

Full Length Article



# Orthogonal gamma-based expansion for the CIR's first passage time distribution

Elvira Di Nardo <sup>a,\*</sup>, Giuseppe D'Onofrio <sup>b</sup>, Tommaso Martini <sup>a</sup>

<sup>a</sup> Dipartimento di Matematica "G. Peano", Università degli Studi di Torino, Via Carlo Alberto 10, 10123 Torino, Italy

<sup>b</sup> Dipartimento di Scienze Matematiche, Politecnico di Torino, 10129 Torino, Italy

## ARTICLE INFO

MSC:

65C20

60G07

62E17

42C10

60-08

Keywords:

Feller square-root process

Hitting times

Fourier series expansion

Cumulants

Laguerre polynomials

Acceptance-rejection method

## ABSTRACT

In this paper we analyze a method for approximating the first-passage time density and the corresponding distribution function for a CIR process. This approximation is obtained by truncating a series expansion involving the generalized Laguerre polynomials and the gamma probability density. The suggested approach involves a number of numerical issues which depend strongly on the coefficient of variation of the first passage time random variable. These issues are examined and solutions are proposed also involving the first passage time distribution function. Numerical results and comparisons with alternative approximation methods show the strengths and weaknesses of the proposed method. A general acceptance-rejection-like procedure, that makes use of the approximation, is presented. It allows the generation of first passage time data, even if its distribution is unknown.

## 1. Introduction

In many applications spanning from finance to engineering, including, among others, computational neuroscience, mathematical biology and reliability theory (see [43] for a thorough exposition), the dynamics of a noisy system is described by a stochastic process  $Y(t)$  evolving in the presence of a threshold  $S(t)$ . The first-passage-time (FPT) problem consists in finding the distribution of the random variable (rv)  $T$ , defined by

$$T = \begin{cases} \inf_{t \geq \tau} \{Y(t) > S(t)\}, & Y(\tau) = y_\tau < S(\tau), \\ \inf_{t \geq \tau} \{Y(t) < S(t)\}, & Y(\tau) = y_\tau > S(\tau), \end{cases} \quad (1.1)$$

representing the time the process  $Y(t)$  crosses the threshold  $S(t)$  for the first time. Although classical and very easy to state, the solution in closed form of this problem is available only in a very few cases, depending on the properties of both  $Y(t)$  and  $S(t)$ . In this paper, we address the FPT problem of a Cox-Ingersoll-Ross (CIR) process  $Y(t)$  through a constant threshold  $S$ . This one-dimensional diffusion process, belonging to the class of Pearson diffusions [20], is frequently involved in the field of mathematical finance starting from the seminal paper [9] by which it is commonly called nowadays. Outside this community the process is often called square-root,

\* Corresponding author.

E-mail addresses: [elvira.dinardo@unito.it](mailto:elvira.dinardo@unito.it) (E. Di Nardo), [giuseppe.donofrio@polito.it](mailto:giuseppe.donofrio@polito.it) (G. D'Onofrio), [tommaso.martini@unito.it](mailto:tommaso.martini@unito.it) (T. Martini).

due to the form of its diffusion coefficient (or volatility), or, due to historical reasons, Feller process from the 1951 paper, in which the process is introduced for the first time [18].

When considering the FPT problem of a CIR process, the literature, also quite recent, is vast but the results are partial and fragmentary, see for instance [2], [10], [21], [22], [23], [24], [32], [34], [49], and [35] for a thorough review of the state of art. Exploiting the Laplace transform of the FPT pdf and the theory of formal power series, a closed form expression of its FPT cumulants is given in [11]. In the same paper, under appropriate assumptions, the FPT pdf has been expanded in series of generalized Laguerre polynomials, involving moments computed from cumulants and weighted by a gamma pdf. The idea of approximating a pdf by truncating a suitable series expansion is not new. Indeed such an approximation is of the Gram-Charlier type with a gamma distribution rather than a normal distribution as reference (parent) distribution and with generalized Laguerre polynomials instead of Hermite polynomials as multipliers. In [40] a general methodology to approximate a pdf based on the knowledge of its moments is introduced, using the product of a suitable weight function, as parent distribution, and a suitable family of associated orthogonal polynomials, as multipliers. One of the main issues of this approximation is that negative values can occur, although the approximate density always entails a unit area. This happens even with the Gram-Charlier series having Gaussian parent distribution. To overcome this drawback, two approaches can be found in the literature. The first one is to use the approximation with a low truncation order and to find constrained regions on the values of the cumulants (or moments) that admit a valid (non-negative) pdf. The suggested truncation is mostly at the fourth-order term because it becomes difficult to manage valid regions for higher orders. Within the FPT framework, this approach was used to approximate the FPT pdf of an Ornstein-Uhlenbeck process [48]. In particular, the restrictions on the first four moments that guarantee the non-negativity of the approximated density are outlined and discussed in [33], along with a thorough examination of when to apply this approximation. Indeed, with this low truncation order, the approximated density may fail to be close to the theoretical one, especially for distributions that are not sufficiently close to the parent distribution. In the case of Gaussian parent distribution and for arbitrary even order, the valid region of cumulants has been found numerically through a semi-definite algorithm [31]. A second way of tackling this issue consists in replacing values of a suitable positive interpolating function to the negative ones assumed by the approximated pdf. In [51], as interpolating function for pdfs with support  $(0, \infty)$ , a second-degree polynomial is suggested in a right-handed neighborhood of the origin. In this paper, taking into account that the FPT pdfs are unimodal for diffusion processes [46], this second approach is developed along a different direction and for the first time within the FPT framework. Firstly, sufficient conditions are given on the sign of the coefficients of the series expansion so that the approximated density may hold non-negative values on the tails and in a right-handed neighborhood of the origin. Secondly, if there are additional intervals in which the approximated pdf turns out to be negative, an appropriate correction is proposed that takes into account sufficient conditions given in [11] allowing the Laguerre-Gamma type expansion for the FPT pdf. The issue concerning the possible negative values of the approximated pdf can be overcome also by considering the FPT cumulative distribution function (cdf). For this reason, in parallel with our discussion, we develop the method for the approximation of the cdf as well. We stress that this approach, new in the FPT context, has some numerical advantages and allows an easier approximation of quantiles.

