Technical Report No. 1/08, January 2008

UNIFORM ERROR BOUNDS IN CONTINUOUS APPROXIMATIONS OF NONNEGATIVE RANDOM VARIABLES USING LAPLACE TRANSFORMS

C. Sangüesa

Uniform error bounds in continuous approximations of nonnegative random variables using Laplace Transforms.

C. Sangüesa

Departamento de Métodos Estadísticos. Universidad de Zaragoza.

Zaragoza. Spain.

e-mail: csangues@unizar.es

Abstract

In this work we deal with approximations for distribution functions of nonnegative random variables. More specifically, we construct continuous approximants using an acceleration technique over a well-know inversion formula for Laplace transforms. We give uniform error bounds using a representation of these approximations in terms of gamma-type operators. We apply our results to certain mixtures of Erlang distributions which contain the class of continuous phase-type distributions.

2000 Mathematics subject classification: Primary: 60E10, 60F05, 41A25; Secondary: 60G50, 41A35.

Key words and phrases: Uniform distance; Laplace transform; gamma distribution; phase-type distribution.

1 Introduction

Frequent operations in probability such as convolution or random summation of random variables, produce probability distributions which are difficult to evaluate in an explicit way. In these cases one needs to use numerical evaluation methods, such as Fast Fourier Transform or recursive methods (see for instance, [6],[8] or [15] in a context of random sums). These methods usually require a previous discretization step of the initial random variables, when these ones are continuous. The usual way to do so, is by means of rounding methods. However, it is not always possible to evaluate the rounded random variable in an explicit way, and it is not always clear by using these methods how the rounding error propagates when one takes successive convolutions. In these cases it seems interesting to consider alternative discretization methods. For instance, when dealing with nonnegative random variables, it has been proposed in the literature the following discretization method based on the Laplace-Stieltjes transform of a random variable ([7, p.233]). Let X be a random variable taking values on $[0,\infty)$ with distribution function F. Denote by $\phi_X(\cdot)$ the Laplace-Stieltjes transform of X, that is

$$\phi_X(t) := Ee^{-tX} = \int_{[0,\infty)} e^{-tu} dF(u), \quad t > 0.$$

For each $t > 0$ we define a random variable $X^{\bullet t}$ taking values on $k/t, \quad k \in \mathbb{N}$, and such that

$$P(X^{\bullet t} = k/t) = \frac{(-t)^k}{k!} \phi_X^{(k)}(t), \quad k \in \mathbb{N}, \tag{1}$$

where $\phi_X^{(k)}$ denotes the k-th derivative ($\phi_X^{(0)} \equiv \phi_X$).

Thus, if we denote by $L_t^* F$ the distribution function of $X^{\bullet t}$ we have that,

$$L_t^* F(x) := P(X^{\bullet t} \le x) = \sum_{k=0}^{[tx]} \frac{(-t)^k}{k!} \phi_X^{(k)}(t), \quad x \ge 0, \tag{2}$$

where $[x]$ indicates the largest integer less than or equal to x. It is interesting to point out that $L_t^* F$ is the distribution function of a normalized Poisson mixture with mixing distribution tX (cf. [1, p.228]). The use of this method allows to obtain the probability mass function in an explicit way in situations in which rounding methods maybe couldn't (see for instance [1] for gamma distributions). Moreover, this method allows an easy representation of L_t^*F in terms of F which makes it possible the study of rates of convergence in the approximation ([1, 2]). In [1] the problem was studied in a general setting, whereas in [2] a detailed analysis was carried out for the case of gamma distributions that is, whose density function is given by

$$f_{a,p}(x) := \frac{a^p x^{p-1} e^{-ax}}{\Gamma(p)}, \quad x > 0. \tag{3}$$

In particular it can be seen in [2] that the error bounds for gamma distributions can be uniformly controlled for shape parameters $p \geq 1$. This property was the starting point in [14] to obtain error bounds for random sums of mixtures of gamma distributions, uniformly controlled on the parameters of the random summation index. In all these papers, the measure of distance considered was the Kolmogorov (or sup-norm) distance. More specifically, for a given real function f, defined on $[0, \infty)$ we denote by $\|f\|$ the sup-norm, that is

$$\|f\| := \sup_{x \geq 0} |f(x)|.$$

It was shown in [2] that for gamma distributions with shape parameter $p \geq 1$, we have that $\|L_t^*F - F\|$ is of order $1/t$, length of the discretization interval. Note that $\|L_t^*F - F\|$ is the Kolmogorov distance between X and $X^{\bullet t}$, as both are nonnegative random variables.

The aim of this paper is twofold. First of all, we will consider a continuous modification of (2) as when the initial distribution function is continuous, a suitable approximation by means of a continuous function can be more accurate than the approximation by a discrete distribution (see Section 2). Secondly we will give conditions under which this continuous modification has rate of convergence of $1/t^2$ instead of $1/t$ (see Section 3). In Section 4 we will consider the case of gamma distributions to see that the error bounds are also uniform on the shape parameter. Finally, in Section 5 we will consider the application of the results in Section 4 to the class of mixtures of Erlang distributions, recently studied in [16]. This class contains many of the distributions used in applied probability (in particular phase-type distributions) and is closed under important operations such as mixtures, convolution or compounding.

2 The approximation procedure

The representation of L_t^*F in (2) in terms of a Gamma process (cf. [1]) will play an important role in our proofs. We recall this representation. Let $(S(u)\ , u \geq 0)$ be a gamma process, in which $S(0) = 0$ and for $u > 0$, each $S(u)$ has a gamma density with parameters $a = 1$ and $p = u$, as given in (3) . Let g be a function defined on $[0, \infty)$. We consider the gamma-type operator L_t given by

$$L_t g(x) := Eg\left(\frac{S(tx)}{t}\right), \quad x \geq 0, \quad t > 0, \tag{4}$$

provided that this operator is well defined, that is, $L_t|g|(x) < \infty,\ x \geq 0,\ t > 0$. Then, whenever F is continuous on $(0, \infty)$, L_t^*F as defined in (2), can be written as (cf. [1, p.228])

$$L_t^*F(x) = L_t F\left(\frac{[tx]+1}{t}\right) = EF\left(\frac{S([tx]+1)}{t}\right) \quad x \geq 0, \quad t > 0. \tag{5}$$

It can be seen that the rates of convergence of $L_t g$ to g are, at most, of order $1/t$ (observe (34) below). Our aim now is to get faster rates of convergence. To this end, we will consider the following operator

$$L_t^{[2]}g(x) := 2L_{2t}g(x) - L_t g(x) = 2Eg\left(\frac{S(2tx)}{2t}\right) - Eg\left(\frac{S(tx)}{t}\right), \quad x \geq 0. \tag{6}$$

This operator will give a rate of uniform convergence from $L_t^{[2]}g$ to g of order $1/t^2$, on the following class of functions functions

$$\mathcal{D} := \{g \in C^4([0, \infty)) : \quad \|x^2 g^{iv}(x)\| < \infty\}. \tag{7}$$

The problem with $L_t^{[2]}g$ is that when tx is not a natural number, $L_t g(x)$ is given in terms of Weyl fractional derivatives of the Laplace transform (cf. [3, p. 92]) and, in general, we are not able to compute them in an explicit way. However, if we modify $L_t^{[2]}g$ using linear interpolation, that is

$$M_t^{[2]}g(x) := (tx-[tx])\left(L_t^{[2]}g\left(\frac{[tx]+1}{t}\right)\right)+([tx]+1-tx)\left(L_t^{[2]}g\left(\frac{[tx]}{t}\right)\right) \quad (8)$$

we observe that the order of convergence of $M_t^{[2]}g$ to g is also $1/t^2$, on the following class of functions

$$\mathcal{D}_1 := \{g \in C^4([0,\infty)) : \quad \|g''(x)\| \leq \infty \text{ and } \|x^2 g^{iv}(x)\| < \infty\}. \quad (9)$$

Moreover, the advantage of using $M_t^{[2]}g$ instead of $L_t^{[2]}g$ to approximate g is the computability. In the following result we note that the last approximation applied to a distribution function F, is related to $L_t^* F$, as defined in (1).

