2r_2)/(r_1-r_2)}{\sqrt{v/(r_1-r_2)}}\right)e^{-2((r_1-r_2)u+r_2v)}.
\end{aligned}$$

We refer to [69, 70] for the asymptotic analysis of the joint buffer overflow probabilities of the type $\mathbb{P}(Q^{(1)} > Au, Q^{(2)} > (1-A)u)$ as $u \to \infty$, for a given $A \in (0,1)$.

Several other issues concerning tandem Lévy systems, including steady-state characteristics and correlation analysis, can be found in [58, 63].

9. Networks

In this section we consider a general class of Lévy-driven queueing networks. We first formally introduce these networks through a Skorokhod formulation. Restricting ourselves to the class of tree networks, the joint Laplace transform of the workloads can be given.

9.1. **Definition, multi-dimensional Skorokhod problem.** We consider a network of n infinite-buffer queues. Queue i is externally fed by the process $J_t^{(i)}$, where it is assumed that $\boldsymbol{J} := (\boldsymbol{J}_t)_t = ((J_t^{(1)}, ..., J_t^{(n)})')_t$ is an n-dimensional Lévy process, with $\boldsymbol{J}_0 = 0$ and $\mathbb{E}|\boldsymbol{J}_1| < \infty$. Let $\boldsymbol{r} = (r_1, ..., r_n)'$, where r_i is the output rate of queue i. The interaction between the queues is given by the *routing matrix* $P = (p_{ij})_{i,j=1,...,n}$, where $p_{ij} \in [0, 1]$ is the fraction of output of station i that is immediately transferred to station j, where a fraction $1 - \sum_{j\neq i} p_{ij}$ leaves the system. We assume that $p_{ii} = 0$ and $\sum_{j=1}^{n} p_{ij} \leq 1$, for all $i = 1, \ldots, n$. We represent such a network by the triplet $(\boldsymbol{J}, \boldsymbol{r}, P)$.
Following [52, 85], the corresponding workload process $(\boldsymbol{Q}_t)_{t\geq 0} = (Q_t^{(1)}, ..., Q_t^{(n)})'_{t\geq 0}$ is the solution of the following multidimensional Skorokhod problem:

(A$^+$) $\boldsymbol{Q}_t$ given by $\boldsymbol{Q}_0 = \boldsymbol{w}$ and, for $t \geq 0$,

$$\boldsymbol{Q}_t = \boldsymbol{w} + \boldsymbol{J}_t - (I - P')\,\boldsymbol{r}t + (I - P')\boldsymbol{L}_t,$$

is non-negative for all $t \geq 0$;

(B$^+$) $\boldsymbol{L}_0 = 0$ and $\boldsymbol{L}_t$ is nondecreasing, and

$$\sum_{i=1}^{n} \int_0^T Q_t^{(i)}\, \mathrm{d}L_t^{(i)} = 0, \quad \text{for all } T > 0,$$

where $\boldsymbol{w} \geq 0$, I is the identity matrix and $\boldsymbol{L} = (\boldsymbol{L}_t)_t = (L_t^{(1)}, ..., L_t^{(n)})'_t$ is the so-called *reflecting process* or *regulator*. The reflecting process $L_t^{(i)}$ has the informal interpretation as accumulated, in time interval $[0, t]$, unused capacity of node i, for $i = 1, ..., n$. It is known that pair $(\boldsymbol{Q}, \boldsymbol{L})$ satisfying (A$^+$) and (B$^+$) exists and that it is unique; see [85].

Example 8. Consider, as introduced in Section 2.4, a single-node Lévy queue driven by $(X_t)_{t\geq 0}$. This system can be described by triplet (J_1, r_1, P), with $P = (p_{11}) = (0)$; here $X_t = J_{1,t} - r_1 t$. Then Q_t solves the corresponding Skorokhod problem. It is easily seen how the solution of the single-dimensional Skorokhod problem (A)-(B), as we introduced in Section 2.4, maps on the conditions (A$^+$)-(B$^+$).

