HAL
open science

On matrix exponential approximations of the infimum of a spectrally negative Levy process

Florin Avram, Andras Horvath, M.R. Pistorius

To cite this version:

Florin Avram, Andras Horvath, M.R. Pistorius. On matrix exponential approximations of the infimum of a spectrally negative Levy process. 2012. hal-00739232

HAL Id: hal-00739232

https://hal.science/hal-00739232

Submitted on 9 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On matrix exponential approximations of the infimum of a spectrally negative Levy process

F. Avram[a], A. Horváth[b], M. R. Pistorius[c]

[a] *Labo. de Math. Appliquées, Université de Pau et des Pays de l'Adour, Pau, France*
[b] *Dipart. di Informatica. Università di Torino Corso Svizzera, 185 10149, Torino, Italy*
[c] *Department of Mathematics Imperial College, London, United Kingdom*

Abstract

We recall four open problems concerning constructing high-order **matrix-exponential approximations** for the infimum of a spectrally negative Levy process (with applications to **first-passage/ruin probabilities**, the **waiting time distribution** in the M/G/1 queue, pricing of barrier options, etc).

On the way, we provide a new approximation, for the perturbed Cramér-Lundberg model, and recall a remarkable family of (not minimal order) approximations of Johnson and Taaffe [JT89], which fit an **arbitrarily high number of moments**, greatly generalizing the currently used approximations of Renyi, De Vylder and Whitt-Ramsay. Obtaining such approximations which fit the Laplace transform at infinity as well would be quite useful.

keywords: Levy process; first passage problem; Pollaczek-Khinchine formula; method of moments; matrix-exponential function; admissible Padé approximation; Johnson-Taaffe approximations; two-point Padé approximations

1. Introduction

Motivation: The problem of approximating distributions based on empirical data like moments is one of the bread and butter problems of applied probability.

*Corresponding author, Tel : +33 (0)5 59 40 75 63, Fax : +33 (0)5 59 40 75 55

In risk theory and related **first passage applications** (queueing, mathematical finance, ...) there is special interest in approximating densities of nonnegative random variables by **affine combinations of exponentials**, also called **GHE densities (generalized hyper-exponential)**, or, more generally, by **matrix exponential distributions** (which allow for "collision of exponents"). One reason for that is that this class captures the asymptotic behavior in the important "light tails case".

There exists already a quite extensive literature, based on inverting the explicit Pollaczek-Khinchine formula for the Laplace transform, which assumes however complete knowledge of the model, for example knowledge of all its moments.

Since input data are never certain, it is interesting to develop **approximations based on finitely many moments** (the coefficients of the power series expansion of the Laplace transform around 0).

Some well-known such approximations in risk theory and queueing are the **Renyi, De Vylder, Gamma**, and **Whitt -Ramsay** approximations, obtained by **fitting one, two or three moments**, and the minimal (but arbitrarily high) order **three moments fitting formulas of Bobbio, Horvath and Telek** [BHT05]. Also useful is the **Cramér-Lundberg** approximation

$$\Psi(x) \sim Ce^{-\gamma x} \tag{1}$$

where $\Psi(x)$ denotes the ruin probability (5) and $-\gamma$ is the so called adjustment coefficient (i.e. the largest negative root of the Cramér-Lundberg equation (13)), which captures the asymptotic behavior in the case of light tail claims. All these approximations may be derived from the explicit Laplace transform provided by the Pollaczek Khinchine formula (6).

Producing higher order approximations in the "intermediate regime" when a finite, but larger number of moments is known, seems a very challenging problem.

Question 1. *In view of the scarcity of approximations fitting more than three moments, it is natural to ask* **what are the difficulties blocking the development of high order moments based approximations***?*

Below, we examine this question in the context of first passage theory for **spectrally negative Levy processes**.

Padé, two-point Padé and other rational approximations. It turns out that most of the approximations currently used are Padé approximations

of Laplace transforms, and that higher order approximations are quite easy to obtain (for example using the Mathematica command **PadeApproximant**), since the conversion from moments to a rational Laplace approximation requires only solving a linear system. It is also quite easy to produce Padé approximations with specified limiting behavior of the Laplace transform at ∞, so called **two-point Padé approximations** – see Example 2.

Other rational approximations of interest are those minimizing in least squares sense the sum of the coefficients – see Beylkin and Monzón [BM05]. Another interesting class are "**Tijms approximations**", which try to incorporate moments fitting with including the exact Cramér-Lundberg asymptotics (1) as dominant term –see [Wil98]. These may also be obtained by using a Padé approximation with a prescribed pole.

The admissibility of Padé approximations in probability. While in principle a great tool due to their easiness of implementation, and their convergence for large n, Padé approximations (and variations, like two point Padé) applied to Laplace transforms in probability have the drawback of the difficulty to guarantee **"admissible inverses"**, i.e. **nonnegative densities** and **non-increasing survival functions**, when fitting three moments or more is desired (note that for fitting two moments m_1, m_2 of a nonnegative random variable, the admissible Gamma approximation

$$f(x) \sim \tilde{f}(x) = \frac{(\mu x)^{\alpha-1}}{\Gamma(\alpha)} \mu e^{-\mu x},\ \alpha = \frac{m_1^2}{m_2 - m_1^2}, \mu = \frac{m_1}{m_2 - m_1^2}$$

provides an easy solution).

Even ensuring the **nonnegativity of combinations of exponentials with fixed given rates** is quite a difficult question (since this involves an infinite number of constraints), still open nowadays, except for two exponentials (when nonnegativity of $\tilde{f}(0)$ and nonnegativity of the coefficient of the asymptotically dominant exponent are clearly necessary and sufficient), and for three exponentials [DL82].

Example 1. *Consider the example due to Harris [HWR92] (see also [Fac03, Ch. 5.4])*

$$f(t) = 2e^{-t} - 6e^{-2t} + 6e^{-3t} = 2e^{-t} - 3(2e^{-2t}) + 2(3e^{-3t})$$
$$\Leftrightarrow \bar{F}(t) = 2e^{-t} - 3e^{-2t} + 2e^{-3t}$$

with **canonical coordinates** $(2, -3, 2)$, *and Laplace transform*

$$f^*(s) = \frac{2\left(s^2+2s+3\right)}{(s+1)(s+2)(s+3)} = \frac{2}{3}\frac{6}{(s+1)(s+2)(s+3)} - \frac{1}{3}\frac{6}{(s+2)(s+3)} + \frac{2}{3}\frac{3}{s+3}$$

Since the "Coxian coordinates" $(2/3, -1/3, 2/3)$, *produced by the partial fractions decomposition above are not nonnegative, this is not a phase-type distribution of order* 3.

However, by an admissibility criteria for **combinations of three negative exponentials** *due to [DL82]* [§] *(see also [Fac03, BFT08] for a criterion which allows colliding exponents), we know this is a proper density. In fact, it is a phase-type density of order* 4 *– see* (2)

Admissibility by phase-type representations. One "lucky case" in which admissibility is automatic is when one has obtained somehow **any phase-type representation** $PH(\boldsymbol{\alpha}, A)$ with $\boldsymbol{\alpha}$ a **probability vector** and A a **subgenerator** matrix (satisfying $A_{ij} \geq 0$ for $i \neq j$ and $A\mathbf{1} \leq \mathbf{0}$). In that case, nonnegativity follows from the probabilistic interpretation of $f(t)$ as the density of the absorbtion time of the corresponding Markovian semigroup.

However, determining when a phase-type representation exists is again notoriously difficult, the so called **positive realization** problem of systems theory [§] .

In example 1, it is possible to show that a phase-type of minimal order 4 is available, by a recursive approach of decomposing the admissibility domain as union of higher order admissibility polytopes associated to phase-type

[§]with exponents equal to $i = 1, 2, 3$, this is $-\omega_2 \leq 2\sqrt{\omega_1\omega_3}$, or, after normalization

$$\omega_1 + 1 \leq \omega_3 \leq \omega_1 + 1 + 2\sqrt{\omega_1}.$$

[§]Finding the minimum possible order of such a representation is even harder, and known currently only for three moments fitting representations). Attesting further to the difficulty of providing admissible phase-type approximations are several interesting recent approaches, like the recursive **minimal order three moments fitting** formulas of [BHT05], and the **EM algorithm** approach [WL11].

representations – see [O'C93].One phase-type representation is:

$$A = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -2 & 2 & 0 \\ 0 & 0 & -3 & 3 \\ 0 & 0 & 0 & -4 \end{pmatrix}, \qquad \alpha = (1/2, 0, 0, 1/2) \tag{2}$$

BUTools http://webspn.hit.bme.hu/ telek/tools/butools/butools.html only obtains a representation of order 5, further illustrating the difficulty of this problem.

Question 2. *Currently, no algorithmic approach for testing admissibility of combinations of more than four given negative exponentials (or four terms Müntz polynomials) is available* [§] *.*

The question of providing admissible approximations (with non-specified rates) fitting more than four given values of a Laplace transform (for example moments) is a priori even more difficult. However, two remarkable exceptions in which this the challenging admissibility problem was solved are the **minimal order three moments fitting admissible approximations** of Bobbio, Horvath and Telek [BHT05], obtained by a recursive approach on the order, and the non-minimal ones of Johnson and Taaffe [JT89], an outcome of the classical moments theory, which work for any number of moments.

Contents and contributions. This problem was motivated by the desire to provide new admissible matrix exponential approximations in ruin theory.

Necessary ruin theory background, including the Pollaczek Khinchine transform, is reviewed in Section 2. Section 3 presents the "key characters" in our ruin application, the aggregate loss and its moments – see (16), (15).