Proposition 2.1 *let X be a nonnegative random variable with Laplace transform ϕ_X. Let $L_t^* F,\ t > 0$ be as defined in* (1)*, and let $M_t^{[2]}F$ be as defined in* (8)*. We have*

$$M_t^{[2]}F\left(\frac{k}{t}\right) = \begin{cases} F(0), & \text{if } k=0; \\ 2L_{2t}^* F\left(\frac{2k-1}{2t}\right) - L_t^* F\left(\frac{k-1}{t}\right), & \text{if } k \in \mathbb{N}^* \end{cases} \quad (10)$$

and

$$M_t^{[2]}F(x) = (tx-[tx])M_t^{[2]}F\left(\frac{[tx]+1}{t}\right) + ([tx]+1-tx)M_t^{[2]}F\left(\frac{[tx]}{t}\right). \quad (11)$$

Proof. Let $t > 0$ be fixed. First, observe that by (8), we can write

$$M_t^{[2]}F\left(\frac{k}{t}\right) = L_t^{[2]}F\left(\frac{k}{t}\right), \quad k \in \mathbb{N}. \quad (12)$$

Now, using (6) and (4), we have

$$M_t^{[2]}F(0) = L_t^{[2]}F(0) = F(0), \quad (13)$$

which shows (10) for $k = 0$. Finally, using (6), (4) and (5), we have for $k \in \mathbb{N}^*$

$$L_t^{[2]} F\left(\frac{k}{t}\right) = 2EF\left(\frac{S(2k)}{2t}\right) - EF\left(\frac{S(k)}{t}\right) = 2L_{2t}^* F\left(\frac{2k-1}{2t}\right) - L_t^* F\left(\frac{k-1}{t}\right). \quad (14)$$

Thus, (12) and (14) show (10) for $k \in \mathbb{N}^*$. Note that (11) is obvious by (8) and (12). This completes the proof of Proposition 2.1. □

3 Error bounds for the approximation

Let $g \in \mathcal{D}$, as defined in (7). Our first aim is to give bounds of $\|L_t^{[2]} g - g\|$ in terms of $\|x^2 g^{iv}(x)\|$. To this end we will use as 'test function' the following one

$$\phi(x) = \begin{cases} 0, & \text{if } x = 0; \\ \dfrac{x^2}{2}\left(\dfrac{3}{2} - \log(x)\right), & \text{otherwise.} \end{cases} \quad (15)$$

Observe that $\phi \in \mathcal{D}$. In fact, by elementary calculus

$$\phi^{'}(x) = x(1 - \log x); \quad \phi^{''}(x) = -\log x; \quad \phi^{'''}(x) = -\frac{1}{x} \quad \text{and} \quad \phi^{iv}(x) = \frac{1}{x^2}. \quad (16)$$

In the next Lemma, we make an explicit computation of $L_t(\phi(x))$, in terms of the Ψ (or digamma) function. Recall that this function is defined as (cf. [4])

$$\Psi(x) := \frac{d}{dx} \log(\Gamma(x)) = \frac{1}{\Gamma(x)} \int_0^\infty \log u \; e^{-u} u^{x-1} du, \quad x > 0 \quad (17)$$

and therefore, using the last equality we have the following probabilistic expression of the psi function in terms of the gamma process:

$$\Psi(x) = ElogS(x), \quad x > 0. \quad (18)$$

We will also make use of the following well known property of this function,

$$\Psi(x+1) = \frac{1}{x} + \Psi(x). \quad (19)$$

Lemma 3.1 *Let ϕ be as defined in* (15)*, and let $L_t,\ t > 0$ be as defined in* (4)*. We have that*

$$L_t \phi(x) = \frac{1}{2t^2}\left(\frac{3(tx)^2}{2} - \frac{tx}{2} - 1 + tx(tx+1)(-\psi(tx) + \log(t))\right), \ x > 0. \quad (20)$$

Proof. Let $t > 0$ and $x > 0$ be fixed. First of all, using elementary calculus, (4) and (19), we can write

$$\begin{aligned} L_t\phi(x) &= E\frac{S(tx)^2}{2t^2}\left(\frac{3}{2} - \log\left(\frac{S(tx)}{t}\right)\right) \\ &= \frac{1}{2t^2}\frac{1}{\Gamma(tx)}\int_0^\infty u^2\left(\frac{3}{2} - \log\left(\frac{u}{t}\right)\right)e^{-u}u^{tx-1}du \\ &= \frac{(tx)(tx+1)}{2t^2}\frac{1}{\Gamma(tx+2)}\int_0^\infty \left(\frac{3}{2} - \log\left(\frac{u}{t}\right)\right)e^{-u}u^{tx+1}du \\ &= \frac{(tx)(tx+1)}{2t^2}\left(\frac{3}{2} - E\log\left(\frac{S(tx+2)}{t}\right)\right). \end{aligned} \tag{21}$$

Therefore, using (18), we can write

$$L_t\phi(x) = \frac{(tx)(tx+1)}{2t^2}\left(\frac{3}{2} - \psi(tx+2) + \log(t)\right). \tag{22}$$

Now, using twice (19), we have

$$\psi(tx+2) = \frac{2(tx)+1}{tx(tx+1)} + \psi(tx). \tag{23}$$

By (22), (23) we obtain

$$L_t\phi(x) = \frac{(tx)(tx+1)}{2t^2}\left(\frac{3}{2} - \frac{2(tx)+1}{tx(tx+1)} - \psi(tx) + \log(t)\right).$$

The result follows using elementary calculus in the expression above. □

In the next Lemma we will study the approximation properties of $L_t\phi$ to ϕ. We will make use of the following inequalities for the psi function.

$$\frac{1}{2x} \le \log(x) - \psi(x) \le \frac{1}{x}, \quad x > 0, \tag{24}$$

and

$$\log(x) - \psi(x) - \frac{1}{2x} \le \frac{1}{12x^2}, \quad x > 0. \tag{25}$$

The first inequality can be found in [4, p. 374] and the second one is an immediate consequence of the fact that the function

$$\psi(x) - \log(x) + \frac{1}{2x} + \frac{1}{12x^2}$$

is completely monotonic (cf. [13, p.304]) and thus, nonnegative.

Lemma 3.2 *Let* ϕ *be as defined in* (15)*, and let* $L_t,\ t>0$ *be as defined in* (4)*. We have*

$$\|L_t\phi(x)-\phi(x)+\frac{x\log x}{2t}+\frac{1}{3t^2}\|\leq\frac{3}{8t^2}. \tag{26}$$

Proof. Let $x>0$ and $t>0$ be fixed. First of all, we can write

$$\phi(x)=\frac{1}{2t^2}\left(\frac{3(tx)^2}{2}-(tx)^2\log(tx)+(tx)^2\log(t)\right). \tag{27}$$

On the other hand,

$$\frac{x\log x}{2t}+\frac{1}{3t^2}=\frac{1}{2t^2}\left((tx)\log tx-(tx)\log t+\frac{2}{3}\right). \tag{28}$$

Therefore, using Lemma 3.1, (27) and (28) we can write

$$\begin{aligned}&L_t\phi(x)-\phi(x)+\frac{x\log x}{2t}+\frac{1}{3t^2}\\&=\frac{1}{2t^2}\left(-\frac{tx}{2}-1-(tx)^2\psi(tx)-(tx)\psi(tx)+(tx)^2\log(tx)+(tx)\log(tx)+\frac{2}{3}\right)\\&=\frac{1}{2t^2}\left((tx)^2\left(\log(tx)-\psi(tx)-\frac{1}{2(tx)}\right)+tx(\log(tx)-\psi(tx))-\frac{1}{3}\right).\end{aligned} \tag{29}$$

By (24) we can write

$$\frac{1}{6}\leq tx(\log(tx)-\psi(tx))-\frac{1}{3}\leq\frac{2}{3}. \tag{30}$$

Thus, using (29), (30) and (25), we obtain (26). □

We are in a position to enunciate the following.

Theorem 3.1 *Let* $g\in\mathcal{D}$*, as defined in* (7) *and let* $L_t^{[2]},\ t>0$ *be as defined in* (6)*. We have*

$$|L_t^{[2]}g(x)-g(x)|\leq\frac{1}{6t^2}\|xg'''(x)\|+\frac{9}{16t^2}\|x^2g^{iv}(x)\|.$$

Proof. Let $g \in \mathcal{D}$. Note firstly that this implies that

$$\|xg'''(x)\| \leq \|x^2 g^{iv}(x)\| < \infty. \tag{31}$$

First of all it is easy to see that

$$\lim_{x\to\infty} g'''(x) = 0. \tag{32}$$

This can be deduced because for all $0 < \alpha < 1$, the fact that $\|x^2 g^{iv}(x)\| < \infty$ implies that $\lim\limits_{x\to\infty} x^{1+\alpha} g^{iv}(x) = 0$. By L'Hopital's rule, we have also that $\lim\limits_{x\to\infty} x^{\alpha} g'''(x) = 0$ thus obtaining (32) easily. Then, taking into account (32), we can write

$$g'''(x) = \int_x^{\infty} g^{iv}(u)du$$

and therefore

$$|xg'''(x)| \leq x \int_x^{\infty} \frac{|u^2 g^{iv}(u)|}{u^2} du \leq \|x^2 g^{iv}(x)\|,$$

thus implying (31).