Example 9. Consider the two-node Lévy tandem network analyzed in Section 8. This network is described by triplet $(\boldsymbol{J}, \boldsymbol{r}, P)$, with

$$\boldsymbol{J}_t = \begin{pmatrix} J_t \\ 0 \end{pmatrix}, \quad \boldsymbol{r} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}, \quad \text{and} \quad P = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

It is a simple verification that the bivariate workload process $\boldsymbol{Q}_t = (Q_t^{(1)}, Q_t^{(2)})'$ solves the related Skorokhod problem. ♢

9.2. **Lévy-driven tree networks.** Above we observed that, both for single-node queues and tandem networks, it is possible to solve the corresponding Skorokhod problem, leading to a representation for the workload process in terms of the driving triplet. In general, however, it is not clear how to explicitly express the pair $(\boldsymbol{Q}, \boldsymbol{L})$ in terms of the triplet $(\boldsymbol{J}, \boldsymbol{r}, P)$. An important large class for which this *is* possible is the class of so-called *tree-type* networks.
In the rest of this section we suppose that $(\boldsymbol{J}, \boldsymbol{r}, P)$ obeys the following properties, see [37]:

(T$_1$) P is strictly upper triangular and the j-th column of P contains exactly one strictly positive element for $j = 2, \ldots, n$;

(T$_2$) if $p_{ij} > 0$, then $p_{ij} > r_j/r_i$;

(T$_3$) the processes $(J_t^{(j)})_{t\geq 0}$ are non-decreasing for $j = 2, \dots, n$;

(T$_4$) $(I - P')^{-1}\,\mathbb{E}\boldsymbol{J}_1 < \boldsymbol{r}$.

The resulting network can be represented by a tree graph, where a directed vertex from node i to node j is put if $p_{ij} > 0$; due to (T$_1$) this corresponds to a 'tree network'. Assumption (T$_2$) is to be interpreted as some sort of work-conserving property, since $Q_t^{(i)} > 0$ implies that for all $j > i$ for which $p_{ij} > 0$ we also have $Q_t^{(j)} > 0$. Condition (T$_4$) ensures stability of the network. Importantly, we do *not* impose the requirement that J_1 be non-decreasing. Observe that both the single-node and tandem Lévy-driven queues, as described in earlier sections, satisfy (T$_1$)-(T$_4$).

9.3. **Representation for the stationary workload.** In order to find the solution of the Skorokhod problem, and hence to get the explicit representation for the workload process, we use that by [85, Thm. D.3] the process $L_t^{(i)}$ (i.e., the regulator of the i-th queue) satisfies the following fixed-point equation:

$$L_t^{(i)} = \max\left\{0, \sup_{s\in[0,t]}\left((\check{P}\boldsymbol{L}_s)_i - w_i - J_s^{(i)} + (I-\check{P})\boldsymbol{r})_i s\right)\right\}, \tag{9.1}$$

where $\check{P} := P'$. Under (T$_1$)-(T$_4$), iterating Eqn. (9.1) yields

$$L_t^{(i)} = \max\left\{0, \sup_{s\in[0,t]}\left(-\sum_{k=0}^{i-1}\left(\check{P}^k\left(\boldsymbol{J}_s - (I-\check{P})\boldsymbol{r}s + \boldsymbol{w}\right)\right)_i\right)\right\}. \tag{9.2}$$

The combination of (9.2) with the fact that $(I-\check{P})^{-1} = I+\check{P}+\check{P}^2+\ldots+\check{P}^{n-1}$ (use T$_1$!), immediately leads to ([37, Thm. 5.1]

$$\boldsymbol{L}_t = \max\left\{0, \sup_{s\in[0,t]}\left(-(I-\check{P})^{-1}(\boldsymbol{w}+\boldsymbol{J}_s) + \boldsymbol{r}s\right)\right\},$$

and hence

$$\begin{aligned}\boldsymbol{Q}_t &= \boldsymbol{w} + \boldsymbol{J}_t - (I-\check{P})\boldsymbol{r}t + (I-\check{P})\boldsymbol{L}_t\\ &= \boldsymbol{w} + \boldsymbol{J}_t - (I-\check{P})\boldsymbol{r}t + \max\left\{0, \sup_{s\in[0,t]}\left(-(\boldsymbol{w}+\boldsymbol{J}_s) + (I-\check{P})\boldsymbol{r}s\right)\right\},\end{aligned} \tag{9.3}$$

where the supremum should be interpreted componentwise. It is readily verified that this is a genuine generalization of the single-node and tandem queues that we discussed before.