Those highlighted so far are not the only issues concerning the use of such an approximation. The choice of the gamma pdf parameters as well as the order of truncation of the series are additional issues that may affect the quality of the approximation. These issues, only briefly sketched in [40], are considered in detail in this paper. For example, the key role played by the coefficient of variation in the choice of the gamma pdf parameters is shown, while the truncation order is controlled by appropriate stopping criteria.

The proposed Laguerre-Gamma expansion has an additional advantage since density estimates can be produced based on sample moments. Indeed, if the FPT moments/cumulants are not known, this approach allows to recover an approximation of the FPT pdf starting from a sample of FPT data. These estimators are known in the literature as orthogonal series estimators and can be very competitive when compared with the classical density estimators such as the kernel density estimator (KDE) or the histogram [26].

Thanks to obtained evaluations of the approximation error, an acceptance-rejection-like method, that makes use of the series expansion, is finally proposed. It allows the generation of FPT data, even if its distribution is unknown, and can be applied to a wide class of pdfs. Note that, although never used for the CIR process, acceptance-rejection methods had already appeared in the FPT context (see for instance [27] and [36]), but the approximation strategy here proposed is a novelty. This method is particularly useful since exact simulation techniques for CIR sample paths are not available, and the existing ones, based on discretization methods or transition densities, exhibit large computational costs if the fixed time step is small.

The paper is organized as follows. In Section 2 we resume the FPT problem for the CIR process recalling the known results useful for carrying out the proposed approximation. In Section 3 we discuss the convergence of the method. Moreover we address some theoretical issues closely related to the approximation such as the choice of the truncation order. The role played by the coefficient of variation of the FPT rv in the choice of the gamma pdf parameters is also discussed. Section 4 suggests how to overcome the two main computational issues arising in dealing with such an approximation: the monotonicity of the approximated cdf and the positivity of the approximated pdf. Numerical results and comparisons with alternative approximation methods are given in Section 5 aiming to discuss the strengths and weaknesses of the proposed approach. We set three different choices of the CIR process parameters and boundaries that corresponds to different forms and statistical properties of the FPT pdf. An application of the Laguerre-Gamma approximation is shown in the last section, which involves sampling FPTs using a technique analogous to the acceptance-rejection method. Concluding remarks close the paper.

## 2. The CIR process and the FPT problem

The CIR process we refer to is the unique strong solution of the stochastic differential equation [18]

$$dY(t) = (-\tau Y(t) + \mu) dt + \sigma \sqrt{Y(t) - c} dW(t), \tag{2.1}$$

where  $W(t)$  is a standard Brownian motion,  $c \leq 0$ ,  $\tau > 0$ ,  $\mu \in \mathbb{R}$ ,  $\sigma > 0$  and  $Y_0 = y_0$ . The state space of the process is the interval  $(c, +\infty)$ . The endpoints  $c$  and  $+\infty$  can or cannot be reached in a finite time, depending on the underlying parameters. According to the Feller classification of boundaries [28],  $c$  is an entrance boundary if it cannot be reached by  $Y(t)$  in finite time, and there is no probability flow to the outside of the interval  $(c, +\infty)$ . In particular,

$$c \text{ is an entrance boundary if } s := 2(\mu - c\tau)/\sigma^2 \geq 1.$$

This will be a standing assumption in the following.