Now, let $t > 0$ and L_t be as defined in (4). As a previous step, we will prove that

$$|L_t g(x) - g(x) - \frac{xg''(x)}{2t} - \frac{xg'''(x)}{3t^2}| \leq \frac{3}{8t^2}\|x^2 g^{iv}(x)\|, \quad x > 0. \tag{33}$$

To this end, let $x > 0$. Using and a Taylor's series expansion of the random point $u = S(tx)/t$ around x, and taking into account that $E(S(x) - x) = 0$, $E(S(x) - x)^2 = x$ and $E(S(x) - x)^3 = 2x$, we can write

$$\begin{aligned} L_t g(x) - g(x) &= Eg\left(\frac{S(tx)}{t}\right) - g(x) \\ &= \frac{E(S(tx) - tx)^2}{2t^2} g''(x) + \frac{E(S(tx) - tx)^3}{6t^3} g'''(x) + \frac{1}{6} E \int_x^{\frac{S(tx)}{t}} g^{iv}(\theta)(\frac{S(tx)}{t} - \theta)^3 d\theta \\ &= \frac{xg''(x)}{2t} + \frac{xg'''(x)}{3t^2} + \frac{1}{6} E \int_x^{\frac{S(tx)}{t}} g^{iv}(\theta)(\frac{S(tx)}{t} - \theta)^3 d\theta. \end{aligned} \tag{34}$$

Then, using (34) we get the bound

$$\left|L_t g(x) - g(x) - \frac{xg''(x)}{2t} - \frac{xg'''(x)}{3t^2}\right| = \frac{1}{6}\left|E\int_x^{\frac{S(tx)}{t}} g^{iv}(\theta)\left(\frac{S(tx)}{t}-\theta\right)^3 d\theta\right|$$
$$\leq \frac{\|x^2 g^{iv}(x)\|}{6} E\int_{\min(x,\frac{S(tx)}{t})}^{\max(x,\frac{S(tx)}{t})} \left|\frac{S(tx)}{t}-\theta\right|^3 \frac{1}{\theta^2} d\theta$$
$$= \frac{\|x^2 g^{iv}(x)\|}{6} E\int_x^{\frac{S(tx)}{t}} \left(\frac{S(tx)}{t}-\theta\right)^3 \frac{1}{\theta^2} d\theta. \quad (35)$$

Let $\phi(\cdot)$ be as in (15). Using (34) and (16) we have

$$L_t\phi(x) - \phi(x) + \frac{x\log x}{2t} + \frac{1}{3t^2} = \frac{1}{6}E\int_x^{\frac{S(tx)}{t}} \left(\frac{S(tx)}{t}-\theta\right)^3 \frac{1}{\theta^2} d\theta. \quad (36)$$

Then, by (35) and (36) we can write

$$|L_t g(x) - g(x) - \frac{xg''(x)}{2t} - \frac{xg'''(x)}{3t^2}| \leq \|x^2 g^{iv}(x)\| \|L_t\phi(x) - \phi(x) + \frac{x\log x}{2t} + \frac{1}{3t^2}\|.$$

Thus, (33) follows applying Lemma 3.2.

Observe that in (33), the only term of order $1/t$ is the one accompanying to the second derivative. We will see that by means of the operator $L_t^{[2]}$, as defined in (6) we can eliminate this term. In fact, using (33) we have

$$L_t^{[2]} g(x) - g(x) = 2\left(L_{2t} g(x) - g(x)\right) - \left(L_t g(x) - g(x)\right)$$
$$= 2\left(L_{2t} g(x) - g(x) - \frac{x}{4t} g''(x) - \frac{x}{12t^2} g'''(x)\right)$$
$$- \left(L_t g(x) - g(x) - \frac{x}{2t} g''(x) - \frac{x}{3t^2} g'''(x))\right) - \frac{x}{6t^2} g'''(x)$$
$$\leq \frac{1}{6t^2}\|xg'''(x)\| + \frac{9}{16t^2}\|x^2 g^{iv}(x)\|. \quad (37)$$

This completes the proof of Theorem 3.1. □

Finally, in the following result we consider the approximation properties of $M_t^{[2]}$.

Theorem 3.2 *Let $g \in \mathcal{D}_1$, as defined in* (9) *and let $M_t^{[2]}$, $t > 0$ be as defined in* (8)*. We have*

$$\|M_t^{[2]} g - g\| \leq \frac{1}{8t^2}\|g''(x)\| + \frac{1}{6t^2}\|xg'''(x)\| + \frac{9}{16t^2}\|x^2 g^{iv}(x)\|.$$

Proof. Note firstly that $g \in \mathcal{D}_1$ implies that $\|xg'''(x)\| < \infty$, thanks to (31). Now let $t > 0$ and $x > 0$ be fixed. We write,

$$M_t^{[2]}g(x) - g(x) = (tx - [tx])\left(L_t^{[2]}g\left(\frac{[tx]+1}{t}\right) - g\left(\frac{[tx]+1}{t}\right)\right)$$
$$+ ([tx]+1-tx)\left(L_t^{[2]}g\left(\frac{[tx]}{t}\right) - g\left(\frac{[tx]}{t}\right)\right)$$
$$+ (tx - [tx])\left(g\left(\frac{[tx]+1}{t}\right) - g(x)\right) + ([tx]+1-tx)\left(g\left(\frac{[tx]}{t}\right) - g(x)\right). \quad (38)$$

Using the usual expansion

$$|g(y) - g(x) - (y-x)g'(x)| \leq \frac{(y-x)^2}{2}\|g''\| \quad (39)$$

and taking into account that

$$(tx - [tx])\left(g\left(\frac{[tx]+1}{t}\right) - g(x)\right) + ([tx]+1-tx)\left(g\left(\frac{[tx]}{t}\right) - g(x)\right)$$
$$= (tx - [tx])\left(g\left(\frac{[tx]+1}{t}\right) - g(x) - \frac{[tx]+1-tx}{t}g'(x)\right)$$
$$+ ([tx]+1-tx)\left(g\left(\frac{[tx]}{t}\right) - g(x) - \frac{[tx]-tx}{t}g'(x)\right), \quad (40)$$

we obtain from the above expression and (39)

$$\left|(tx - [tx])\left(g\left(\frac{[tx]+1}{t}\right) - g(x)\right) + ([tx]+1-tx)\left(g\left(\frac{[tx]}{t}\right) - g(x)\right)\right|$$
$$\leq \left((tx-[tx])\frac{([tx]+1-tx)^2}{2t^2} + ([tx]+1-tx)\frac{([tx]-tx)^2}{2t^2}\right)\|g''\|$$
$$= \frac{(tx-[tx])([tx]+1-tx)}{2t^2}\|g''\| \leq \frac{1}{8t^2}\|g''\|, \quad (41)$$

the last inequality as as for each $k \in \mathbb{N}$, the supremum of $(u-k)(k+1-u),\ k \leq u \leq k+1$ is attained at $u = k+1/2$. On the other hand, taking into account Theorem 3.1 we have

$$\left|(tx - [tx])\left(L_t^{[2]}g\left(\frac{[tx]+1}{t}\right) - g\left(\frac{[tx]+1}{t}\right)\right)\right.$$
$$\left.+([tx]+1-tx)\left(L_t^{[2]}g\left(\frac{[tx]}{t}\right) - g\left(\frac{[tx]}{t}\right)\right)\right|$$
$$\leq \|L_t^{[2]}g - g\| \leq \frac{1}{6t^2}\|xg'''(x)\| + \frac{9}{16t^2}\|x^2g^{iv}(x)\|. \quad (42)$$

The result follows by (38), (41) and (42). $\square$

4 Application to gamma distributions

In this Section we will study the case of gamma distributions, that is, with density function as given in (3). It is not hard to see that these distributions are in the class $\mathcal{D}_1$, for a shape parameter $p = 1$ or $p \geq 2$, and therefore, we are a position of apply Theorem 3.2. The aim of this Section is to show that in fact, the bounds in this Theorem can uniformly bounded on the shape parameter, which will be an advantage when dealing with mixtures of these distributions. From now on, we denote by

$$f_p(x) := \begin{cases} \dfrac{e^{-x}x^{p-1}}{\Gamma(p)}, & x > 0, \quad \text{if } p \in \mathbb{R} \setminus \{0, -1, -2, \dots\}; \\ 0, & x > 0, \quad \text{if } p \in \{0, -1, -2, \dots\}, \end{cases} \tag{43}$$

Note that for $p > 0$ the function above is the density of a gamma random variable as in (3) with scale parameter $a = 1$. Results for another scale parameter will follow by a change of scale (see Proposition 5.2 below). First of all we will consider the case $p = 1$, that is an exponential random variable. As the distribution function of this random variable has no computational problems, it makes no sense to approximate it. However, when we consider the problem of approximating a general mixture of Gamma distributions, the exponential distribution could be a component.