In order to characterize the distribution of the steady state workload $\boldsymbol{Q}$, we follow the setup given in [37]. To this end we introduce the process

$$\boldsymbol{X}_t := (I-\check{P})^{-1}\boldsymbol{J}_t - \boldsymbol{r}t.$$

Let, as in the tandem case, $\bar{\boldsymbol{X}} = \left(\bar{X}^{(1)}, \ldots, \bar{X}^{(n)}\right)'$ be defined by $\bar{X}^{(i)} := \sup_{t\geq 0} X_t^{(i)}$. Then it is convenient to consider the transformed version of the workload process $\tilde{\boldsymbol{Q}}_t := (I-\check{P})^{-1}\boldsymbol{Q}_t$. By (9.3) we have, for any $t \geq 0$,

$$\tilde{\boldsymbol{Q}}_t = \max\left\{(\boldsymbol{x}+\boldsymbol{X}_t), \sup_{0\leq s\leq t}(\boldsymbol{X}_t - \boldsymbol{X}_s)\right\},$$

where $\boldsymbol{x} := (I-\check{P})^{-1}\boldsymbol{w}$. Due to the stationarity of increments of $\boldsymbol{X}$, we have

$$\tilde{\boldsymbol{Q}}_t \stackrel{\mathrm{d}}{=} \max\left\{(\boldsymbol{x}-\boldsymbol{X}_{-t}), \sup_{-t\leq s\leq 0}(-\boldsymbol{X}_s)\right\}.$$

Since $\boldsymbol{x} - \boldsymbol{X}_{-t} \to -\infty$, as $t \to \infty$, then

$$\widetilde{\boldsymbol{Q}}_t \xrightarrow{\mathrm{d}} \sup_{s \le 0}(-\boldsymbol{X}_s) \stackrel{\mathrm{d}}{=} \bar{\boldsymbol{X}}.$$

Finally we arrive at the following representation of the stationary workload for the $(\boldsymbol{J}, \boldsymbol{r}, P)$ network:

$$\boldsymbol{Q}_t \stackrel{\mathrm{d}}{=} (I - \check{P})\bar{\boldsymbol{X}}.$$

This argument extends to the representation of the joint distribution of the workload, *age of the busy period* $\boldsymbol{B}$ and *age of the idle period* $\boldsymbol{I}$. More precisely, we can define $\boldsymbol{B}_t = (B_t^{(1)}, ..., B_t^{(n)})'$ by

$$B_t^{(i)} := t - \sup\{s \le t : Q_s^{(i)} = 0\}$$

and $\boldsymbol{I}_t = (I_t^{(1)}, ..., I_t^{(n)})'$ by

$$I_t^{(i)} := t - \sup\{s \le t : Q_s^{(i)} > 0\}.$$

Let $\boldsymbol{G} := (G_1, ..., G_n)'$, with $G_i := \inf\{s \ge 0 : X_s^{(i)} = \sup_{t \ge 0} X_t^{(i)}\}$ (cf. the definitions for the tandem in Section 8) and $\boldsymbol{H} := (H_1, ..., H_n)'$, with $H_i := \inf\{s \ge 0 : X_s^{(i)} \ne \sup_{t \ge s} X_t^{(i)}\}$. The following result can be found in [37].

Theorem 17. *Let* (T_1)*-*(T_4) *hold for the tree fluid network* $(\boldsymbol{J}, \boldsymbol{r}, P)$*. Then for any initial condition* $\boldsymbol{Q}_0 = \boldsymbol{w}$*, the triplet of vectors* $(\boldsymbol{Q}_t, \boldsymbol{B}_t, \boldsymbol{I}_t)'$ *converges in distribution to* $(I - \check{P})(\bar{\boldsymbol{X}}, \boldsymbol{G}, \boldsymbol{H})'$ *as* $t \to \infty$*.*

With $\langle \cdot\,, \cdot \rangle$ denoting the usual inner product, the Laplace transform $\mathbb{E}e^{-\langle \boldsymbol{\alpha}, \bar{\boldsymbol{X}} \rangle - \langle \boldsymbol{\vartheta}, \boldsymbol{G} \rangle}$, for $\boldsymbol{\alpha}, \boldsymbol{\vartheta} \in \mathbb{R}_+^n$, was computed in [37, Thm. 3.3], and obeys a so-called *quasi-product formula*. It leads to the following result.

Theorem 18. *Let* (T_1)*-*(T_4) *hold for the tree fluid network* $(\boldsymbol{J}, \boldsymbol{r}, P)$*. For* $\boldsymbol{\alpha}, \boldsymbol{\vartheta} \in \mathbb{R}_+^n$*,*

$$\mathbb{E}e^{-\langle \boldsymbol{\alpha}, \boldsymbol{Q} \rangle - \langle \boldsymbol{\vartheta}, \boldsymbol{B} \rangle} = \prod_{j=1}^{n-1} \frac{\mathbb{E}e^{-[\sum_{\ell=j}^n \vartheta_\ell] G_j - \sum_{\ell=j}^n \tilde{\alpha}_\ell X_{G_j}^{(\ell)}}}{\mathbb{E}e^{-[\sum_{\ell=j+1}^n \vartheta_\ell] G_j - \sum_{\ell=j+1}^n \tilde{\alpha}_\ell X_{G_j}^{(\ell)}}} \times \mathbb{E}e^{-\tilde{\alpha}_n \bar{X}^{(n)} - \vartheta_n G_n},$$

with $\tilde{\boldsymbol{\alpha}} = (\tilde{\alpha}_1, ..., \tilde{\alpha}_n)' = (I - \check{P})\boldsymbol{\alpha}$*.*