First order Padé approximations of the Pollaczek Khinchine transform due to Renyi and DeVylder are reviewed in Section 4. We also provide here a new approximation, for the perturbed Cramér-Lundberg model – see

[§] a solution might however be possible by the approach of Faybusovich [Fay02, Thm 4,5], who offers a general representation of the Koszul-Vinberg characteristic function [See12] of the positivity cone generated by any Chebyshev system, as Pfaffian of a matrix of multiple integrals (and taking logarithm yields a "self-concordant barrier" function).

Theorem 1. The remarkable Johnson-Taaffe approximations are reviewed in Section 5, Theorem 2. Using their approach, we may "update" a second order approximation due to Ramsay to make it work for arbitrary claims having three moments – see Theorem 3.

In the case of random sums, it is possible to apply the Johnson-Taaffe approach both to the individual summands – we call this a "Ramsay-type approximation", and directly to the sum – yielding "Beekman-Bowers-type approximations". Theorem 4 in Section 6 provides a comparison between these two methods, by comparing their explicit **3 moments Johnson-Taaffe orders**.

Section 7 discusses two-point Padé approximations –see Theorem 5, whose admissibility is an open problem for the moment. Finally, in Section 8 we provide some numerical examples.

2. Ruin theory background

The perturbed Cramér Lundberg risk process models the reserves of an insurance companyby:

$$X(t) = u + c\,t - S(t) + \sigma W(t), \qquad S(t) = \sum_{k=1}^{N_\lambda(t)} Z_k \tag{3}$$

used in collective risk theory to describe the surplus $X = \{X(t), t \geq 0\}$ of an insurance company. Here,

1. u is the initial capital,
2. $c\,t$ represents the premium income up to time t,
3. Z_k are i.i.d. positive random variables representing the claims made, with cumulative distribution function and density denoted by $F(x)$ and $f(x)$, and (some) moments denoted by $m_i, i = 1, 2, ...,$
4. $N = \{N_\lambda(t), t \geq 0\}$ is an independent Poisson process with intensity λ modeling the times at which the claims occur, and
5. $W(t)$ is an infinite variation spectrally negative perturbation, for example a standard Wiener motion, and $\sigma > 0$ is a scale parameter.

Since the jumps of X are all negative, the moment generating function $\mathbb{E}[\mathrm{e}^{sX(t)}]$ exists for all $s \geq 0$ and $t \geq 0$, and is log-linear in t. The symbol/Laplace exponent/cumulant generating function $\kappa(s)$ is defined by

$$\kappa(s) = \log\left(\mathbb{E}_0 \mathrm{e}^{sX(1)}\right) = \boxed{s\left(c - \lambda \bar{F}^*(s) + \tilde{\kappa}(s)\right)}, \tag{4}$$

where $\bar{F}^*(s)$ denotes the Laplace transform of the survival function of the claims, and $s\tilde{\kappa}(s)$ is the symbol of the perturbation $\sigma W(t)$. For example, in the Wiener case $\tilde{\kappa}(s) = \frac{\sigma^2}{2}s$.

Let T be the first passage time of a stochastic process $X(t)$ below 0:

$$T := \inf\{t \geq 0 : X(t) < 0\}.$$

The objects of interest in classical ruin theory are the "finite-time" and "eventual" ruin probabilities

$$\Psi(t,u) = P_u[T \leq t], \qquad \boxed{\Psi(u) = P_u[T < \infty] = P_0[L > u]},$$

where $L = -\underline{X}$, also called maximal aggregate loss, is the negative of all-time infimum of the process (3), started from 0.

The ultimate ruin probability $\Psi(u)$ for (3) is not identically 1 iff the Levy drift/profit rate

$$\boxed{p := c - \lambda m_1 > 0 := \lambda m_1 \theta} \tag{5}$$

is positive, in which case adding the condition $\lim_{x\to\infty} \Psi(x) = 0$ determines it uniquely.

The Pollaczek-Khinchine formula. Taking Laplace transform of the Kolmogorov equation for the **ultimate ruin probabilities** of the perturbed Cramér-Lundberg model yields the Pollaczek-Khinchine formula:

$$\Psi^*(s) = \frac{1}{s} - \frac{\kappa'(0)}{\kappa(s)} = \frac{\rho(1 - f_e^*(s)) + \tilde{\kappa}(s)/c}{s(1 - \rho f_e^*(s) + \tilde{\kappa}(s)/c)} := \frac{1}{s}(1 - \psi^*(s)) \tag{6}$$

The first expression, in terms of the symbol $\kappa(s)$ of the Levy process involved, is valid for all spectrally negative Lévy processes, – see for example [Kyp06]). The second emphasizes the fact that the result in the Cramér-Lundberg case depends only on the "equilibrium density" of the claims, defined by

$$f_e(x) := \bar{F}(x)/m_1, \tag{7}$$

the estimation of which may be a convenient starting point. The third expression is equivalent to $\Psi(x) = \int_x^\infty \tilde{\psi}(u)du, \qquad x > 0$ where $\tilde{\psi}(u)$ is the **density of the continuous part of the distribution of the aggregate**

loss L (while $\Psi(x)$ is the survival function of L). The Pollaczek Khinchine formula for $\psi^*(s)$

$$\boxed{\psi^*(s) = \mathbb{E}e^{-sL} = 1 - s\Psi^*(s) = \frac{\kappa'(0)}{\kappa(s)/s} = \frac{1-\rho}{1-\rho f_e^*(s) + \tilde{\kappa}(s)/c}}. \tag{8}$$

In the case $\sigma = 0$, a beautiful probabilistic interpretation of the (non-perturbed) Pollaczek Khinchine formula (8) was discovered (independently) by Benes, Kendall and Dubordieu, by expanding the denominator into a geometric series:

$$\boxed{\psi^*(s) = \frac{1-\rho}{1-\rho\widehat{f_e}(s)} = \sum_{k=0}^{\infty}(1-\rho)\rho^k \widehat{f_e^k}(s)} \tag{9}$$

where $f_e(x) = \bar{F}(x)/m_1$ is the equilibrium distribution of Z_i. This reveals that $\psi^*(s)$ is the Laplace transform of a geometric sum of convolutions of the equilibrium/stationary distribution, i.e. $\psi(x) = (1-\rho)\sum_{n=0}^{\infty}\rho^k f_e^{(*,k)}(x)$ [§] , which may be visualized by examining the "ladders" of the paths (the amounts by which the process $-ct + S(t)$ jumps to new maxima).

This is the so called Pollaczek Khinchine "ladder decomposition" $L = \sum_{n=1}^{N} L_n$ of the "maximal aggregate loss random variable" L, with N a geometric r.v. $\Pr[N = k] = (1-\rho)\rho^k, k = 0, 1, ..., \rho = \frac{\lambda m_1}{c}$, representing the number of ladders.

Remark 1. The ladder decomposition, $\sigma > 0$. *In the Brownian perturbed case, a beautiful probabilistic interpretation of the Pollaczek Khinchine formula* (8) *was recently discovered by Dufresne-Gerber and rederived in an elementary way by Kella, by rewriting* (8) *as:*

$$\psi^*(s) = \frac{1-\rho}{1-\rho f_e^*(s) + \tilde{\kappa}(s)/c} = \frac{1}{1+\tilde{\kappa}(s)/c}\frac{1-\rho}{1-\rho f_e^*(s)\frac{1}{1+\tilde{\kappa}(s)/c}} \tag{10}$$

reflects the fact that L is an independent sum of a "first creep at the current infimum" (which in the diffusion case is an exponential of rate $\frac{2c}{\sigma^2}$) and of an

[§] Another interpretation of $\psi^*(s)$ is as Laplace transform of the stationary waiting time of the M/G/1 queue –see for example [Asm03, Thm VIII.5.7].

alternating geometric sum of "compound Poisson ladders and further creeps" when $\sigma > 0$ – see [DG91, Fig 2], [Kel11].

In the case that the Brownian perturbation $\sigma W(t)$ is replaced by a **general spectrally negative perturbation** *Y with non-zero expectation $E[Y(1)] \geq 0$, a similar ladder decomposition holds true. Let*

$$\kappa^*(s) = s\tilde{\kappa}(s)$$

denote the Laplace exponent of Y. Note that by the Wiener-Hopf factorisation– see e.g. [Kyp06], $\tilde{\kappa}(s)$ is Laplace exponent of the possibly killed downward ladder process (and s is the Laplace exponent of the up-crossing ladder process), that is, $\tilde{\kappa}$ takes the form

$$\tilde{\kappa}(s) = c + \delta s + \int_{(0,\infty)} (1 - \mathrm{e}^{-sx})\nu(dx),$$

where the killing rate c and the drift δ are non negative constants and the Lévy measure ν satisfies the integrability condition

$$\int_{(0,\infty)} [1 \wedge x]\nu(dx) < \infty.$$

Then, the formula (10) still holds, providing a decomposition of L as an independent sum of the **increment of the ladder process of the perturbation Y at an independent exp(c)- random time** *and the geometric sum of further such increments and "compound Poisson ladder height increments". In particular, if the perturbation is given by a completely asymmetric stable process, i.e. $\kappa^*(s) = s^\alpha$, $\alpha \in (1,2)$, then $\kappa^*_-(s) = s^{\alpha-1}$ and we identify $\frac{1}{1+\kappa^*_-(s)/c}$ as the Laplace transform of the non-negative random variable $\tilde{Y}^{\alpha-1}_{e(c)}$, where $Y^{\alpha-1}$ is a stable subordinator with parameter $\alpha - 1$.*

Remark 2. *Note that the aggregate loss L is the mixture of a discrete mass of $\psi^*(\infty) = 1 - \Psi(0) = 1 - \rho$ at 0, and of a continuous random variable. Letting $\tilde{\psi}^*(s)$ denote the Laplace transform of the density $\tilde{\psi}(u)$, note the decomposition*

$$\psi^*(s) = 1 - \rho + \rho\tilde{\psi}^*(s) \;\Leftrightarrow\; \psi(x) = (1-\rho)\delta_0(x) + \rho\tilde{\psi}(x)$$
$$\tilde{\psi}^*(s) := \frac{\psi^*(s) - \psi^*(\infty)}{\psi^*(0) - \psi^*(\infty)} = \frac{\psi^*(s) - (1-\rho)}{\rho} = (1-\rho)\frac{f_e^*(s)}{1 - \rho f_e^*(s)}, \quad (11)$$

where we denoted by $\psi(u)$ the inverse Laplace transform of $\psi^*(s)$, given by the generalized function $\psi(u) = \tilde{\psi}(u) + (1-\rho)\delta_0(u)$.