Lemma 4.3 *Let $F(x) = 1 - e^{-x},\ x \geq 0$. For $t > 0$, let $M_t^{[2]}F$ be as defined in (8). We have that*

$$\|M_t^{[2]}F - F\| \leq \left(\frac{1}{8} + \frac{1}{6e} + \frac{9}{4e^2}\right)\frac{1}{t^2}$$

Proof. First of all, note that $|F^{(k)}(x)| = e^{-x}$, and that $\sup_{x\geq 0} x^k e^{-x} = k^k e^{-k},\ k = 1, 2, \ldots$. Thus, we have

$$\|F''\| = 1, \quad \|xF''(x)\| = e^{-1} \quad \text{and} \quad \|x^2 F^{iv}(x)\| = 2^2 e^{-2} \tag{44}$$

The conclusion follows taking into account Theorem 3.2. $\square$

Now we will deal with the case $p \geq 2$ in (43). The following Lemma will be useful in order to bound the derivatives of this density.

Lemma 4.4 *Let $f_p(\cdot),\ p > 0$ be as defined in* (43). *We have for all $n \in \mathbb{N}$*

$$\begin{aligned} \frac{d^n}{dx^n} f_p(x) &= \frac{e^{-x} x^{p-n-1}}{\Gamma(p)} \sum_{i=0}^{n} \binom{n}{i} (-1)^i \left(\prod_{j=1}^{n-i} (p-j) \right) x^i \\ &= \sum_{i=0}^{n} \binom{n}{i} (-1)^i f_{p-n+i}(x), \quad x > 0, \end{aligned} \tag{45}$$

in which $\prod_{j=1}^{0} (p-j) = 1$.

Proof. Let $n \in \mathbb{N}$, $p > 0$ and $x > 0$. We recall (43) and apply Leibniz's rule for derivatives to write

$$\begin{aligned} \frac{d^n}{dx^n} f_p(x) &= \frac{1}{\Gamma(p)} \sum_{i=0}^{n} \binom{n}{i} \frac{d^i}{dx^i} e^{-x} \cdot \frac{d^{n-i}}{dx^{n-i}} x^{p-1} \\ &= \frac{1}{\Gamma(p)} \sum_{i=0}^{n} \binom{n}{i} (-1)^i e^{-x} \left(\prod_{j=1}^{n-i} (p-j) \right) x^{p-1-(n-i)}, \end{aligned}$$

which proves the first inequality in (45). The second equality follows because

$$f_{p-n+i}(x) = \frac{e^{-x} x^{p-n-1}}{\Gamma(p)} \left(\prod_{j=1}^{n-i} (p-j) \right) x^i, \quad i = 0, 1, \ldots n \tag{46}$$

Actually, for $p - n + i \in \mathbb{R} \setminus \{0, -1, -2 \ldots\}$ we recall (43) and the fact that $\frac{\Gamma(p)}{\Gamma(p-n+i)} = \prod_{j=1}^{n-i} (p-j)$. For $p - n + i \in \{0, -1, -2 \ldots\}$, observe that both terms in (46) are equal to 0. $\square$

The aim of the following results is to get bounds of the quantities required in Theorem 3.2, depending on the shape parameter p, but also decreasing on this parameter.

First of all we formulate a technical lemma in which we define certain decreasing functions, which will be used to bound the weighted derivatives of f_p. Its proof is rather long, although only elementary calculus is required.

Lemma 4.5 *We have*

(i) The function

$$g_1(p) := \frac{1}{\Gamma(p)} e^{-(p-1)} (p-1)^{p-1}, \quad p > 1, \quad (g_1(1) = 1), \tag{47}$$

is decreasing in p.

(ii) The function

$$g_2(p) := \frac{1}{\Gamma(p)} e^{-\left(p-\frac{1}{2}+\frac{1}{2}\sqrt{4p-3}\right)} \left(p - \frac{1}{2} + \frac{1}{2}\sqrt{4p-3}\right)^{p-1/2}, \quad p \geq 1 \tag{48}$$

is decreasing in p.

(iii) The function

$$g_3(p) := \frac{1}{\Gamma(p)} e^{-(p-1-\sqrt{p-1})} (\sqrt{p-1} - 1)^{p-2} (\sqrt{p-1})^{p-1}, \quad p > 2, \tag{49}$$

$(g_3(2) = 1)$, *is decreasing in p.*

(iv) The function

$$g_4(p) := \frac{1}{\Gamma(p)} e^{-(p-\sqrt{3p-2})} (p - \sqrt{3p-2})^{p-2} (\sqrt{3p-2} - 1)^3, \quad p > 2 \tag{50}$$

$(g_4(2) = 1)$ *is decreasing in p.*

Proof. Parts (i) and (ii) are proven in [14]. (see Lemmas 5.1 and 5.2 in this paper for (i) and (ii), respectively).

To show part (iii), define the auxiliary function

$$l_3(u) := -\log\Gamma(u^2+1) - u(u-1) + (u^2-1)\log(u-1) + u^2\log(u), \quad u > 1. \quad (51)$$

Note that $g_3(\cdot)$, as defined in (49), can be expressed as

$$g_3(p) = e^{l_3(\sqrt{p-1})}, \quad p > 2. \quad (52)$$

We will show firsty that l_3 is decreasing. In fact, it follows by calculus (recall, (17)) that

$$l_3'(u) = 2u(-\psi(u^2+1) + \log(u(u-1))) + 2, \quad u > 1. \quad (53)$$

Now, we use (24) to write

$$l_3'(u) \le 2u\left(-\log(u^2+1) + \frac{1}{u^2+1} + \log(u(u-1))\right) + 2, \quad u > 1. \quad (54)$$

Divide the right hand side by $1/(2u)$ and call

$$d_3(u) := -\log(u^2+1) + \frac{1}{u^2+1} + \log(u(u-1)) + \frac{1}{u}, \quad u > 1.$$

It can be checked by calculus that

$$d_3'(u) = \frac{1 - 2u + 4u^2 + u^4}{u^2(u-1)(u^2+1)^2} \ge 0, \quad u > 1$$

and that $\lim_{u\to\infty} d_3(u) = 0$. Then, we conclude that $d_3(u) \le 0$, $u > 1$, and therefore, by (54), that $l_3'(u) \le 2ud_3(u) \le 0$, thus showing that $l_3(\cdot)$ is decreasing. This implies, recalling (52) that $g_3(p)$ is decreasing, thus concluding the proof of part (iii). The proof of (iv) is very similar to the proof of (iii). Firstly, define

$$l_4(u) := -\log\Gamma\left(\frac{u^2+2}{3}\right) - \frac{(u-1)(u-2)}{3} + \frac{u^2-4}{3}\log\left(\frac{(u-1)(u-2)}{3}\right) + 3\log(u-1), \quad u > 2.$$

We observe that

$$g_4(p) = e^{l_4(\sqrt{3p-2})}, \quad p > 2. \quad (55)$$

We will show that $l_4(\cdot)$ is decreasing. Thus, taking derivatives it can be checked that

$$l_4'(u) = \frac{2u}{3}\left(-\psi\left(\frac{u^2+2}{3}\right) + \log\left(\frac{(u-1)(u-2)}{3}\right)\right) + \frac{2u}{u-1}, \quad u > 2.$$

Now, using (25), we have

$$l_4'(u) \leq \frac{2u}{3}\left(-\log\left(\frac{(u^2+2)}{3}\right) + \frac{3}{2(u^2+2)} + \frac{3}{4(u^2+2)^2} + \log\left(\frac{(u-1)(u-2)}{3}\right)\right) + \frac{2u}{u-1}, \quad u > 2. \tag{56}$$

To show that $l_4'(u) \leq 0$, we divide by $2u/3$ the above expression and define

$$d_4(u) := -\log\left(\frac{(u^2+2)}{3}\right) + \frac{3}{2(u^2+2)} + \frac{3}{4(u^2+2)^2} + \log\left(\frac{(u-1)(u-2)}{3}\right) + \frac{3}{u-1}, \quad u > 2.$$

After some computations we see that

$$d_4'(u) = 3\frac{2u^4 - 2u^3 + 13u^2 - 10u + 24}{(u-1)^2(u^2+2)^3(u-2)} \geq 0, \quad u > 2.$$