Example 10. Consider an n-node tandem system, i.e., the network $(\boldsymbol{J}, \boldsymbol{r}, P)$ with the routing matrix P such that $p_{i,i+1} > 0$ for $i = 1, ..., n-1$, and $p_{ij} = 0$ otherwise. Assume that (T_1)-(T_4) apply and that $\boldsymbol{J}$ has mutually independent components with $J^{(1)} \in \mathscr{S}_+$. Importantly, in contrast to Section 8, we allow independent inputs to nodes $2, ..., n$; in addition, at each station some output may leave the system. Then, applying Thm. 18, for $\boldsymbol{\alpha}, \boldsymbol{\vartheta} \in \mathbb{R}_+^n$,

$$\begin{aligned}
\mathbb{E}e^{-\langle \boldsymbol{\alpha}, \boldsymbol{Q} \rangle - \langle \boldsymbol{\vartheta}, \boldsymbol{B} \rangle} &= -\mathbb{E}X_1^{(n)} \frac{\psi_n(\vartheta_n) - \alpha_n}{\vartheta_n - \varphi_n(\alpha_n)} \\
\times \quad & \prod_{j=1}^{n-1} \frac{\psi_j\left(\sum_{\ell=j+1}^n \theta_\ell^J(\alpha_\ell) + \sum_{\ell=j+1}^n (p_{\ell-1,\ell} r_{\ell-1} - r_\ell)\alpha_\ell + \sum_{\ell=j}^n \vartheta_\ell\right) - \alpha_j}{\psi_j\left(\sum_{\ell=j+1}^n \theta_\ell^J(\alpha_\ell) + \sum_{\ell=j+1}^n (p_{\ell-1,\ell} r_{\ell-1} - r_\ell)\alpha_\ell + \sum_{\ell=j+1}^n \vartheta_\ell\right) - p_{j,j+1}\alpha_{j+1}} \\
\times \quad & \prod_{j=1}^{n-1} \frac{\sum_{\ell=j+1}^n \theta_\ell^J(\alpha_\ell) + \sum_{\ell=j+1}^n (p_{\ell-1,\ell} r_{\ell-1} - r_\ell)\alpha_\ell + \sum_{\ell=j+1}^n \vartheta_\ell - \varphi_j(p_{j,j+1}\alpha_{j+1})}{\sum_{\ell=j+1}^n \theta_\ell^J(\alpha_\ell) + \sum_{\ell=j+1}^n (p_{\ell-1,\ell} r_{\ell-1} - r_\ell)\alpha_\ell + \sum_{\ell=j}^n \vartheta_\ell - \varphi_j(\alpha_j)},
\end{aligned}$$

with $\theta_i^J(\alpha) := -\log \mathbb{E}e^{-\alpha J_1^{(i)}}$, $\alpha \geq 0$, $\varphi_i(\alpha) = \log \mathbb{E}e^{-\alpha X_1^{(i)}}$ and $\psi_i(\cdot) = \varphi_i^{-1}(\cdot)$; see also [37, Thm. 6.1]. We note that if

$$\boldsymbol{J}_t = \begin{pmatrix} J_t^{(1)} \\ 0 \end{pmatrix}, \quad P = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

and choosing $\vartheta_1 = \vartheta_2 = 0$, then we recover Thm. 16. We finally mention that the case $J^{(1)} \in \mathscr{S}_-$ can be solved as well, in a way that is similar to the approach followed in Section 8.2. $\diamondsuit$

Besides the Laplace transforms of the age $\boldsymbol{B}$ of the busy periods, Thm. 18 also enables to find the Laplace transforms of the length of the steady-state running busy periods. We refer to [37, Cor. 6.1] for details. Other interesting problems related to Lévy networks, such as stability and the applicability of (quasi-)product form solutions, are analyzed in a series of papers by Kella and Whitt [55, 56, 57, 58, 62].

10. Concluding remarks

In this survey we have highlighted a set of important results on queues with Lévy input. An obvious disclaimer is in place here: with this field being large, some relevant contributions may have been overlooked. Also, given the connection between Lévy-driven queues and risk theory in a Lévy-environment — cf. Eqn. (2.2)—, perhaps not all relations with the vast finance and insurance literature have been fully exploited.