The behavior at ∞ of $\psi^*(s)$ distinguishes between the nonperturbed ($\sigma = 0$) and perturbed case ($\sigma > 0$):

$$\lim_{s\to\infty} \psi^*(s) = \lim_{s\to\infty} 1 - s\Psi^*(s) = 1 - \Psi(0) = \begin{cases} 1 - \frac{\lambda m_1}{c} = 1 - \rho = \frac{p}{c}, & \sigma = 0 \\ 0, & \sigma > 0 \end{cases} \tag{12}$$

The Laplace transform $\psi^*(s)$ is an essential quantity in the theory of Lévy processes

Remark 3. *The roots of the denominator $\kappa(s) = 0$ in the Pollaczek Khinchine formula determine the asymptotic behavior of ultimate ruin probabilities. More generally, an important role is played by the roots of the Cramér Lundberg equation*

$$\kappa(s) = q, q > 0. \tag{13}$$

Remark 4. *In practice, the true distribution of the claims (and interarrival times) is of course unknown, and since the Pollaczek Khinchine formula requires this knowledge, it should be viewed more as a theoretical than a practical tool. It may be argued that the most reliable information available in insurance data is contained in the first few integer moments, and thus it seems natural to start building approximations by fitting moments, or, equivalently, by Padé interpolation of the Laplace transform at the origin.*

3. The moments of the aggregate loss

From now on, we will assume a classic Brownian perturbation. As an alternative to Ramsay's approximation of the ladder distribution, we may approximate directly the aggregate loss distribution. One advantage is that the second moment of the aggregate loss satisfies automatically the second order representability constraint $c_v \geq 1/2$. This allows focusing on the third moment constraint.

Consider the expansion

$$\Psi^*(s) = \frac{\eta_{2,\sigma}/2 - \eta_3 s/6 + ...}{p + \eta_{2,\sigma} s/2 - \eta_3 s^2/6 + ...} \Leftrightarrow \tag{14}$$

$$\psi^*(s) = \frac{p}{p + \eta_{2,\sigma} s/2 - \eta_3 s^2/6 + ...}$$

$$= \frac{\theta}{1 + \theta - \left(1 - (\frac{m_2}{m_1} + \frac{\sigma^2}{2\lambda m_1})s/2 + \frac{m_3}{m_1}s^2/6 - \frac{m_4}{m_1}s^3/4! + ...\right)}$$

$$= \frac{1-\rho}{1 - \rho\left(1 - (\frac{m_2}{m_1} + \frac{\sigma^2}{2\lambda m_1})s/2 + \frac{m_3}{m_1}s^2/6 - \frac{m_4}{m_1}s^3/4! + ...\right)}$$

where $\eta_i = \lambda m_i, i = 0, 1, 2, ...$ are the moments of the Lévy measure, and $\eta_{2,\sigma} = \eta_2 + \sigma^2$.

Remark 5. *When $\sigma = 0$, the expression in the last parenthesis of* (14)

$$\widehat{f}(s) := 1 - \frac{m_2}{2m_1}s + \frac{m_3}{3m_1}s^2/2 - \frac{m_4}{4m_1}s^3/3! + ... = \mathbb{E}[e^{-sL_i}|\{L_i > 0\}]$$

has moments $\tilde{m}_i = \frac{m_{i+1}}{(i+1)m_1}$, identifying L_i as the famous equilibrium/stationary excess/ladder variable generated by Z_i, with density $f_e(x) = \bar{F}(x)/m_1$, and **stationary excess moments** $\tilde{m}_i = \frac{m_{i+1}}{(i+1)m_1}$.

Remark 6. *Let $l_i = \mathbb{E}L^i, \lambda_i = l_i/i!, i \geq 1$ denote the moments and "factorially reduced moments" obtained by "normalizing" with respect to the exponential moments of the maximal aggregate loss. These may be easily obtained, either by the recursion equivalent of the Pollaczek Khinchine formula:*

$$\theta\lambda_n = \tilde{\mu}_n + \sum_{k=1}^{n-1} \tilde{\mu}_k \lambda_{n-k}, \qquad n \geq 1, \tilde{\mu}_k := \frac{\tilde{m}_k}{k!} = \frac{\mu_{k+1}}{m_1}$$

or by expanding (14) *in power series. The first factorially reduced moments are:*

$$\lambda_1 = \mathbb{E}L = \frac{\tilde{m}_1}{\theta} + \frac{\sigma^2}{p} = \frac{\lambda m_2 + \sigma^2}{2p}, \tag{15}$$

$$\lambda_2 = \mathbb{E}L^2/2! = \frac{\tilde{m}_2}{2\theta} + \lambda_1^2, \ \lambda_3 = \mathbb{E}L^3/3! = \frac{\tilde{m}_3}{3!\theta} + 2\frac{\tilde{m}_1\tilde{m}_2/2}{\theta^2} + \frac{\tilde{m}_1^3}{\theta^3}$$

$$\lambda_4 = \mathbb{E}L^4/4! = \frac{\tilde{m}_4}{4!\theta} + \frac{2\tilde{m}_1\tilde{m}_3/3! + (\tilde{m}_2/2)^2}{\theta^2} + \frac{3\tilde{m}_1^2\tilde{m}_2/2}{\theta^3} + \frac{\tilde{m}_1^4}{\theta^4},$$

and the mass of the continuous part is $\lambda_0 := \Pr[L > 0] = \rho = \frac{1}{1+\theta}$.

Remark 7. *When $\sigma = 0$, the factorially reduced moments λ_k admit an interesting decomposition:*

$$\psi^*(s) = (1-\rho)\frac{1}{1-\rho\widehat{f_e}(s)} = \sum_{k=0}^{\infty}(1-\rho)\rho^k(f_e^*(s))^k =$$
$$1 + \sum_{k=1}^{\infty}\lambda_k(-s)^k = 1 + \sum_{k=1}^{\infty}(-s)^k \sum_{i_1+i_2+\ldots=k}^{\prime} \psi_{i_1}\psi_{i_2}\ldots$$

where $\sum'$ denotes sum over all decompositions of k as a sum, and

$$\psi_i = \frac{\lambda m_{i+1}}{p(i+1)!} = \theta^{-1}\frac{\tilde{m}_i}{i!} = \theta^{-1}\widetilde{\mu}_i > 0,$$

where $\tilde{m}_i := \frac{m_{i+1}}{(i+1)m_1}, \mu_i := \frac{m_i}{i!}$ and $\tilde{\mu}_i := \frac{\tilde{m}_i}{i!}$.

Remark 8. *The moments of the conditioned continuous r.v. $L|L > 0$, necessary for applying certain results from the literature, may be obtained by dividing by $\rho = \frac{1}{1+\theta}$.*

Remark 9. *Note that the corresponding moments l_i may also be viewed as moments of the ruin function:*

$$\begin{aligned}
\Psi_0 &:= \lambda_1 = \int_0^{\infty}\Psi(x)dx = \frac{\tilde{m}_1}{\theta},\ \Psi_1 := \lambda_2 = \int_0^{\infty} x\Psi(x)dx = \frac{\tilde{m}_2}{2\theta} + (\frac{\tilde{m}_1}{\theta})^2, \\
\Psi_k &:= \frac{\int_0^{\infty} x^k\Psi(x)dx}{k!} = \lambda_{k+1}, \ldots
\end{aligned}$$

4. Renyi, De Vylder, and a new simple approximation for ruin probabilities

Approximations of ultimate ruin probabilities. The problem of approximating ultimate ruin probabilities $\Psi(u)$ for the Cramér Lundberg model (3) using data on the distribution $F(u)$ of the claims is a classic of applied probability, dating back before 1900. Its roots may be traced back to the Danish mathematician TN. Thiele, who founded the first insurance company, Hafnia, who is also the **inventor of cumulants** and of **Thiele continued fractions** useful for Laplace transform inversion [Coh07].

In this section, we make the observation that the exponential mixture approximations recalled in the introduction are particular cases of Padé approximations of Laplace transforms (a theme already present in Thiele's research preoccupations). We also provide a new simple approximation in this vein for the perturbed Cramér-Lundberg model in Theorem 1.