This means that d_4 is increasing. As $\lim_{u\to\infty} d_4(u) = 0$, we have that $d_4(u) \leq 0,\ u > 2$. Then, using (56) we conclude that

$$\frac{3}{2u} l_4'(u) \leq d_4(u) \leq 0, \quad u > 2.$$

This shows that $l_4(u)$ is decreasing. Using this fact and taking into account (55), we obtain (iv). The proof of Lemma 4.5 is complete. □

Lemma 4.6 *Let f_p be as in* (43) *We have*

(i) $\sup_{x\geq 0} |f_p(x)| = g_1(p), \quad p \geq 1.$

(ii) $\sup_{x\geq 0} |x f_p'(x)| = g_2(p), \quad p \geq 1.$

(iii) $\sup_{x\geq 0} |f_p'(x)| = g_3(p), \quad p \geq 2.$

(iv) $\sup_{x\geq 0} |xf_p''(x)| \leq \max\{g_1(p-1), g_2(p-1)\}, \quad p \geq 2.$

(v) $\sup_{x\geq 0} |x^2 f_p'''(x)| \leq g_4(p) + 3g_2(p-1) + g_1(p-1), \quad p \geq 2.$

Proof. To show part (i), it is clear that, for $p \geq 1$,

$$\sup_{x\geq 0} f_p(x) = f_p(p-1) = \frac{e^{-(p-1)}(p-1)^{p-1}}{\Gamma(p)},$$

and (i) follows recalling (47). To show part (ii) we have (cf. [14] Remark 3.2. and Lemma 5.2)

$$\sup_{x\geq 0} |xf_p'(x)| = \frac{1}{\Gamma(p)} \left(p - \frac{1}{2} + \frac{1}{2}\sqrt{4p-3}\right)^{p-1/2} e^{-p-\frac{1}{2}+\frac{1}{2}\sqrt{4p-3}}, \quad p > 1, \quad (57)$$

and (ii) follows recalling (48). To show part (iii), by (45), we have for $p \geq 2$,

$$f_p'(x) = \frac{1}{\Gamma(p)} e^{-x} x^{p-2}(p-1-x), \quad x > 0, \quad (58)$$

$$f_p''(x) = \frac{1}{\Gamma(p)} e^{-x} x^{p-3}((p-1)(p-2) - 2(p-1)x + x^2), \quad x > 0, \quad (59)$$

and it can be checked easily that the zeroes of $f_p''(x)$ are $p_1 := p - 1 - \sqrt{p-1}$ and $p_2 := p - 1 + \sqrt{p-1}$. Therefore, $|f_p'(x)|$ must attain its maximum value either at p_1 or p_2. Actually p_1 corresponds to the maximum. To show that we will see that

$$\frac{f_p'(p_1)}{|f_p'(p_2)|} = e^{2\sqrt{p-1}} \left(\frac{\sqrt{p-1}-1}{\sqrt{p-1}+1}\right)^{p-2} \geq 1, \quad p \geq 2. \quad (60)$$

To show the last inequality in (60), taking logarithms we will prove that

$$r_1(p) := 2\sqrt{p-1} + (p-2)\Big(\log(\sqrt{p-1}-1) - \log(\sqrt{p-1}+1)\Big) \geq 0, \quad p > 2. \quad (61)$$

Call

$$\rho_1(b) := \frac{2b}{b^2-1} + (\log(b-1) - \log(b+1)), \quad b > 1.$$

Note that

$$r_1(p) = (p-2)\rho_1(\sqrt{p-1}), \quad p > 2. \quad (62)$$

We will firstly prove that

$$\rho_1(b) \geq 0, \quad b > 1. \tag{63}$$

To show (63), it is readily seen that $\rho_1'(b) = -4(b^2-1)^{-2},\ b > 1$, so that ρ_1 is decreasing. As $\lim_{b\to\infty} \rho_1(b) = 0$, we have (63). This implies also (61), recalling (62). Therefore, we conclude that

$$\sup_{x>0} |f_p'(x)| = f_p'(p_1) = \frac{1}{\Gamma(p)} e^{-(p-1-\sqrt{p-1})}(\sqrt{p-1}-1)^{p-2}(\sqrt{p-1})^{p-1}, \tag{64}$$

this, together with (49), shows (iii).

To show part (iv), note that using (45), we can write $f_p'(x) = f_{p-1}(x) - f_p(x)$ and therefore,

$$x f_p''(x) = x f_{p-1}'(x) - x f_p'(x), \quad x > 0, \quad p \geq 2. \tag{65}$$

On the other hand, we see in (59) that $f_{p-1}'(x)$ and $f_p'(x)$ have the same sign for $0 < x < p-2$ and $p-1 < x < \infty$ and therefore, using part (ii), and Lemma 4.5(i), we can write

$$\sup_{x\notin[p-2,p-1]} |x f_p''(x)| \leq \max(g_2(p-1), g_2(p)) = g_2(p-1). \tag{66}$$

On the other hand we have by (59)

$$x f_p''(x) = \frac{1}{\Gamma(p)} e^{-x} x^{p-2}((p-1)(p-2) - 2(p-1)x + x^2) \tag{67}$$

using the above expression and taking into account that for $p-2 \leq x \leq p-1$

$$e^{-x} x^{(p-2)} \leq e^{-p-2}(p-2)^{p-2} \quad \text{and} \quad |(p-1)(p-2)-2(p-1)x+x^2| = p-1, \tag{68}$$

the last inequality as $|(p-1)(p-2) - 2(p-1)x + x^2|,\ p-2 \leq x \leq p-1$ attains its maximum value at $p-1$. From (67) and (68), we conclude that

$$\sup_{x\in[p-2,p-1]} |x f_p''(x)| \leq \frac{1}{\Gamma(p)} e^{-(p-2)}(p-2)^{p-2}(p-1) = g_1(p-1), \tag{69}$$

the last inequality by (i). Thus (66) and (69) conclude the proof of (iv). To show (v), let $p \geq 2$. We have firstly, by (45)

$$\begin{aligned} f_p'''(x) &= f_{p-3}(x) - 3f_{p-2}(x) + 3f_{p-1}(x) - f_p(x) \\ &= \frac{e^{-x}x^{p-4}}{\Gamma(p)}((p-1)(p-2)(p-3) - 3(p-1)(p-2)x + 3(p-1)x^2 - x^3) \\ &= \frac{e^{-x}x^{p-4}}{\Gamma(p)}((p-1-x)^3 + 3(p-1)(x-(p-2)) - (p-1)), \; x > 0. \quad (70) \end{aligned}$$

Therefore, if we call

$$h_p(x) := \frac{e^{-x}x^{p-2}}{\Gamma(p)}(p-1-x)^3, \quad x > 0.$$

We have, recalling (58)

$$\begin{aligned} x^2 f_p'''(x) &= \frac{e^{-x}x^{p-2}}{\Gamma(p)}((p-1-x)^3 - 3(p-1)(x-(p-2)) - (p-1)) \\ &= h_p(x) + 3xf_{p-1}'(x) - f_{p-1}(x), \quad x \geq 0. \quad (71) \end{aligned}$$

We will firstly see that

$$\sup_{x \geq 0} |h_p(x)| = g_4(p), \quad (72)$$

with $g_4(\cdot\,)$ as defined in (50). Note that

$$h_p'(x) = \frac{e^{-x}x^{p-3}}{\Gamma(p)}(p-1-x)^2(x^2 - 2px + (p-1)(p-2)), \quad x > 0$$

The maximum value of $|h_p|$ will be attained at the roots of the last polynomials, being $p_1 := p + \sqrt{3p-2}$ and $p_2 := p - \sqrt{3p-2}$. To check which value attains the maximum, call $u := \sqrt{3p-2}$. Note that $p_1 = (u+1)(u+2)/3$ and $p_2 = (u-1)(u-2)/3$. Then, with this notation we will prove that

$$\frac{|h_p(p_2)|}{|h_p(p_1)|} = e^{2u}\left(\frac{(u-1)(u-2)}{(u+1)(u+2)}\right)^{\frac{u^2-4}{3}}\left(\frac{u-1}{u+1}\right)^3 \geq 1, \quad u > 2. \quad (73)$$

To show the last inequality in (73), taking logarithms, we will show that

$$\rho_2(u) := 2u + \frac{u^2-4}{3}\log\left(\frac{(u-1)(u-2)}{(u+1)(u+2)}\right) + 3\log\left(\frac{u-1}{u+1}\right) \geq 0 \quad u > 2. \quad (74)$$