Many problems in this area are still open; we mention here a few directions. In the first place, many results are restricted to the spectrally one-sided case, where in practical situations the underlying Lévy process often has two-sided jumps — see however [11]. Another domain in which the results are still quite partial, is that of the Lévy-driven networks: hardly any results are available when the underlying network does not satisfy the conditions (T_1)-(T_4) — see however [76].

The variety of open questions, that emerge from analyzing Lévy-driven queueing systems, enforces the current research to lie at the interface between such areas as extreme value theory, fluctuation theory, stochastic geometry, large deviations, stochastic simulations etc. This makes the theory of Lévy-driven queues especially stimulating for applied and theoretical probability.

REFERENCES

[1] J. ABATE and W. WHITT (1988). The correlation functions of RBM and M/M/1. *Stoch. Models* 4, pp. 315-359.

[2] J. ABATE and W. WHITT (1995). Numerical inversion of Laplace transforms of probability distributions. *ORSA J. Comp.*, 7, pp. 36-43.

[3] J. ABATE and W. WHITT (1997). Asymptotics for M/G/1 low-priority waiting-time tail probabilities. *Queueing Syst.* 25, pp. 173-223.

[4] A. DE ACOSTA (1994). Large deviations for vector-valued Lévy processes. *Stoch. Proc. Appl.* 51, pp. 75-115.

[5] L.N. ANDERSEN and M. MANDJES (2009). Structural properties of reflected Lévy processes. *Queueing Syst.* 63, pp. 301-322.

[6] D. APPLEBAUM (2004). *Lévy Processes and Stochastic Calculus.* Cambridge University Press, Cambridge, UK.

[7] D. APPLEBAUM (2004). Lévy processes — from probability to finance and quantum groups. *Not. Am. Math. Soc.* 51, pp. 1336-1347.

[8] S. ASMUSSEN (2003). *Applied Probability and Queues.* Springer, New York, NY, US.

[9] S. ASMUSSEN (1998). Subexponential asymptotics for stochastic processes: extremal behavior, stationary distributions and first passage probabilities. *Ann. Appl. Probab.* 8, pp. 354-374.

[10] S. ASMUSSEN and H. ALBRECHER (2010). *Ruin Probabilities*, 2nd edition. World Scientific, Singapore.

[11] S. ASMUSSEN, F. AVRAM, and M. PISTORIUS (2004). Russian and American put options under exponential phase-type Lévy models. *Stoch. Proc. Appl.* 109, pp. 79-111.

[12] S. ASMUSSEN and K. BINSWANGER (1997). Simulation of ruin probabilities for subexponential claims. *ASTIN Bull.* 27, pp. 297-318.

[13] S. ASMUSSEN and P. GLYNN (2007). *Stochastic Simulation: Algorithms and Analysis*. Springer, New York, NY, USA.

[14] S. ASMUSSEN, C. KLÜPPELBERG, and K. SIGMAN (1999). Sampling at subexponential times, with queueing applications. *Stoch. Proc. Appl.* 79, pp. 265-286.

[15] S. ASMUSSEN and O. KELLA (2000). A multi-dimensional martingale for Markov additive processes and its applications. *Adv. Appl. Probab.* 32, pp. 376-393.

[16] S. ASMUSSEN and D. KROESE (2006). Improved algorithms for rare event simulation with heavy tails. *Adv. Appl. Probab.* 38, pp. 545-558.

[17] S. ASMUSSEN and M. PIHLSGÅRD (2007). Loss rates for Lévy processes with two reflecting barriers. *Math. Oper. Res.* 32, pp. 308-321.

[18] B. AVI-ITZHAK (1965). A sequence of service stations with arbitrary input and regular service times. *Man. Sci.* 11, pp. 565-571.

[19] G. BAXTER and M. DONSKER (1957). On the distribution of the supremum functional for the processes with stationary independent increments. *Trans. Amer. Math. Soc.* 85, pp. 73-87.

[20] R. BEKKER, S. BORST, O. BOXMA, and O. KELLA (2004). Queues with workload-dependent arrival and service rates. *Queueing Syst.* 46, pp. 537-556.

[21] R. BEKKER, O. BOXMA, and O. KELLA (2008). Queues with delays in two-state strategies and Lévy input. *J. Appl. Probab.* 45, pp. 314-332.

[22] R. BEKKER, O. BOXMA, and J. RESING (2009). Lévy processes with adaptable exponent. *Adv. Appl. Probab.* 41, pp. 117-205.