Remark 10. *Concerning our application which involves random sums, we make the observation that it is possible to use Padé, Johnson-Taaffe and any other rational approximations of Laplace transforms at three levels:*

1. *for the density of the claims, based on the estimates of* $m_i, i \geq 1$
2. *for the equilibrium density of the claims, using the estimates of the equilibrium moments* $\tilde{m}_i, i \geq 1$, *and the profit rate* p. *The second level is intuitively superior to the first, since the equilibrium density is monotonically decreasing, even when the claims density isn't* [§] . *We will call this Ramsay type approximation.*
3. *for the aggregate loss density transform* $\psi^*(s)$ *(or, equivalently, the ccdf transform* $\Psi^*(s)$*), using directly the moments* $\lambda_{i+1}, i \geq 0$, *given in* (15). *This amounts to* **working directly with the Pollaczek-Khinchine formula**, *instead of approximating its denominator, and so intuitively, should be better, at least under certain conditions.* [§] . *We will call this Beekman-Bowers type approximation.*

One Padé approximation we consider here is:

$$(\psi^*)^{(n)}(s) = \mathcal{P}ade_{(m,n)}(\psi^*(s)) = \mathcal{P}ade_{(m,n)}([\frac{p}{\kappa(s)/s}]_N) \tag{16}$$

where $\mathcal{P}ade_{(m,n)}$ denotes the "classic Padé approximation" based on the Taylor series around 0, where $m = n$ or $m = n + 1$ for the classic/perturbed Cramér Lundberg process, respectively, and where $[f(s)]_N$ denotes truncation of a power series to its first N terms, with $N = m + n + 1$ (for "theoretical models" where an expression for $\psi^*(s)$ is available, we may also take

[§]For example, this gives rise to a smaller JT index, as illustrated in Example 5

[§]More precisely, the third method is expected to be better in the case of light tails claims and heavy traffic, while the second is expected to be better in the case of heavy tails claims and light traffic. The fact that both methods are better sometimes is illustrated in Theorem 4, from the point of view of yielding a smaller JT index. Identifying the domains within which methods two and three are preferable in "boundary cases" is not an easy task. Note also that a mixture of the two has been also proposed [Sak04].

$N = \infty$). The first case is applicable to the classic, and the second to the perturbed Cramér Lundberg model. As mentioned, the motivation of (16) is that the classic De Vylder approximation, is precisely the one point Padé approximation of $\Psi^*(s)$, around 0, of orders $(n-1, n)$, with $n = 1$.

A second class of Padé approximations we experiment with is:

$$(\psi^*)^{(n)}(s) = \mathcal{P}ade_{((m,n,m_1,n_1))}(\psi^*(s)) = \mathcal{P}ade_{((m,n,m_1,n_1))}([\frac{p}{\kappa(s)/s}]_N) \quad (17)$$

where $\mathcal{P}ade_{((m,n,m_1,n_1))}$ denotes a two point Padé approximation based on the power series around 0 and ∞. These are indispensable when dealing with the perturbed model.

Example 2. *The simplest Ramsay-type approximation is the* **one moment Renyi exponential approximation of the equilibrium density** *(which may also be viewed as a Padé* $(0,1)$ *approximation of the aggregate loss density, which imposes also the correct limiting behavior* $\lim_{s\to\infty} s\Psi^*(s) = \Psi(0) = \rho = 1/(1+\theta)$ *of the Laplace transform at* ∞*). This amounts to looking for an approximation of the form*

$$\Psi^*(s) \approx \frac{\rho}{s + b_0} \Leftrightarrow f_e^*(s) \approx \frac{b_0/(1-\rho)}{s + b_0/(1-\rho)},$$

where $f_e^*(s)$ *denotes the Laplace transform of the stationary excess density of the claims* (7) *(the two being related by the Pollaczek Khinchine formula* (6)*).*

Fitting one moment yields $\frac{1-\rho}{b_0} = \tilde{m}_1$*, where* $\tilde{m}_1 = \frac{m_2}{2m_1}$ *(the first moment of the stationary excess density), and*

$$\Psi^*(s) \approx \frac{\rho}{\frac{1-\rho}{\tilde{m}_1} + s}, \qquad \Leftrightarrow \qquad \Psi(x) \approx \rho e^{-x(1-\rho)/\tilde{m}_1}, \quad (18)$$

which is also correct in the limit $\rho \to 1$ *when* $\tilde{m}_1$ *exists see [Kal97], [Gra00, (31)].*

De Vylder's exponential approximation. The simplest Beekman-Bowers type approximation is De Vylder's (19), one of the most popular approximations for ruin probabilities, due to its simplicity and asymptotic correctness [Gra00], despite of its its being expressed in terms of only the first three moments of the claims.

Derivation of De Vylder as a two moments Padé approximation. De Vylder's approximation was obtained originally by equating the first three moments of the original Cramér Lundberg process with those of a new process with exponential claims, and with different arrival intensity and premium rate. We check now, using the moments (15) of the aggregate loss density, that De Vylder's formula coincides with the Padé $(0,1)$ approximation around 0 of the Laplace transform $\Psi^*(s)$ of the ruin probabilities. We start by expanding in power series the numerator and denominator of the Pollaczek-Khinchine formula (6):

$$\begin{aligned}
&\Psi^*(s) = \frac{\lambda m_2/2 - \lambda m_3 s/6 + ...}{p + \lambda m_2 s/2 - \lambda m_3 s^2/6 + ...} \approx \frac{a}{s+\alpha} \\
&\Leftrightarrow as(p + \lambda m_2 s/2 - \lambda m_3 s^2/6 + ...) \\
&\approx (s+\alpha)(\lambda m_2 s/2 - \lambda m_3 s^2/6 + ...) \\
&\Leftrightarrow \begin{cases} ap = \alpha\lambda m_2/2, \\ am_2/2 = m_2/2 - \alpha m_3/6 \end{cases} \quad \Leftrightarrow \begin{cases} a = \frac{3\lambda m_2^2}{3\lambda m_2^2 + 2pm_3}, \\ \alpha = \frac{6pm_2}{3\lambda m_2^2 + 2pm_3} \end{cases}
\end{aligned}$$

and so

$$\Psi^*_{DV}(s) = \frac{a}{s+\alpha} = \frac{\frac{3\lambda m_2^2}{3\lambda m_2^2+2pm_3}}{s + \frac{6pm_2}{3\lambda m_2^2+2pm_3}} = \frac{3\lambda m_2^2}{s(3\lambda m_2^2 + 2pm_3) + 6pm_2} \tag{19}$$

Question 3. *Identify domains within which methods two and three are "preferable" in some sense, in the simplest case of exponential approximations, i.e.* **compare the Renyi and De Vylder approximations**.

Example 3. The Cramér Lundberg process with Brownian perturbation $\sigma B(t)$. *The Laplace exponent is*

$$\kappa(s) = ps + \frac{\kappa_2}{2}s^2 + \sum_{k=3}^{\infty}(-s)^k \frac{\lambda m_k}{k!}, \tag{20}$$

where λm_k are the moments of the Lévy measure, and $\kappa_2 = \lambda m_2 + \sigma^2$.

Now we have two further unknowns of interest, the probability of "creeping ruin" and that of "ruin by jump". The respective Laplace transforms satisfy

$$\Psi^*_d(s) = \frac{\sigma^2/2}{\kappa'(0)}\psi^*(s) = \frac{\sigma^2/2}{p}\psi^*(s), \qquad \Psi^*_j(s) = \frac{1}{s} - \frac{\psi^*(s)}{s} - \frac{\sigma^2/2}{p}\psi^*(s) \tag{21}$$

The Padé $K = 0, L = 1$ approximations are unreasonable, since they cannot satisfy the boundary conditions $\Psi_d(0) = 1, \Psi_j(0) = 0$. To satisfy those as well as the equation

$$\Psi_d^*(s) = \frac{\sigma^2/2}{p}\psi^*(s) \Leftrightarrow 1 - s\Psi_j^*(s) = (s + \frac{2p}{\sigma^2})\Psi_d^*(s) \tag{22}$$

which follows from (21)*, we must use at least a $K = 1, L = 2$ approximation.*

Theorem 1. *Consider the exponential approximation*

$$\Psi_d(x) = \frac{a_d - s_1}{s_2 - s_1}e^{-s_1 x} + \frac{s_2 - a_d}{s_2 - s_1}e^{-s_2 x} \tag{23}$$

$$\Psi_j(x) = \frac{a_j}{s_2 - s_1}[e^{-s_1 x} - e^{-s_2 x}], \tag{24}$$

for the "creeping ruin" and "ruin by jump", which satisfy $\Psi_d(0) = 1, \Psi_j(0) = 0$. Then, by fitting the first two moments of the aggregate loss L, one is led to the following admissible approximation:

$$a_d = \frac{3m_2}{m_3},\ a_j = a_d\frac{\lambda m_2}{\sigma^2} = \frac{3\lambda m_2^2}{\sigma^2 m_3}. \tag{25}$$

and $-s_1$ and $-s_2$ the negative roots of $s^2 + (a_d + a_j + \frac{2p}{\sigma^2})s + a_d\frac{2p}{\sigma^2} = 0$ (whose discriminant is non-negative).

Proof: *$\Psi^*(s)$, $\Psi_d^*(s) = \frac{s+a_d}{s^2+b_1 s+b_0}$ are quotients of monic polynomials (to satisfy $\lim_{s\to\infty} s\Psi_d^*(s) = \Psi_d(0) = 1$), with three free coefficients, but the second condition in* (22) *imposes one more condition $b_0 = \frac{2p}{\sigma^2}a_d$ (so that $\Psi_j^*(s) = \frac{a_j}{s^2+b_1 s+a_d\frac{2p}{\sigma^2}}$), leaving only two free coefficients. Finally, fitting the first two coefficients around* 0 *of*

$$\psi^*(s) = 1 - s\Psi^*(s) = \frac{2p}{\sigma^2}\Psi_d^*(s) \approx \frac{2p}{\sigma^2}\frac{s + a_d}{s^2 + (a_d + a_j + \frac{2p}{\sigma^2})s + a_d\frac{2p}{\sigma^2}} \tag{26}$$

yields:

$$\begin{cases} \frac{\sigma^2}{2p}(1 + \frac{a_j}{a_d}) = \frac{\lambda m_2 + \sigma^2}{2p} \Leftrightarrow \frac{a_j}{a_d} = \frac{\lambda m_2}{\sigma^2} \\ \frac{\sigma^2}{2p}\frac{a_j}{a_d^2} = \frac{\lambda m_3}{6p} \end{cases}$$

with solution (25)*. Then, assuming $s_1 < s_2$, Laplace inversion yields* (23)*.*

Moreover, both Ψ_j and Ψ_d are admissible. Indeed, this is obvious for Ψ_j, since its initial value and its dominant coefficient $\frac{a_j}{s_2-s_1}$ are non-negative.