Note that

$$\rho_2'(u) = 2 + \frac{2u}{3}\log\left(\frac{(u-1)(u-2)}{(u+1)(u+2)}\right) + \frac{u^2-4}{3}\left(\frac{1}{u-1} + \frac{1}{u-2} - \frac{1}{u+1} - \frac{1}{u+2}\right)$$
$$+ 3\left(\frac{1}{u-1} - \frac{1}{u+1}\right) = \frac{4u^2}{u^2-1} + \frac{2u}{3}\log\left(\frac{(u-1)(u-2)}{(u+1)(u+2)}\right) \quad u > 2.$$

We will show that $\rho_2'(u) \le 0, u > 2$. In fact,

$$\frac{d}{du}\frac{3}{2u}\rho_2'(u) = \frac{36}{(u+1)^2(u-1)^2(u^2-4)^2} \ge 0, \quad u > 2.$$

and then $3(2u)^{-1}\rho_2'(u)$ is increasing. As $\lim_{u\to\infty} 3(2u)^{-1}\rho_2'(u) = 0$, we conclude that $3(2u)^{-1}\rho_2'(u) \le 0$, and thus that $\rho_2'(u) \le 0$. Therefore, $\rho_2(u)$ is decreasing. This, together with the fact that $\lim_{u\to\infty}\rho_2(u) = 0$, proves (74), and therefore (73). Then $\|h_p\| = h_p(p_2) = g_4(p)$, thus proving (72). Now, the proof of part (iv) follows easily recalling (71) and using (72) and parts (i) and (ii). □

As an immediate consequence of Theorem 3.2 and Lemma 4.6 we have the following

Corollary 4.1 *Let F_p be a gamma distribution of shape parameter $p \ge 2$, that is whose density function is given by* (43). *Let $M_t^{[2]}$, $t > 0$ be as defined in* (8). *We have*

$$\|M_t^{[2]}F_p - F_p\| \le \left(\frac{17}{12} + \frac{27}{16e}\right)\frac{1}{t^2} \approx \frac{2.0375}{t^2}$$

Proof. Let $p \ge 2$ be fixed. The result is an immediate consequence of Theorem 3.2, as $F_p' = f_p$ as defined in (43). Therefore by Lemma 4.6(iii) and Lemma 4.5 (ii) we have that

$$\|F_p''\| = \|f_p'\| = g_3(p) \le g_3(2) = 1. \tag{75}$$

On the other hand, we see we have by Lemma 4.5 (i) that

$$g_1(p-1) \le g_1(1) = 1 \quad \text{and} \quad g_2(p-1) \le g_2(1) = e^{-1}, \quad p \ge 2 \tag{76}$$

Thus, using the above inequalities and Lemma 4.6(iv), we have

$$\|xF_p'''(x)\| = \|xf_p''(x)\| \leq 1. \tag{77}$$

Finally by Lemma 4.6(v), Lemma 4.5 (iv) and (76) we have

$$\|x^2 F_p^{iv}(x)\| = \|x^2 f_p'''(x)\| \leq g_4(2) + 3g_2(1) + g_1(1) = 2 + 3e^{-1}. \tag{78}$$

Using (75), (77), (78), and Theorem 3.2, we obtain the result. This completes the proof of Corollary 4.1. □

5 Applications to mixtures of Erlang distributions and phase-type distributions

In this Section we apply the results in the previous Section to mixtures of Erlang distributions, and to random sums of them. In order to make the study for an arbitrary scale parameter, we see in the following result the behaviour of $M_t^{[2]}F$ under changes of scale.

Proposition 5.2 *Let X be a random variable with distribution function F. For a given $c > 0$ denote by F^c the distribution function of cX. Let $M_t^{[2]}F$ and $M_t^{[2]}F^c$, $t > 0$ be the respective approximations for F and F^c, as defined in (8). We have that*

$$M_t^{[2]}F^c(x) = M_{ct}^{[2]}F(x/c), \quad x \geq 0. \tag{79}$$

Therefore,

$$\|M_t^{[2]}F^c - F^c\| = \|M_{ct}^{[2]}F - F\|. \tag{80}$$

Proof. Let $t > 0$ and $c > 0$ be fixed. First of all, we will see that

$$M_t^{[2]}F^c\left(\frac{k}{t}\right) = M_{ct}^{[2]}F\left(\frac{k}{ct}\right), \quad k \in \mathbb{N}, \tag{81}$$

and therefore, (79) is satisfied for points in the set $k/t,\ k \in \mathbb{N}$. To this end, we use (12) and (6), and take into account that

$$F^c(x) = F(x/c), \quad x \geq 0, \tag{82}$$

to write, for all $k \in \mathbb{N}$,

$$\begin{aligned} M_t^{[2]} F^c\left(\frac{k}{t}\right) &= 2EF^c\left(\frac{S(2k)}{2t}\right) - EF^c\left(\frac{S(k)}{t}\right) \\ &= 2EF\left(\frac{S(2k)}{2ct}\right) - EF\left(\frac{S(k)}{ct}\right) = M_{ct}^{[2]} F\left(\frac{k}{ct}\right), \end{aligned} \tag{83}$$

thus proving (81). For a general $x > 0$, we use (8) and (81), to see that

$$\begin{aligned} M_t^{[2]} F^c(x) = (tx - [tx]) M_t^{[2]} F^c\left(\frac{[tx]+1}{t}\right) + ([tx]+1-tx) M_t^{[2]} F^c\left(\frac{[tx]}{t}\right) \\ (tx - [tx]) M_{ct}^{[2]} F\left(\frac{[tx]+1}{ct}\right) + ([tx]+1-tx) M_{ct}^{[2]} F\left(\frac{[tx]}{ct}\right) = M_{ct}^{[2]} F\left(\frac{x}{c}\right), \end{aligned}$$

the last inequality being trivial, as $tx = (ct)\dot{(}x/c)$. This concludes the proof of (79). Finally, (80) follows easily from (79) and (82), as we have

$$\sup_{x>0} |M_t^{[2]} F^c(x) - F^c(x)| = \sup_{x>0} |M_{ct}^{[2]} F(x/c) - F(x/c)|$$

This concludes the proof of Proposition 5.2. □

As an application of the results in the previous Section, we will consider the class of (possibly infinite) mixtures of Erlang distributions recently studied by Willmot and Woo (cf. [16]). More specifically let $F_{(a,j)},\ a > 0,\ j \in \mathbb{N}^*$, be the distribution function corresponding to the density $f_{(a,j)}$ given in (3). (an Erlang j distribution with scale parameter a). By convention $F_{(a,0)}\ a > 0$, will mean the distribution degenerate at the point 0. We will consider a finite number of scale parameters arranged in increasing order ($0 < a_1 < \cdots < a_n$), and a set of nonnegative numbers $p_{ij},\ i = 1, \ldots n,\ j = 0, 1, 2, \ldots$, such that $\sum_{i=1}^{n} \sum_{j=0}^{\infty} p_{ij} = 1$, and define the class of distribution function $\mathcal{ME}(a_1, \ldots, a_n)$ given as

$$F(x) = \sum_{i=1}^{n} \sum_{j=1}^{\infty} p_{ij} F_{a_i,j}(x) \tag{84}$$

As it was pointed out in Willmot and Woo (cf. [16]), a distribution as in (84) admits an alternative expression by using only the maximum of the scale parameters, that is

$$F(x) = \sum_{j=0}^{\infty} p_j F_{a_n,j}(x) \tag{85}$$

(see [16] for more details). Moreover, the class (85) is a wide class containing many of the distributions considered in applied probability, such as (obviously) finite mixtures of Erlangs, but also the class of phase-type distributions (see Proposition 5.4 below). Every random variable having a representation as in (84) can be approximated by means of $M_t^{[2]}$, as it is shown in the following.