[23] V. BENEŠ (1957). On queues with Poisson arrivals. *Ann. Math. Stat.* 3, pp. 670-677.

[24] S. BERNSTEIN (1929). Sur les fonctions absolument monotones. *Acta Math.* 52, pp. 1–66.

[25] J. BERTOIN and R. DONEY (1994). Cramér's estimate for Lévy processes, *Statist. Probab. Lett.* 21, pp. 363-365.

[26] J. BERTOIN (1998). *Lévy Processes*. Cambridge University Press, Cambridge, UK.

[27] N. BINGHAM and R. DONEY (1974). Asymptotic properties of subcritical branching processes I: the Galton-Watson process. *Adv. Appl. Probab.* 6, pp. 711-731.

[28] N. BINGHAM, C. GOLDIE, and J. TEUGELS (1987). *Regular Variation.* Cambridge University Press, Cambridge, UK.

[29] O. BOXMA, J. IVANOVS, K. KOSIŃSKI, and M. MANDJES (2009). Lévy-driven polling systems and continuous-state branching processes. *Submitted*.

[30] O. BOXMA, M. MANDJES, and O. KELLA (2008). On a queueing model with service interruptions. *Probab. Engg. Inf. Sci.* 22, pp. 537-555.

[31] L. BREUER (2010). A quintuple law for Markov additive processes with phase-type jumps. *J. Appl. Probab.* 47, pp. 441-458.

[32] P. BROCKWELL, S. RESNICK, and R. TWEEDIE (1982). Storage processes with general release rule and additive inputs. *Adv. Appl. Probab.* 14, pp. 392-433.

[33] E. ÇINLAR (1972). Markov additive processes, II. *Probab. Th. Rel. Flds.* 24, pp. 95-121.

[34] R. CONT and P. TANKOV (2003). *Financial modelling with Jump Processes.* Chapman & Hall/CRC Press, Boca Rato, FL, USA.

[35] D. COX and W. SMITH (1961). *Queues*. Methuen, London, UK.

[36] B. D'AURIA, J. IVANOVS, O. KELLA, and M. MANDJES (2010). First passage process of a Markov additive process, with applications to reflection problems. *Submitted.*

[37] K. DĘBICKI, T. DIEKER, and T. ROLSKI (2007). Quasi-product forms for Lévy-driven fluid networks. *Math. Oper. Res.* 32, pp. 629-647.

[38] K. DĘBICKI, M. MANDJES, and M. VAN UITERT (2007). A tandem queue with Lévy input: a new representation of the downstream queue length. *Probab. Engg. Inf. Sci.* 21, pp. 83-107.

[39] K. DĘBICKI, A. ES-SAGHOUANI, and M. MANDJES (2010). Transient asymptotics of Lévy-driven queues. *J. Appl. Probab.* 47, pp. 109-129.

[40] K. DĘBICKI, K. KOSIŃSKI, and M. MANDJES (2010). On the infimum attained by a reflected Lévy process. *Forthcoming.*

[41] A. DE MEYER and J. TEUGELS (1980). On the asymptotic behaviour of the distributions of the busy period and service time in M/G/1. *J. Appl. Probab.* 17, pp. 802-813.

[42] P. DEN ISEGER (2006). Numerical transform inversion using Gaussian quadrature. *Probab. Engg. Inf. Sci.*, 20, pp. 1-44.

[43] T. DIEKER (2006). Applications of factorization embeddings for Lévy processes. *Adv. Appl. Probab.* 38, pp. 768-791.

[44] T. DIEKER and M. MANDJES (2009). Extremes of Markov-additive processes with one-sided jumps, with queueing applications. *Meth. Comp. Appl. Probab.*

[45] P. DUBE, F. GUILLEMIN, and R. MAZUMDAR (2004). Scale functions of Lévy processes and busy periods of finite-capacity M/GI/1 queues. *J. Appl. Probab.* 41, pp. 1145-1156.

[46] A. ES-SAGHOUANI and M. MANDJES (2008). On the correlation structure of a Lévy-driven queue. *J. Appl. Probab.* 45, pp. 940-952.

[47] H. FRIEDMAN (1965). Reduction methods for tandem queueing systems. *Oper. Res.* 13, pp. 121-131.

[48] H. FURRER (1997). *Risk theory and heavy-tailed Lévy processes* Ph.D. thesis, Eidgenössische Technische Hochschule, Zürich. `http://e-collection.ethbib.ethz.ch/eserv/eth:22556/eth-22556-02.pdf`.