The same is true for Ψ_d; indeed, we may check that its dominant coefficient is non-negative, i.e. that $s_1 \leq a_d$, by noting that $s^2 + (a_d + a_j + 2p/\sigma^2)s + a_d 2p/\sigma^2$ is negative at $s = -a_d$ (since $(a_d)^2 + (a_d + a_j + 2p/\sigma^2)(-a_d) + a_d 2p/\sigma^2 = -a_d a_j < 0$, with a_d and a_j being positive). Therefore, $s_1 < a_d < s_2$.

Remark 11. *It is easy to check that this approximation is exact for exponential claims. Indeed, in that case the density transform in the parametrization* (26) *is:*

$$\psi^*(s) = \frac{p}{\kappa(s)} = \frac{1}{1 + \frac{\sigma^2}{2p}s + \frac{\lambda}{p}s\frac{m_1}{s+m_1^{-1}}} = \frac{s + m_1^{-1}}{(s + m_1^{-1})(1 + \frac{\sigma^2}{2p}s) + \frac{\lambda}{p}sm_1}$$
$$= \frac{s + m_1^{-1}}{\frac{\sigma^2}{2p}s^2 + s(\frac{\lambda}{p}m_1 + \frac{\sigma^2}{2p}m_1^{-1} + 1) + m_1^{-1}} = \frac{2p}{\sigma^2}\frac{s + m_1^{-1}}{s^2 + s(\frac{2\lambda m_1}{\sigma^2} + m_1^{-1} + \frac{2p}{\sigma^2}) + m_1^{-1}\frac{2p}{\sigma^2}}$$

with $a_d = 1/m_1, a_j = 2\lambda m_1/\sigma^2$, and it is easy to check that this coincides with our approximation (25)*, in the case of exponential claims.*

Remark 12. *Another admissible approximation exact in the exponential case, but fitting now only one moment is $a_j = 2\lambda m_1/\sigma^2$, $a_d = 1/\tilde{m}_1 = \frac{2m_1}{m_2}$ (with similar admissibility proof).*

5. Johnson-Taaffe approximations

This elegant moment fitting method approximates by **common order Erlang mixtures approximations** [§] :

$$\boxed{f(x) \sim \sum_{i=1}^{K} w_i e_n(\rho_i, x) := \int_0^\infty e_n(\rho, x)\nu_K(d\rho)} \tag{27}$$

[§]One hint which may explain this choice comes from discretizing the well-behaved Post-Widder Laplace inversion method [Jag78], which leads to Erlang mixtures of common order. Note also that pure Erlang distributions are the unique extremal points minimizing the coefficient of variability, within phase-type distributions representable at order n, and that this is a dense subclass of phase-type distributions.

where $e_n(\rho, x)$ is the Erlang density. Finding a discrete measure $\nu_K(d\rho)$ with positive coefficients w_i requires solving a classic De Prony system with respect to **Erlang reduced moments**

$$\mu_1 = \frac{m_1}{n}, \mu_2 = \frac{m_1}{n(n+1)}, ..., \mu_k = \frac{m_k}{n_{(k)}}, \quad n_{(k)} = n(n+1)...(n+k-1)$$

obtained dividing by the moments $n_{(k)}$ of the scaled Erlang density of shape n.

Example 4. *When $n = 1$, one finds the* **factorially reduced moments**

$$\mu_1 = \frac{m_1}{1}, \mu_2 = \frac{m_1}{2!}, \mu_3 = \frac{m_1}{3!}, ...$$

which intervene in constructing GHE approximations (which have positive weights, if the moments m_k are far enough from the boundaries of the Stieltjes space of moments $\frac{m_2}{m_1^1} \geq 1, \frac{m_3 m_1}{m_2^1} \geq 1$).

It turns out that Erlang reducing leads to "positive moment sequences" μ_k satisfying $\frac{\mu_2}{\mu_1^1} \geq 1, \frac{\mu_3 \mu_1}{\mu_2^1} \geq 1, ...$ (i.e. having positive Hankel matrices of moments up to any desired degree), for n big enough. Then, the classic Stieltjes method of creating discrete moment fitting measures with **positive weights** may be applied.

Theorem 2. Johnson-Taaffe three moments approximation. *Let $\boldsymbol{m} = (m_1, m_2, m_3)$ denote the first three moments of a nonnegative r.v. Let*

$$\begin{aligned} n^* &= \lceil \max\{(m_0 m_2/m_1^2 - 1)^{-1}, (m_1 m_3/m_2^2 - 1)^{-1} - 1\} \rceil \\ &= \lceil \max\{\frac{1}{\hat{m}_2 - 1}, \frac{2\hat{m}_2 - \hat{m}_3}{\hat{m}_3 - \hat{m}_2}\} \rceil \end{aligned}$$

denote the **explicit "JT index of degree"** 3 *giving the smallest "JT sector" – see [BHT05, Fig 3-7] – containing our target moments. Let $\boldsymbol{\mu} = (\mu_i, i = 1, ...3)$ denote the Erlang reduced moments of order n^*, and let*

$$\boxed{b_0 = \mu_2 - \mu_1^2, b_1 = \mu_3 - \mu_1 \mu_2, b_2 = \mu_3 \mu_1 - \mu_2^2}$$

denote the coefficients of the denominator of the classic second order Padé approximant, and let

$$s_{1,2} = \frac{2b_0}{b_1 \mp \sqrt{b_1^2 - 4b_0 b_2}}, \ \omega_1 = \frac{1 - s_1 \mu_1}{s_2 - s_1}, \omega_2 = 1 - \omega_1$$

denote its roots and (non-negative) partial fractions coefficients.

Then, the Erlang reduced moments satisfy $\boldsymbol{\mu} \in \{b_0 > 0, b_2 > 0\}$*, and*

$$f(x) = \sum_{i=1}^{2} \omega_i e_{n^*}(s_i, x),$$

is a nonnegative, decreasing "common order Erlang n^**" density fitting our three moments.*

Using this result, we can "fix" Ramsay's approximation by sums of two exponentials [Ram92] to make it work for any valid three moments, at the price of using higher (not minimal) order approximations – see Theorem 3.

Remark 13. In principle, $2k+1$ **moments fitting Johnson-Taaffe Padé approximations** *could be obtained by determining numerically higher order JT indices ensuring the positivity of corresponding Hankel determinants. The Hankel determinants that must be taken into consideration are:*

$$\begin{aligned}\mathcal{M}_k(\mathcal{P}) = \{&m_4(m_2 - m_1^2) + m_2(m_3 m_1 - m_2^2) + m_3(m_2 m_1 - m_3^2) \geq 0,\\ &m_5 m_4 \left(m_1 m_3 - m_2^2\right) + m_4 \left(-m_3^3 + 2m_2 m_4 m_3 - m_1 m_4^2\right) \geq 0, ..., \}\end{aligned} \quad (28)$$

The inequalities above define implicitly the **JT index of degree** 5*.*

Theorem 3. Ramsay updated *Consider the ruin problem for the classic (nonperturbed) Cramér-Lundberg model* (3)*.*

1. *Imposing the correct limiting behavior* $\lim_{s\to\infty} s\Psi^*(s) = \Psi(0) = \rho = 1/(1+\theta)$ *at* ∞*, the Padé* $(1,2)$ *approximations for the equilibrium density and ruin transforms are:*

$$\begin{aligned} f_e^*(s) &\approx \frac{N}{D} := \frac{b_0 + a_1 s}{b_0 + b_1 s + b_2 s^2},\\ \Psi^*(s) &\approx \rho \frac{D-N}{D-\rho N} := \rho \frac{b_2 s + b_1 - a_1}{b_2 s^2 + \tilde{b}_1 s + \tilde{b}_0}, \end{aligned}$$

$$\begin{aligned} &b_0 = \tilde{\mu}_2 - \tilde{\mu}_1^2, \qquad b_1 = \tilde{\mu}_3 - \tilde{\mu}_2\tilde{\mu}_1, \qquad b_2 = \tilde{\mu}_1\tilde{\mu}_3 - \tilde{\mu}_2^2\\ &a_1 = b_1 - \tilde{\mu}_1 b_0 = \tilde{\mu}_3 - 2\tilde{\mu}_1\tilde{\mu}_2 + \tilde{\mu}_1^3\\ &\tilde{b}_1 s = b_1 - \rho a_1, \qquad \tilde{b}_0 = (1-\rho) b_0, \end{aligned}$$

where $\tilde{\mu}_2$ *are* **factorially reduced equilibrium moments***.*

2. *For the approximate inverse transform $f_e(x)$ to be hyperexponential (in particular, nonnegative and completely monotone), it is sufficient that $b_0 > 0, b_2 > 0$.*
3. *Higher order moments fitting is possible by Erlang approximations of order n^* big enough, where for three moments $n^* = \lceil \max\{\frac{m_1^2}{m_2 - m_1^2}, \frac{m_2^2}{m_1 m_3 - m_2^2} - 1\}\rceil$.*

6. Comparison between the JT indices of Ramsay and Beekman-Bowers type approximations

In this section we present a simple comparison between these two approaches, based on comparing their Johnson-Taaffe indices.