Proposition 5.3 *Let F be a distribution function of the form $\mathcal{ME}(a_1, \ldots, a_n)$, $0 < a_1 < \cdots < a_n$, as in* (84). *Let $M_t^{[2]}$, $t > 0$ be as defined in* (8). *We have*

$$\|M_t^{[2]} F - F\| \leq \left(\frac{17}{12} + \frac{27}{16e} \right) \frac{\sum_{i=1}^{n} (\sum_{j=1}^{\infty} p_{ij}) a_i^2}{t^2}. \tag{86}$$

Proof. Let $t > 0$ and $0 < a_1 < \cdots < a_n$ be fixed. Note that the linearity of $M_t^{[2]}$ allows us to write

$$M_t^{[2]} F(x) = \sum_{i=1}^{n} \sum_{j=0}^{\infty} p_{ij} M_t^{[2]} F_{a_i,j}(x), \quad x \geq 0. \tag{87}$$

By Corollary 4.1 we can write, for a scale parameter 1,

$$\|M_t^{[2]} F_{1,j} - F_{1,j}\| \leq \left(\frac{17}{12} + \frac{27}{16e} \right) \frac{1}{t^2}, \quad j = 2, 3, \ldots \tag{88}$$

Moreover, using Lemma 4.3 we have

$$\|M_t^{[2]} F_{1,1} - F_{1,1}\| \leq \left(\frac{1}{2} + \frac{1}{6e} + \frac{9}{4e^2} \right) \frac{1}{t^2} \leq \left(\frac{17}{12} + \frac{27}{16e} \right) \frac{1}{t^2} \tag{89}$$

Let now the general scale parameters a_i, $i = 1, \ldots, n$. We use that given X a gamma random variable of scale parameter 1, then, X/a_i is a gamma random variable of scale parameter a_i, and therefore, using Proposition 5.2, (88) and (89), we have for each $a_i,\ i = 1, \ldots, n$ and $j \in \mathbb{N}^*$

$$\|M_t^{[2]} F_{a_i,j} - F_{a_i,j}\| = \|M_{t/a_i}^{[2]} F_{1,j} - F_{1,j}\| \leq \left(\frac{17}{12} + \frac{27}{16e}\right) \frac{a_i^2}{t^2}. \tag{90}$$

It can be checked, using (6) and (8) that $M_t^{[2]} F_{a_i,0} = F_{a_i,0}$. Thus using (87) and (90) we have

$$\begin{aligned} \|M_t^{[2]} F - F\| &\leq \sum_{i=1}^{n} \sum_{j=1}^{\infty} p_{ij} \|M_t^{[2]} F_{a_i,j} - F_{a_i,j}\| \\ &\leq \left(\frac{17}{12} + \frac{27}{16e}\right) \frac{\sum_{i=1}^{n} (\sum_{j=1}^{\infty} p_{ij}) a_i^2}{t^2}. \end{aligned} \tag{91}$$

This completes the proof of Proposition 5.3. □

Let $(X_i)_{i \in \mathbb{N}^*}$ be a sequence of independent, identically distributed nonnegative random variables. Let M be a random variable concentrated on the nonnegative integers, independent of $(X_i)_{i \in \mathbb{N}^*}$. Consider the random variable

$$\sum_{i=1}^{M} X_i, \tag{92}$$

with the convention that the empty sum is 0.

As a consequence of the previous result, we can provide error bounds for compound distributions of mixtures of Erlangs, as stated in the following

Corollary 5.2 *Let G be a compound distribution of mixtures of Erlangs, that is the distribution function of a random sum as in* (92)*, in which the sequence of $(X_i)_{i \in \mathbb{N}^*}$ has a common distribution $\mathcal{ME}(a_1, \ldots, a_n), 0 < a_1 < \cdots < a_n$, as defined in* (84)*. Let $M_t^{[2]}$ be as in* (8)*. We have that*

$$\|M_t^{[2]} G - G\| \leq \left(\frac{17}{12} + \frac{27}{16e}\right) \frac{(1 - G(0)) a_n^2}{t^2},$$

Proof. The proof is immediate taking into account that a mixture of Erlangs $\mathcal{ME}(a_1, \ldots, a_n), 0 < a_1 < \cdots < a_n$ can be expressed as in (85), and compound distributions of these random variables are also mixtures of Erlang (cf. [16, p.106], with a slight modification in the coefficients, as we allow a point mass at 0), that is, we can write

$$G(x) = \sum_{j=0}^{\infty} q_j F_{a_n,j}(x), \quad x \geq 0,$$

in which $\{q_j,\ j = 0, 1, \dots\}$ form a probability mass function. Note that, obviously, $q_0 = G(0)$. Using the above expression the result is immediate by Proposition 5.3. $\Box$

The class of phase type distributions, of great importance in applied probability, can be expressed as mixtures of Erlangs. Phase-type distribution are defined as the time until absorption in a continuous-time Markov chain with one absorbent state (cf., for instance [10, Ch.II], or [5, Ch.VIII], and the references therein). A phase-type distribution can be expressed in terms of a matrix exponential as follows. Consider a vector $\alpha = (\alpha_1, \dots \alpha_n)$ of nonnegative numbers such that $\alpha_1 + \cdots + \alpha_n \leq 1$. Let A be a $n \times n$ matrix with negative diagonal entries, non-negative off-diagonal entries and non-positive row sums. A nonnegative random variable X is a phase type distribution $PH(\alpha, A)$ if its distribution function is written as

$$F(x) = 1 - \alpha e^{xA} \mathbf{1'}, \quad x \geq 0,$$

in which $\mathbf{1'}$ represent the transpose of the n th dimensional vector $\mathbf{1} = (1, \dots, 1)$. Note that phase-type distributions are absolutely continuous random variables when $\alpha_1 + \cdots + \alpha_n = 1$, having positive mass at 0 (of magnitude $1 - (\alpha_1 + \cdots + \alpha_n)$) when $\alpha_1 + \cdots + \alpha_n < 1$. Phase-type distributions have been extensively studied both from a theoretical and practical point of view. In particular, the following property of phase-type distributions, due to Maier (cf.[11, p.591]) allows us to give expressions of phase-type distributions in terms of mixtures of Erlangs.

Let f be the density of an absolutely continuous phase-type distribution. There exists some $c > 0$ verifying

$$c_j := \frac{d^j}{dx^j} e^{cx} f(x) \bigg|_{x=0} > 0, \quad j \in \mathbb{N}. \tag{93}$$

We are in a position to enunciate the following.

Proposition 5.4 *Let F be a phase-type distribution $PH(\alpha, A)$, with $\alpha_1 + \cdots + \alpha_n > 0$. Let $c > 0$ be such that the absolutely continuous part of F satisfies property* (93)*. Then F can be expressed as a mixture of Erlangs, that is*

$$F(x) = \sum_{j=0}^{\infty} p_j F_{c,j}(x), \quad x \geq 0, \tag{94}$$

in which $p_0 = 1 - (\alpha_1 + \cdots + \alpha_n)$.

Proof. To prove (a) assume firstly that F is absolutely continuous, that is, $\alpha_1 + \cdots + \alpha_n = 1$. Then, its density is given by $f(x) = -\alpha e^{xA} A\mathbf{1}', \quad x > 0$. We choose $c > 0$ verifying (93). Note that we can write

$$e^{cx} f(x) = -\alpha e^{x(cI-A)} A\mathbf{1}', \quad x \geq 0. \tag{95}$$

It can be easily checked that the function $-\alpha e^{x(cI-A)} A\mathbf{1}', \quad x \in \mathbb{R}$ is analytic in $\mathbb{R}$, so that if we consider the Taylor's series expansion of this function around 0, and take into account (93) and (95), we have

$$e^{cx} f(x) = \sum_{j=0}^{\infty} c_j \frac{x^j}{j!}, \quad x > 0,$$

from which we can write (recall (3))

$$f(x) = \sum_{j=0}^{\infty} \frac{c_j}{c^{j+1}} \frac{c^{j+1} x^j e^{-cx}}{j!} = \sum_{j=0}^{\infty} \frac{c_j}{c^{j+1}} f_{c,j+1}(x), \quad x > 0$$

and in this way we obtain the expression of f in terms of a mixture of Erlang densities with shape parameter c (by construction the coefficients are non-negative, and integrating both sides in the above expression we see that their sum is 1). As a consequence we can write

$$F(x) = \sum_{j=1}^{\infty} \frac{c_{j-1}}{c^j} F_{c,j}(x), \quad x \geq 0, \tag{96}$$

thus having F expressed as a mixture of Erlangs, as in (94). Now assume that $0 < \alpha_1 + \cdots + \alpha_n < 1$. This means that F has a point mass at 0 of magnitude $p_0 := 1 - (\alpha_1 + \cdots + \alpha_n)$. The absolutely continuous part of F (F^{ac}) is a phase type distribution ($PH(\bar{\alpha}, A)$), with $\bar{\alpha} = (\alpha_1 + \cdots + \alpha_n)^{-1}\alpha$. Let $c > 0$ be such that F^{ac} verifies property (95). We can write thanks to (96)

$$F(x) = p_0 F_{0,c}(x) + (1-p_0)F^{ac}(x) = p_0 F_{0,c}(x) + \sum_{j=1}^{\infty}(1-p_0)\frac{c_{j-1}}{c^j} F_{c,j}(x), \quad x \geq 0$$

This completes the proof of Proposition 5.4. $\square$

Remark 5.1 Similar expansions to that given in Proposition 5.4 can be found in [10, p. 58]. These expansions are obtained using a representation $PH(\alpha, A)$ of the distribution under consideration. Note that if we denote by $\|A\|$ the maximum absolute value of the entries of A, it is easy to check using (95) (cf. [12, p.751]), that $c = \|A\|$ verifies (93). However, as the representation of a phase type is not unique this value might not be the optimum one. Observe also that the error bound given in (86) indicates that we should take c as small as possible. This problem then, is closely connected with Conjecture 6 in [12], concerning the minimum c satisfying (93) and its relation with a phase-type representation having $\|A\|$ as small as possible. To the best of our knowledge, this conjecture remains, nowadays, unsolved.