[49] P. GLYNN and W. WHITT (1994). Logarithmic asymptotics for steady-state tail probabilities in a single-server queue. *J. Appl. Probab.* 31, pp. 131-156.

[50] P. GLYNN and M. MANDJES (2010). Simulation-based computation of the workload correlation function in a Lévy-driven queue. To appear in: *J. Appl. Probab.*

[51] J. HARRISON (1985). *Brownian Motion and Stochastic Flow Systems.* Wiley, New York, NY, US.

[52] J. HARRISON and M. REIMAN (1981). Reflected Brownian motion on an orthant, *Ann. Probab.* 9, pp. 302-308.

[53] H. KASPI (1982). On the symmetric Wiener-Hopf factorization for Markov-additive processes. *Probab. Th. Rel. Flds.* 59, pp. 179-196.

[54] J. IVANOVS, O. BOXMA, and M. MANDJES (2010). Singularities of the generator of a Markov additive process with one-sided jumps. To appear in: *Stoch. Proc. Appl.*

[55] O. KELLA (2000). Non-product form of two-dimensional fluid networks with dependent Lévy inputs, *J. Appl. Probab.* 37, 1117-1122.

[56] O. KELLA (1997). Stochastic storage networks: stationarity and the feedforward case, *J. Appl. Probab.* 34, 498–507.

[57] O. KELLA (1996). Stability and nonproduct form of stochastic fluid networks with Lévy inputs. *Ann. Appl. Probab.* 6, 186-199.

[58] O. KELLA (1993). Parallel and tandem fluid networks with dependent Lévy inputs. *Ann. Appl. Probab.* 3, pp. 682-695.

[59] O. KELLA (1992). Concavity and reflected Lévy process. *J. Appl. Probab.* 29, pp. 209-215.

[60] O. KELLA, O. BOXMA, and M. MANDJES (2006). A Lévy precess reflected at a Poisson age process. *J. Appl. Probab.* 43, pp. 221-230.

[61] O. KELLA and M. SVERCHKOV (1994). On concavity of the mean function and stochastic ordering for reflected processes with stationary increments. *J. Appl. Probab.* 31, pp. 1140-1142.

[62] O. KELLA and W. WHITT (1992). Useful martingales for stochastic storage processes with Lévy input. *J. Appl. Probab.* 33, pp. 1169-1180.

[63] O. KELLA and W. WHITT (1992). A tandem fluid network with Lévy input. In: *Queueing and Related Models*, pp. 112-128, U. Bhat and I. Basawa (eds.). Oxford University Press, Oxford, UK.

[64] A. KHINTCHINE (1932). Matematischeskaya teoriya statsionarnoi ocheredi. *Matematicheskii Sbornik* 30, pp. 73-84.

[65] C. KLÜPPELBERG, A. KYPRIANOU, and R. MALLER (2004). Ruin probabilities and overshoots for general Lévy insurance risk processes. *Ann. Appl. Probab.* 14, pp. 1766-1801.

[66] L. KRUK, J. LEHOCZKY, K. RAMANAN, and S. SHREVE (2007). An explicit formula for the Skorokhod map on $[0, a]$. *Ann. Probab.* 35, pp. 1740-1768.

[67] A. KYPRIANOU (2006). *Introductory Lectures on Fluctuations of Lévy Processes with Applications.* Springer, Berlin, Germany.

[68] A. KYPRIANOU and Z. PALMOWSKI (2008). Fluctuations of spectrally negative Markov additive processes. *Séminaire de Probabilité XLI*, pp. 121-135.

[69] P. LIESHOUT and M. MANDJES (2007). Tandem Brownian queues. *Math. Meth. Oper. Res.* 66, pp. 275-298.

[70] P. LIESHOUT and M. MANDJES (2008). Asymptotic analysis of Lévy-driven tandem queues. *Queueing Syst.* 60, pp. 203-226.

[71] D. LINDLEY (1959). Discussion of a paper of C.B. Winsten. *J. Roy. Stat. Soc. B* 21 , pp. 22-23.

[72] M. MANDJES (2004). Packet models revisited: tandem and priority systems. *Queueing Syst.* 47, pp. 363-377.

[73] M. MANDJES (2007). *Large Deviations of Gaussian Queues.* Wiley, Chichester, UK.

[74] M. MANDJES, Z. PALMOWSKI, and T. ROLSKI (2010). Quasi-stationary workload of a Lévy-driven queue. *Forthcoming.*

[75] T. MIKOSCH, S. RESNICK, H. ROOTZÉN, and A. STEGEMAN (2002). Is network traffic approximated by stable Lévy motion or fractional Brownian motion? *Ann. Appl. Probab.* 12, pp. 23-68.