The moments of the aggregate loss L satisfy

$$\widehat{l}_2 = \frac{l_2}{l_1^2} = 2 + \theta \frac{\tilde{m}_2}{\tilde{m}_1^2} = 2 + \theta \widehat{\tilde{m}}_2 \geq 2 + \theta > 2, \tag{29}$$

$$\widehat{l}_3 = \frac{l_3}{l_1 l_2} = 3 + \frac{\frac{\tilde{m}_3}{\theta} + \frac{\tilde{m}_1 \tilde{m}_2}{\theta^2}}{\frac{\tilde{m}_2 \tilde{m}_1}{\theta^2} + 2(\frac{\tilde{m}_1}{\theta})^3} = 3 + \frac{\theta \widehat{\tilde{m}}_2 (1 + \theta \widehat{\tilde{m}}_3)}{2 + \theta \widehat{\tilde{m}}_2}$$

The second moment satisfies the necessary inequality $\widehat{l}_2 \geq \frac{n+1}{n}$ [BHT05] already with $n = 1$. We only need to investigate the necessary inequality for the third moment, which is:

$$n \geq \boxed{J(L) := \frac{2\widehat{l}_2 - \widehat{l}_3}{\widehat{l}_3 - \widehat{l}_2}} = \frac{\left(\widehat{\tilde{m}}_2 \widehat{\tilde{m}}_3 - 2\widehat{\tilde{m}}_2^{\,2}\right)\theta^2 - 4\widehat{\tilde{m}}_2 \theta - 2}{\theta^2\left(\widehat{\tilde{m}}_2^{\,2} - \widehat{\tilde{m}}_2 \widehat{\tilde{m}}_3\right) - 2} \Leftrightarrow$$

$$\frac{n+2}{n+1} \leq \boxed{\eta(L) := \frac{\widehat{l}_3}{\widehat{l}_2}} = \frac{\theta \widehat{\tilde{m}}_2 \left(\theta \widehat{\tilde{m}}_3 + 4\right) + 6}{\left(\theta \widehat{\tilde{m}}_2 + 2\right)^2}$$

Here, $J(L) = \frac{2\widehat{l}_2 - \widehat{l}_3}{\widehat{l}_3 - \widehat{l}_2}$ is a "partial JT index", based on the third moment admissibility condition, and the "normalized moment" $\eta(L) = \frac{\widehat{l}_3}{\widehat{l}_2}$ is a monotone transformation (since $J = \frac{2-\eta}{\eta - 1}$ is a decreasing function in the relevant range $\eta \in (1, \infty)$), which has been already used in the literature.

A rough indication of the performance of the Beekman-Bowers and Ramsay methods will be obtained now by checking which of $\eta(L)$, $\eta(L_i)$ is higher.

Theorem 4. *a) The partial J index of the aggregate loss is strictly smaller than $J(L_i) = \frac{2\widehat{\tilde{m}}_2 - \widehat{\tilde{m}}_3}{\widehat{\tilde{m}}_3 - \widehat{\tilde{m}}_2}$ iff*

$$\theta\widehat{\tilde{m}}_2 < \frac{3/2\widehat{\tilde{m}}_2 - \widehat{\tilde{m}}_3}{\widehat{\tilde{m}}_3 - \widehat{\tilde{m}}_2} = \frac{3/2 - \eta(L_i)}{\eta(L_i) - 1}$$

In particular, if $\eta(L_i) = \frac{\widehat{\tilde{m}}_3}{\widehat{\tilde{m}}_2} \geq \frac{3}{2}$, in which case $\eta(L) < \frac{3}{2}$ holds as well, a two terms exponential mixture distribution matching the first three moments exists only for the equilibrium ladders L_i.

b) More generally, let n denote the unique integer such that $\frac{n+2}{n+1} \leq \eta(L_i) \leq \frac{n+1}{n}, n \geq 2$, is satisfied. If furthermore $x_1 < \theta\widehat{\tilde{m}}_2 < x_2$, where $x_{1,2} = \frac{2\mp\sqrt{2}\sqrt{n^2+n-a(n^2-1)}}{(n+1)a-(n+2)}$, then a n'th order Erlang mixture distribution matching the first three moments exists only for the equilibrium ladders L_i.

c) For any θ, there exists a mixture of Erlang n distribution matching the first three moments of L if

$$n \geq \frac{\widehat{\tilde{m}}_3}{\widehat{\tilde{m}}_3 - \widehat{\tilde{m}}_2}.$$

Proof: a) Consider the unimodal function

$$\eta(a, x) = \frac{\widehat{l}_3}{\widehat{l}_2} = \frac{ax^2 + 4x + 6}{(x+2)^2},$$

where $a = \frac{\widehat{\tilde{m}}_3}{\widehat{\tilde{m}}_2}, x = \theta\widehat{\tilde{m}}_2$, which takes values

$$\frac{3a-2}{2a-1} \leq \eta(a, x) \leq \max[\eta(a, \infty) = a, \eta(a, 0) = 3/2]$$

(the lower minimal value is achieved for $x^* = \frac{1}{a-1} \Leftrightarrow \theta^* = \frac{1}{\widehat{\tilde{m}}_3 - \widehat{\tilde{m}}_2}$).

b) Fitting L is possible when

$$\frac{ax^2 + 4x + 6}{(x+2)^2} > \frac{n+2}{n+1} \Leftrightarrow ((n+1)a - (n+2))x^2 - 4x + 2(n-1) > 0.$$

In the prescribed range of a, the discriminant $\Delta = 2(n^2 + n - a(n^2 - 1))$ is always positive, and the inequality holds when $x < x_1$ or when $x > x_2$, where $x_{1,2} = \frac{2\mp\sqrt{2}\sqrt{n^2+n-a(n^2-1)}}{(n+1)a-(n+2)}$.

c) Follows by minimizing $\eta(\theta)$ in θ, which yields $\theta^* = \frac{1}{\widehat{\tilde{m}}_3 - \widehat{\tilde{m}}_2}$ and $\eta(\theta^*) = \frac{3\widehat{\tilde{m}}_3 - 2\widehat{\tilde{m}}_2}{2\widehat{\tilde{m}}_3 - \widehat{\tilde{m}}_2}, J(\theta^*) = \frac{\widehat{\tilde{m}}_3}{\widehat{\tilde{m}}_3 - \widehat{\tilde{m}}_2}$.

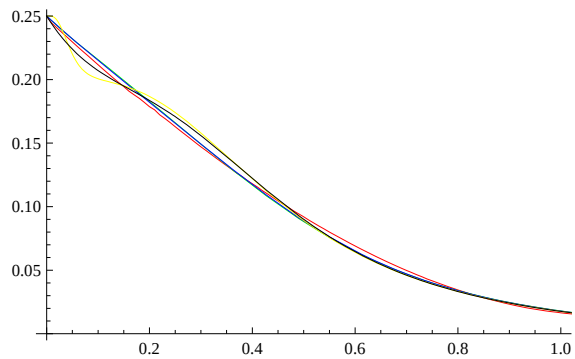

Figure 1: Ruin probabilities for uniform claims. Red line: original calculated by numerical inverse Laplace; green line: by 3 moments [JT89] on the uniform itself; blue line: by 3 moments [BHT05] on the uniform itself; yellow line: by 3 moments [JT89] on the equilibrium of the uniform; black line: by 3 moments [BHT05] on the equilibrium of the uniform.

Example 5. *Consider a three moments fitting example: the* $U[0,1]$ *rv, with moments* $m_1 = \frac{1}{2}, m_2 = \frac{1}{3}, m_3 = \frac{1}{4}, m_4 = \frac{1}{5}$. *The Johnson-Taaffe index is* 7*, and so this procedure yields* 14 *phases, while the Bobbio, Horvath and Telek [BHT05] and He-Zhang methods [HZ06] yield only* 9 *phases.*

For the equilibrium distribution, the moments are $\tilde{m}_1 = \frac{1}{3}, \tilde{m}_2 = \frac{1}{6}, \tilde{m}_3 = \frac{1}{10}$, *the normalized moments are*

$$\widehat{\tilde{m}}_2 = \frac{\tilde{m}_2}{\tilde{m}_1^2} = \frac{3}{2}, \widehat{\tilde{m}}_3 = \frac{\tilde{m}_3}{\tilde{m}_2 \tilde{m}_1} = \frac{9}{5}$$

and the Johnson-Taaffe index is 4. *The Bobbio, Horvath and Telek method yields the minimal order of* 3 *phases.*

Finally consider the JT approach for a "Beekman-Bowers" approximation. Computing the partial JT index, we find

$$\frac{4\left(9\theta^2 + 30\theta + 10\right)}{9\theta^2 + 40}.$$

By Theorem 4 a), this is strictly less than the 4 *required for the Ramsay approach when* $\theta \leq 1$. *Taking into account the integer part, we find out that* $\theta \leq 1.5$ *still yields a J index less or equal to* 4.

7. Admissible two point Padé approximations

We consider now briefly including further information beyond moments, for example by using two-point Padé approximations (of special interest in the case of Levy processes with infinite variation paths – see Example). Ensuring admissibility in this case is however an open problem.

We may formally expand the ruin transform at infinity as well:

$$\Psi^*(s) = \int_0^\infty e^{-sx} \sum_{k=0}^\infty \Psi^{(k)}(0) \frac{x^k}{k!} = \sum_{k=0}^\infty \Psi^{(k)}(0) s^{-k-1} \tag{30}$$

Here,

$$\Psi(0) = \begin{cases} 1 & \text{if } \sigma > 0 \\ \frac{\eta_1}{c} = \rho & \text{if } \sigma = 0 \end{cases}$$

is well known, and the derivatives at 0:

$$\begin{aligned} \Psi'(0) &= -\frac{\eta_0}{c}(1 - \frac{\eta_1}{c}) = -\frac{p\lambda}{c^2} = -\frac{\rho(1-\rho)}{m_1}, \\ \Psi''(0) &= -\Psi'(0)(f(0) - \frac{\eta_0}{c}), \\ \Psi^{(3)}(0) &= -\Psi'(0)\left(f'(0) + 2\frac{\eta_0}{c} f(0) - (\frac{\eta_0}{c})^2\right), \\ \Psi^{(4)}(0) &= -\Psi'(0)\left(f''(0) + 2\frac{\eta_0}{c} f'(0) - \frac{\eta_0}{c} f(0)^2 + 3(\frac{\eta_0}{c})^2 f(0) - (\frac{\eta_0}{c})^3\right) \end{aligned} \tag{31}$$

may be obtained recursively, by differentiating the integro-differential equation for $\Psi(x)$.