Remark 5.2 It is well known that phase-type distributions have a rational Laplace transform. Thus, we can easily approximate a phase-type distribution using Proposition 2.1, as this methods is based on the Laplace transform and its successive derivatives. Moreover, for a given random variable X, we have (cf.

[1, p. 228] that $tX^{\bullet t}$, as defined in (1) represents the number of Poisson events (of rate 1) during a random interval tX, so that if X is continuous phase type, we deduce (see [10, p.50]) that $tX^{\bullet t}$ is discrete phase type. On the other hand, Corollary 5.4 says us that phase-type distributions can be expressed as mixtures of Erlangs, so that we can give error bounds in the approximation using Proposition 5.3. We can also use our discretization method to approximate the distribution of random sums having phase type summands. This can be done by a straightforward application of Proposition 2.1 if we have a closed form expression for the Laplace transform of the random sum. Otherwise, we can use (1) to discretize the individual summands, and use afterwards computational methods existing in the literature to calculate the probability mass function of the discretized random sum (see [6],[8] or [15] for general methods, and [9], for a simple recursive formula when the summands are of discrete phase-type). Afterwards, we would use Proposition 2.1 to get the final approximation. The error bounds in this case, would be given by using Corollary 5.4. Finally, recall that a random sum of phase-type distributions is itself a phase type distribution, when the random summation index is of phase-type. In this case, we could apply matrix-analytic methods to compute the distribution function of the random sum (cf. [5] and the references therein). However, when the random summation index is not a phase-type distribution, the resulting compound distribution might not be of phase type (cf. [10, p.56]).

Acknowledgments

I would like to thank José Garrido, for suggesting me the final applications in phase-type distributions when I was at Concordia University. This research has been partially supported by the research grants 2006-CIE-05 (University of Zaragoza) MTM2007-63683 and PR 2007-0295 (Spanish Government), E64

(DGA) and by FEDER funds.

References

[1] ADELL, J. A. AND DE LA CAL, J. (1993). On the uniform convergence of normalized Poisson mixtures to their mixing distribution, *Statist. Probab. Lett.* **18,** 227-232.

[2] ADELL, J. A. AND DE LA CAL, J. (1994). Approximating gamma distributions by normalized negative binomial distributions, *J. Appl. Probab.* **31,** 391-400.

[3] ADELL J. A. AND SANGÜESA C. (1999). *Direct and converse inequalities for positive linear operators on the positive semi-axis*, *J. Austral. Math. Soc. Ser. A* **66**, 90–103.

[4] ALZER, H. (1997) On some inequalities for the gamma and psi functions, *Math. Comp.*, **66**, 373-389.

[5] ASMUSSEN, S. (2000). *Ruin probabilities*, World Scientific, Singapore.

[6] EMBRECHTS, P., GRÜBEL, R. AND PITTS, S. M. (1993) Some applications of the fast Fourier transform algorithm in insurance mathematics, *Statist. Neerlandica*, **47**, 59-75.

[7] FELLER, W. (1971). *An Introduction to Probability Theory and its Applications*, Vol II, 2nd. edition. Wiley, New York.

[8] GRÜBEL, R. AND HERMESMEIER, R. (2000). Computation of compound distributions II: discretization errors and Richardson extrapolation. *Astin Bull.* **30,** 309–331.

[9] HIPP, C. (2006). Speedy convolution algorithms and Panjer recursions for phase-type distributions. *Insurance Math. Econom.* **38,** 176–188.

[10] Latouche, G. and Ramaswami, V. (1999). *Introduction to Matrix Analytic Methods in Stochastic Modelling*, ASA-SIAM, Philadelphia.

[11] Maier, R.S. (1991) The algebraic construction of phase-type distributions, *Comm. Statist. Stochastic Models*, **7**, 573-602.

[12] O'Cinneide, C. A. (1999) Phase-type distributions: open problems and a few properties., *Comm. Statist. Stochastic Models*, **15**, 731–757.

[13] Qi, F., Cui, R., Chen, C., Guo, B. (2005) Some completely monotonic functions involving polygamma functions and an application. *J. Math. Anal. Appl.* **310** , 303–308.

[14] Sangüesa, C. (2007). Error bounds in approximations of compound distributions using gamma-type operators, *To appear in Insurance Math. Econom.*

[15] Sundt, B. (2002). Recursive evaluation of aggregate claims distributions, *Insurance Math. Econom.* **30**, 297–322.

[16] Willmot, G. E. and Woo, J. K. (2007). On the class of Erlang mixtures with risk theoretical applications, *N. Am. Actuar. J.* **11,** 99–105

List of Recent Technical Reports

74. Xiaowen Zhou, *On a Classical Risk Model with a Constant Dividend Barrier*, November 2004

75. K. Balasubramanian and M.I. Beg, *Three Isomorphic Vector Spaces–II*, December 2004

76. Michael A. Kouritzin and Wei Sun, *Rates for Branching Particle Approximations of Continuous–Discrete Filters*, December 2004

77. Rob Kaas and Qihe Tang, *Introducing a Dependence Structure to the Occurences in Studying Precise Large Deviations for the Total Claim Amount*, December 2004

78. Qihe Tang and Gurami Tsitsiashvili, *Finite and Infinite Time Ruin Probabilities in the Presence of Stochastic Returns on Investments*, December 2004

79. Alexander Melnikov and Victoria Skornyakova, *Efficient Hedging Methodology Applied to Equity–Linked Life Insurance*, February 2005

80. Qihe Tang, *The Finite Time Ruin Probability of the Compound Poisson Model with Constant Interest Force*, June 2005

81. Marc J. Goovaerts, Rob Kaas, Roger J.A. Laeven, Qihe Tang and Raluca Vernic, *The Tail Probability of Discounted Sums of Pareto–Like Losses in Insurance*, August 2005

82. Yogendra P. Chaubey and Haipeng Xu, *Smooth Estimation of Survival Functions under Mean Residual Life Ordering*, August 2005

83. Xiaowen Zhou, *Stepping–Stone Model with Circular Brownian Migration*, August 2005

84. José Garrido and Manuel Morales, *On the Expected Discounted Penalty Function for Lévy Risk Processes*, November 2005

85. Ze–Chun Hu, Zhi–Ming Ma and Wei Sun, *Extensions of Lévy–Khintchine Formula and Beurling–Deny Formula in Semi–Dirichlet Forms Setting*, February 2006

86. Ze–Chun Hu, Zhi–Ming Ma and Wei Sun, *Formulae of Beurling–Deny and Lejan for Non-Symmetric Dirichlet Forms*, February 2006

87. Ze–Chun Hu and Wei Sun, *A Note on Exponential Stability of the Non-Linear Filter for Denumerable Markov Chains*, February 2006

88. H. Brito–Santana, R. Rodríguez–Ramos, R. Guinovart–Díaz, J. Bravo–Castillero and F.J. Sabina, *Variational Bounds for Multiphase Transversely Isotropic Composites*, August 2006

89. José Garrido and Jun Zhou, *Credibility Theory for Generalized Linear and Mixed Models*, December 2006

90. Daniel Dufresne, José Garrido and Manuel Morales, *Fourier Inversion Formulas in Option Pricing and Insurance*, December 2006

91. Xiaowen Zhou, *A Superprocess Involving Both Branching and Coalescing*, December 2006

92. Yogendra P. Chaubey, Arusharka Sen and Pranab K. Sen, *A New Smooth Density Estimator for Non–Negative Random Variables*, January 2007

93. Md. Sharif Mozumder and José Garrido, *On the Relation between the Lévy Measure and the Jump Function of a Lévy Process*, October 2007

94. Arusharka Sen and Winfried Stute, *A Bi-Variate Kaplan-Meier Estimator via an Integral Equation*, October 2007

95. C. Sangüesa, *Uniform Error Bounds in Continuous Approximations of Nonnegative Random Variables Using Laplace Transforms*, January 2008

Copies of technical reports can be requested from:

Dr. Wei Sun
Department of Mathematics and Statistics
Concordia University
1455 de Maisonneuve Blvd. West,
Montreal, QC, H3G 1M8 CANADA