[76] M. MIYAZAWA and T. ROLSKI (2009). Tail asymptotics for a Lévy-driven tandem queue with an intermediate input. *Queueing Syst.* 63, pp. 323-353.

[77] J. NEVEU (1961). Une généralisation des processus à accroissements positifs indépendents. *Abh. Math.*

Sem. Univ. Hamburg, pp. 36-61.

[78] T. OTT (1977). The covariance function of the virtual waiting-time process in an M/G/1 queue. *Adv. Appl. Probab.* 9, pp. 58-168.

[79] Z. PALMOWSKI and T. ROLSKI (2006). On the exact asymptotics of the busy period in GI/G/1 queues. *Adv. Appl. Probab.* 38, pp. 792-803.

[80] M. PISTORIUS (2004). On exit and ergodicity of the completely asymmetric Lévy process reflected at its infimum. *J. Th. Prob.* 17, pp. 183-220.

[81] F. POLLACZEK (1930). Über eine Aufgabe der Wahrscheinlichkeitsrechnung. *Math. Zeit.* 32, pp. 64-100 and 729-850.

[82] S. PORT (1989). Stable processes with drift on the line. *Trans. Am. Math. Soc.* 313, pp. 805-841.

[83] N. PRABHU (1998). *Stochastic Storage Processes*, 2nd edition. Springer, New York, NY, USA.

[84] E. REICH (1958). On the integrodifferential equation of Takács I. *Ann. Math. Stat.* 29, pp. 563-570.

[85] P. ROBERT (2003). *Stochastic networks and queues*, Springer, Berlin, Germany.

[86] L. ROGERS (1983). Wiener-Hopf factorization of diffusions and Lévy processes. *Proc. London Math. Soc.* 47, pp. 177-191.

[87] L. ROGERS (1994). Fluid models in queueing theory and Wiener-Hopf factorisation of Markov chains. *Ann. Appl. Probab.* 4, pp. 390-413.

[88] I. RUBIN (1974). Path delays in communication networks. *Appl. Math. Opt.* 1: pp. 193-221.

[89] G. SAMORODNITSKY and M. TAQQU (1994). *Stable non-Gaussian random processes: stochastic models with inifinite variance.* Chapman and Hall, New York, NY, USA.

[90] K. SATO (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, UK.

[91] D. SIEGMUND (1976). The equivalence of absorbing and reflecting barrier problems for stochastically monotone Markov processes. *Ann. Probab.* 4 , pp. 914-924.

[92] D. SIEGMUND (1976). Importance sampling in the Monte Carlo study of sequential tests. *Ann. Stat.* 4, pp. 673-684.

[93] L. TAKÁCS (1965). On the distribution of the supremum of stochastic processes with exchangeable increments. *Trans. Am. Math. Soc.* 119, pp. 367–379.

[94] M. TAQQU, W. WILLINGER, and R. SHERMAN (1997). Proof of a fundamental result in self-similar traffic modeling. *Comp. Comm. Rev.* 27, pp. 5-23.

[95] R. WERON (2001). Lévy-stable distributions revisited: tail index > 2 does not exclude the Lévy-stable regime. *International Journal of Modern Physics C*, 12, pp. 209-223.

[96] W. WHITT (2002). *Stochastic-Process Limits*. Springer, New York, NY, USA.

[97] D. WILLIAMS (1991). *Probability with Martingales*. Cambridge University Press, Cambridge, UK.

[98] V. ZOLOTAREV (1964). The first passage time of a level and the behaviour at infinity for a class of processes with independent increments. *Th. Probab. Appl.* 9, pp. 653-661.

[99] B. ZWART (2001). *Queueing Systems with Heavy Tails.* Ph.D. thesis, Eindhoven University of Technology. `http://alexandria.tue.nl/extra2/200112999.pdf`.

INSTYTUT MATEMATYCZNY, UNIVERSITY OF WROCŁAW, PL. GRUNWALDZKI 2/4, 50-384 WROCŁAW, POLAND.
E-mail address: `Krzysztof.Debicki@math.uni.wroc.pl`

KORTEWEG-DE VRIES INSTITUTE FOR MATHEMATICS, UNIVERSITY OF AMSTERDAM, THE NETHERLANDS; EURANDOM, EINDHOVEN UNIVERSITY OF TECHNOLOGY, THE NETHERLANDS; CWI, AMSTERDAM, THE NETHERLANDS
E-mail address: `M.R.H.Mandjes@uva.nl`