Theorem 5. Two-point Padé-Ramsay approximation *Imposing both the correct limiting behavior* $\lim_{s\to\infty} s\Psi^*(s) = \Psi(0) = \rho = 1/(1+\theta)$ *at* ∞*, and the first derivative at* ∞ $\Psi'(0) = -\frac{\rho(1-\rho)}{m_1}$ *leads to the two-point Padé* $(1,2)$ *approximation:*

$$f_e^*(s) \approx \frac{b_0 + a_1 s}{b_0 + b_1 s + b_2 s^2}, \quad \Psi^*(s) \approx \rho \frac{b_2 s + b_1 - a_1}{\tilde{b}_0 + \tilde{b}_1 s + b_2 s^2}, \tag{32}$$

where $a_1 = \frac{b_2}{m_1}$ *and the coefficients* $b_i, \tilde{b}_i$*, obtained fitting the first two equilibrium/aggregate loss moments, are:*

$$\begin{aligned} &b_2 = (2m_1 m_3 - 3m_2^2)/6, \; b_1 = (m_3 - 3m_1 m_2)/3, \; b_0 = m_2 - 2m_1^2, \\ &\tilde{b}_1 = b_1 - \rho a_1, \qquad \tilde{b}_0 = (1-\rho) b_0. \end{aligned}$$

Proof: The theorem follows from the identity

$$s\Psi^*(s) = \rho\frac{1-f_e^*(s)}{1-\rho f_e^*(s)} = \rho\frac{b_2s^2+(b_1-a_1)s}{b_2s^2+(b_1-\rho a_1)s+b_0(1-\rho)} \tag{33}$$

by taking into account that we are looking for an approximation of the form

$$\Psi^*(s) \approx \rho\frac{b_2s+b_1-\frac{b_2}{m_1}}{b_2s^2+\tilde{b}_1s+\tilde{b}_0} := \rho\frac{\tilde{N}}{\tilde{D}},$$

which will satisfy $\lim_{s\to\infty} s\frac{\tilde{N}}{\tilde{D}} = 1, \lim_{s\to\infty} s(s\frac{\tilde{N}}{\tilde{D}}-1) = \Psi'(0)/\rho = -\frac{1-\rho}{m_1}$.

8. Numerical results

8.1. Without perturbation

In the following we illustrate the application of the approximations provided in Theorem 3 and 5.

8.1.1. Mixed exponential claim distribution

With mixed exponential claim distribution

$$f(x) = \frac{315e^{-5x}}{128} + \frac{7e^{-4x}}{8} + \frac{27e^{-3x}}{64} + \frac{3e^{-2x}}{16} + \frac{7e^{-x}}{128}$$

and $\lambda = 1, c = 2/5$ the ruin probability is

$$\Psi(x) = \frac{245e^{-9x/2}}{32768} + \frac{135e^{-7x/2}}{8192} + \frac{567e^{-5x/2}}{16384} + \frac{735e^{-3x/2}}{8192} + \frac{19845e^{-x/2}}{32768}$$

We approximated the ruin probabilities by the Renyi and De Vylder's first order exponential approximation given in (18) and (19), and by the approximations provided in Theorem 3 and 5. In Figure 2 we show the ruin probabilities with the different approximations and in Figure 3 the relative error with respect to the exact solution. As expected, the approximation in Theorem 5 is better near $x = 0$, and that may be exploited to obtain a better approximation by switching between the formulas in Theorem 5, Theorem 3 when they cross, starting with the first.

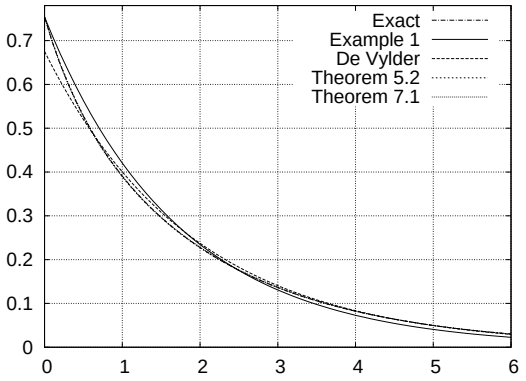

Figure 2: Ruin probabilities with mixed exponential claims.

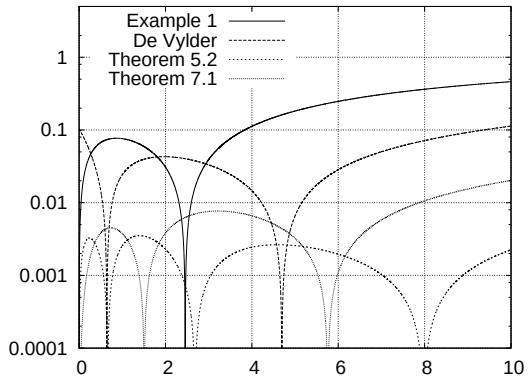

Figure 3: Absolute relative error of ruin probabilities with mixed exponential claims.

x	exact $\Psi(x)$	Renyi	DeVylder	Theorem 3	Theorem 5
0.	0.909091	0.909091	0.882867	0.909091	0.909091
300.	0.521143	0.529743	0.522539	0.521107	0.522526
600.	0.308668	0.30869	0.309273	0.308713	0.309268
900.	0.182866	0.179879	0.183048	0.182888	0.183047
1200.	0.108338	0.104818	0.10834	0.108347	0.10834
1500.	0.0641841	0.0610794	0.0641226	0.0641869	0.0641233
1800.	0.0380254	0.035592	0.037952	0.0380257	0.0379527
2100.	0.0225279	0.0207401	0.0224625	0.0225272	0.0224631
2400.	0.0133465	0.0120856	0.0132948	0.0133456	0.0132953
2700.	0.00790706	0.00704247	0.00786872	0.0079062	0.00786908
3000.	0.00468448	0.00410377	0.00465722	0.0046838	0.00465748

Table 1: Ruin probabilities with Gamma claims ($\alpha = 0.01, \beta = 100$) and its approximations

8.1.2. Gamma distribution with $\alpha = 0.01, \beta = 100$

Let us consider now the Gamma claim distribution

$$\frac{e^{-\frac{x}{\beta}} x^{\alpha-1} \beta^{-\alpha}}{\Gamma(\alpha)}$$

with parameters $\alpha = 0.01, \beta = 100$ and with claim arrival intensity $\lambda = 1$ and loading factor $\theta = 0.1$, which appeared frequently in the literature ([Ram92], [Gra00]). As for the previous example, we approximated the ruin probabilities by the formulas given in Renyi, (19) and Theorem 3 and 5. Ruin probabilities themselves are given in Table 1 while in Table 2 we provide absolute relative errors.

8.1.3. Gamma distribution with $\alpha = 2.5, \beta = 1$

Next we consider Gamma distributed claims with $\alpha = 2.5, \beta = 1$ and $\lambda = 2/5, c = \frac{4}{5}\left(-1 + 4\sqrt{2}\right)$. This case is interesting as it has been shown in [ACH11] that direct, moment based Padé approximation of the claim distribution does not result in valid distributions. The approximations presented in this paper leads instead to valid ruin probabilities. Ruin probabilities themselves are given in Table 3 while in Table 4 we provide absolute relative errors.

8.2. With perturbation

In this section we illustrate the application of the two approximations given in Theorem 1 and Remark 12, respectively.

x	Renyi	DeVylder	Theorem 3	Theorem 5
0	0	0.0288462	0	0
300	0.0165011	0.00267814	0.0000688428	0.00265325
600	0.0000714085	0.0019599	0.000146799	0.00194387
900	0.0163373	0.000993297	0.000119325	0.000986105
1200	0.0324864	0.0000177571	0.0000819863	0.000019394
1500	0.0483709	0.000957418	0.0000440626	0.00094697
1800	0.0639946	0.00193166	6.12674×10^{-6}	0.00191241
2100	0.0793618	0.00290493	0.000031798	0.00287691
2400	0.0944767	0.00387725	0.0000697035	0.00384047
2700	0.109343	0.00484863	0.000107635	0.00480311
3000	0.123966	0.00581907	0.000145554	0.00576482

Table 2: Relative error of approximate ruin probabilities with Gamma claims ($\alpha = 0.01, \beta = 100$)

x	exact $\Psi(x)$	Renyi	DeVylder	Theorem 3	Theorem 5
0.	0.268422	0.268422	0.299749	0.268422	0.268422
0.5	0.22854	0.217791	0.237348	0.22894	0.228126
1.	0.189678	0.176711	0.187938	0.189655	0.189069
1.5	0.154441	0.143379	0.148813	0.154172	0.154016
2.	0.124037	0.116334	0.117834	0.123743	0.123926
2.5	0.0986589	0.0943911	0.0933036	0.0984496	0.0988216
3.	0.0779451	0.0765868	0.07388	0.0778418	0.0782763
3.5	0.0612929	0.0621407	0.0584999	0.0612758	0.0616894
4.	0.0480435	0.0504196	0.0463215	0.0480817	0.04843
4.5	0.0375759	0.0409093	0.0366785	0.0376414	0.0379079
5.	0.0293456	0.0331929	0.0290429	0.0294185	0.0296037

Table 3: Ruin probabilities with Gamma claims ($\alpha = 2.5, \beta = 1$) and its approximations

x	Renyi	DeVylder	Theorem 3	Theorem 5
0.	0.	0.116709	0.	0.
0.5	0.0470337	0.0385392	0.00175079	0.0018135
1.	0.0683674	0.00917815	0.000125885	0.00321361
1.5	0.0716262	0.0364389	0.00174365	0.00275474
2.	0.062096	0.0500072	0.00236637	0.000891164
2.5	0.0432581	0.0542808	0.00212105	0.00164952
3.	0.0174272	0.0521542	0.00132515	0.00424903
3.5	0.0138333	0.045568	0.000278168	0.00647046
4.	0.0494556	0.0358424	0.000794076	0.00804433
4.5	0.0887094	0.023884	0.00174297	0.00883326
5.	0.1311	0.0103171	0.0024838	0.00879521

Table 4: Relative error of approximate ruin probabilities with Gamma claims ($\alpha = 2.5, \beta = 1$)

8.2.1. Mixed exponential claim distribution

As without perturbation, in this case the exact ruin probabilities can be calculated by symbolic inversion of the Laplace transform. We applied the approximations given in Theorem 1 and Remark 12 for three values of σ, namely, 0.1, 0.5 and 2. Figures 4-6 show the exact and approximate ruin probabilities and we depicted the two components of the ruin probability (by diffusion and by jump) as well. The two approximations result in distinguishable curves only in the case $\sigma = 0.1$. The relative errors for the three values of σ are provided in Figure 7-9. The error is smaller for larger values of σ and for most values of x the better approximation of $\Psi(x)$ is by the approach of Theorem 1.

Acknowledgement: We thank Jiandong Ren for useful comments.

[ACH11] F. Avram, DF Chedom, and A. Horvath. On moments based Padé approximations of ruin probabilities. *Journal of computational and applied mathematics*, 2011.

[Asm03] S. Asmussen. *Applied probability and queues*, volume 51. Springer Verlag, 2003.

[BFT08] N.G. Bean, M. Fackrell, and P. Taylor. Characterization of matrix-exponential distributions. *Stochastic Models*, 24(3):339–363, 2008.

[BHT05] A. Bobbio, A. Horváth, and M. Telek. Matching three moments

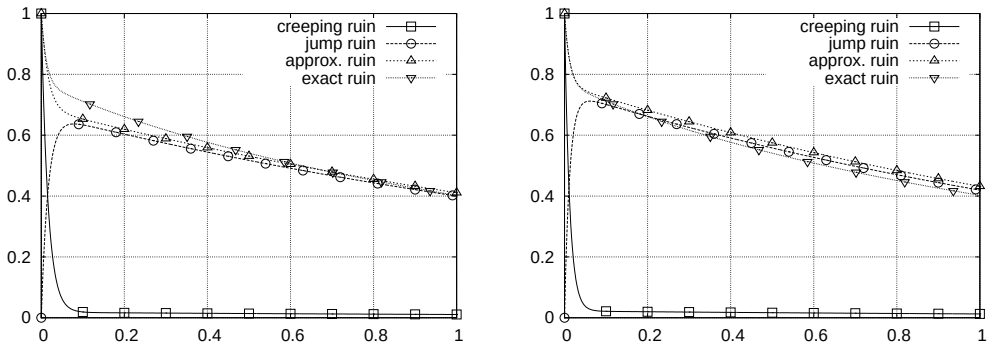

Figure 4: Ruin probabilities with mixed exponential claim distribution with $\sigma = 0.1$, with the approximation given in Theorem 1 (left) and the one given in Remark 12 (right)

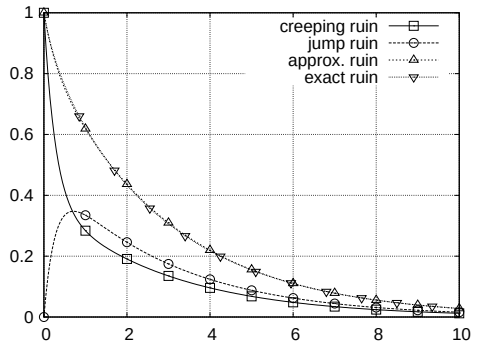

Figure 5: Ruin probabilities with mixed exponential claim distribution with $\sigma = 0.5$, with the approximations of Theorem 1 and Remark 12 (the two approximations look the same)

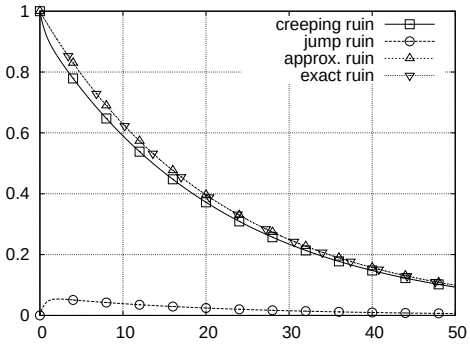

Figure 6: Ruin probabilities with mixed exponential claim distribution with $\sigma = 2$, with the approximations of Theorem 1 and Remark 12 (the two approximations look the same)

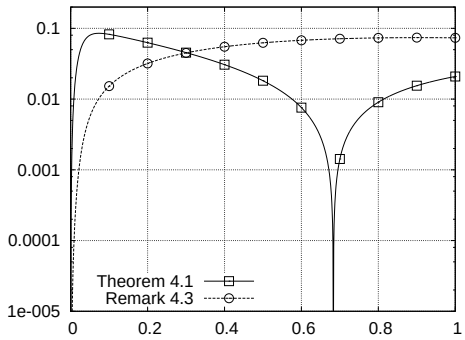

Figure 7: Absolute relative error of the approximations given in Theorem 1 and Remark 12 with mixed exponential claims and $\sigma = 0.1$.

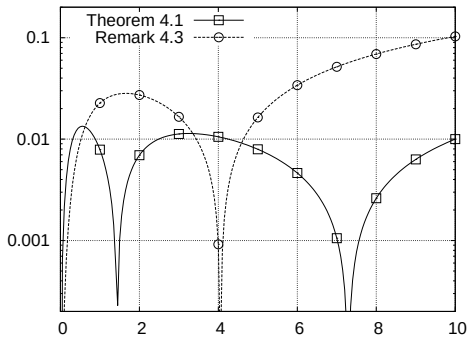

Figure 8: Absolute relative error of the approximations given in Theorem 1 and Remark 12 with mixed exponential claims and $\sigma = 0.5$.

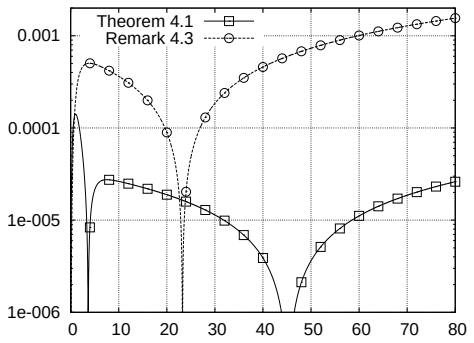

Figure 9: Absolute relative error of the approximations given in Theorem 1 and Remark 12 with mixed exponential claims and $\sigma = 2$.

with minimal acyclic phase type distributions. *Stochastic models*, 21(2-3):303–326, 2005.

[BM05] G. Beylkin and L. Monzón. On approximation of functions by exponential sums. *Applied and Computational Harmonic Analysis*, 19(1):17–48, 2005.

[Coh07] A.M. Cohen. *Numerical methods for Laplace transform inversion*. Springer Verlag, 2007.

[DG91] F. Dufresne and H.U. Gerber. Risk theory for the compound poisson process that is perturbed by diffusion. *Insurance: mathematics and economics*, 10(1):51–59, 1991.

[DL82] M. Dehon and G. Latouche. A geometric interpretation of the relations between the exponential and generalized erlang distributions. *Advances in Applied Probability*, pages 885–897, 1982.

[Fac03] M.W. Fackrell. *Characterization of matrix-exponential distributions*. PhD thesis, The University of Adelaide, 2003.

[Fay02] L. Faybusovich. Self-concordant barriers for cones generated by chebyshev systems. *SIAM Journal on Optimization*, 12(3):770–781, 2002.

[Gra00] J. Grandell. Simple approximations of ruin probabilities. *Insurance: Mathematics and Economics*, 26(2):157–173, 2000.

[HWR92] C.M. Harris, G.M. William, and F.B. Robert. A note on generalized hyperexponential distributions. *Stochastic Models*, 8(1):179–191, 1992.

[HZ06] Q.M. He and H. Zhang. Spectral polynomial algorithms for computing bi-diagonal representations for phase type distributions and matrix-exponential distributions. *Stochastic Models*, 22(2):289–317, 2006.

[Jag78] DL Jagerman. An inversion technique for the laplace transform with application to approximation. *Bell System Tech. J*, 57(3):669–710, 1978.

[JT89] M.A. Johnson and M.R. Taaffe. Matching moments to phase distributions: Mixtures of erlang distributions of common order. *Stochastic Models*, 5(4):711–743, 1989.

[Kal97] V.V. Kalashnikov. *Geometric sums: bounds for rare events with applications: risk analysis, reliability, queueing*, volume 413. Springer, 1997.

[Kel11] O. Kella. The class of distributions associated with the generalized pollaczek-khinchine formula. *Arxiv preprint arXiv:1111.7099*, 2011.

[Kyp06] A.E. Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications*. Springer Verlag, 2006.

[O'C93] C.A. O'Cinneide. Triangular order of triangular phase-type distributions*. *Stochastic Models*, 9(4):507–529, 1993.

[Ram92] C.M. Ramsay. A practical algorithm for approximating the probability of ruin. *Transactions of the Society of Actuaries*, 44:443–461, 1992.

[Sak04] T. Sakurai. Approximating m/g/1 waiting time tail probabilities. 2004.

[See12] A. Seeger. Epigraphical cones ii. *J. Convex Anal*, 19(1), 2012.

[Wil98] G.E. Willmot. On a class of approximations for ruin and waiting time probabilities. *Operations research letters*, 22(1):27–32, 1998.

[WL11] G.E. Willmot and X.S. Lin. Risk modelling with the mixed erlang distribution. *Applied Stochastic Models in Business and Industry*, 27(1):2–16, 2011.