Denote with  $g(t) = \frac{d}{dt} \mathbb{P}\{T \leq t\}$  the pdf of the FPT rv  $T$  as defined in (1.1). Its Laplace transform  $\tilde{g}(z)$  is such that  $\tilde{g}(z) = 1$  if  $y_0 \equiv S$  and  $\tilde{g}(z) < +\infty$  for any different  $y_0$  [35]. Its closed form expression is [16]

$$\tilde{g}(z) = \frac{\Phi\left(\frac{z}{\tau}, s, \frac{2\tau(y_0 - c)}{\sigma^2}\right)}{\Phi\left(\frac{z}{\tau}, s, \frac{2\tau(S - c)}{\sigma^2}\right)}, \quad z > 0 \tag{2.2}$$

where  $\Phi(a, b, z) = {}_1F_1(a; b; z)$  is the confluent hypergeometric function of the first kind (or Kummer’s function). The Laplace transform (2.2) cannot be inverted explicitly, except for the case  $S = 0$ , see for instance [34], but information on the moments can be obtained by direct derivation or from cumulants as described in the next subsection.

### 2.1. FPT cumulants and moments

Recall that, if  $T$  has moment generating function  $\mathbb{E}[e^{zT}] < \infty$  for all  $z$  in an open interval about 0, then its cumulants  $\{c_k(T)\}_{k \geq 1}$  are such that

$$\sum_{k \geq 1} c_k(T) \frac{z^k}{k!} = \log \mathbb{E}[e^{zT}]$$

for all  $z$  in some (possibly smaller) open interval about 0. Using the logarithmic polynomials<sup>1</sup>  $\{P_k\}$ , the FPT cumulants of the CIR process can be expressed as [11]

$$c_k(T) = (-\tau)^{-k} [c_k^*(y_0) - c_k^*(S)], \quad k \geq 1$$

where

$$c_k^*(w) = P_k \left[ h_1 \left( \frac{2\tau(w - c)}{\sigma^2} \right), h_2 \left( \frac{2\tau(w - c)}{\sigma^2} \right), \dots, h_k \left( \frac{2\tau(w - c)}{\sigma^2} \right) \right], \tag{2.3}$$

with  $h_j(y) = j! \sum_{n \geq j} \binom{n}{j} \frac{y^n}{n! \langle s \rangle_n}$ , for  $j = 1, 2, \dots, k$ ,  $\binom{n}{j}$  the unsigned Stirling numbers of first type and  $\langle \cdot \rangle_n$  the  $n$ -th rising factorial.

FPT moments of the CIR process are obtained from cumulants [11] using the complete Bell polynomials<sup>2</sup>  $\{Y_k\}$  and  $\{c_k^*\}$  given in (2.3), that is

$$\mathbb{E}[T^k] = \frac{(-1)^k}{\tau^k} \sum_{i=0}^k \binom{k}{i} Y_{k-i} [c_1^*(y_0), \dots, c_{k-i}^*(y_0)] Y_i [-c_1^*(S), \dots, -c_i^*(S)] \tag{2.4}$$

for  $k \geq 1$ . An alternative way to compute moments from cumulants is the well-known recursion formula [14]

$$\mathbb{E}[T^k] = c_k(T) + \sum_{i=1}^{k-1} \binom{k-1}{i-1} c_i(T) \mathbb{E}[T^{k-i}]. \tag{2.5}$$

This formula is particularly convenient from a computational point of view and has been used to recover FPT moments from the knowledge of cumulants.

### 2.2. The FPT pdf and cdf

Under suitable hypotheses, a closed form expression of the FPT pdf has been given in [11] using the moments  $\{\mathbb{E}[T^k]\}_{k \geq 1}$ . Indeed, suppose

<sup>1</sup> See Appendix for their definition.

<sup>2</sup> See Appendix for their definition.

$$f_{\alpha,\beta}(t) = \beta(\beta t)^\alpha \frac{e^{-\beta t}}{\Gamma(\alpha + 1)}, \quad t > 0 \tag{2.6}$$

the gamma pdf with scale parameter  $\alpha + 1 > 0$  and shape parameter  $\beta > 0$ . For  $\alpha > -1$ , let the polynomial sequence  $\{Q_k^{(\alpha)}(t)\}_{k \geq 0}$  be defined as

$$Q_k^{(\alpha)}(t) = (-1)^k \left( \frac{\Gamma(\alpha + 1 + k)}{k! \Gamma(\alpha + 1)} \right)^{-1/2} L_k^{(\alpha)}(t), \tag{2.7}$$

where  $L_k^{(\alpha)}(t)$  is the  $k$ -th generalized Laguerre polynomial

$$L_k^{(\alpha)}(t) = \sum_{i=0}^k \binom{k + \alpha}{k - i} \frac{(-t)^i}{i!}, \quad k \geq 1$$

with  $L_0^{(\alpha)}(t) = 1$ . For any fixed  $t > 0$ , the FPT pdf admits the following expansion (see Theorem 2 in [11]):

$$g(t) = \frac{\beta(\beta t)^\alpha e^{-\beta t}}{\Gamma(\alpha + 1)} \sum_{k \geq 0} a_k^{(\alpha)} Q_k^{(\alpha)}(\beta t), \tag{2.8}$$

where  $a_k^{(\alpha)} = \mathbb{E}[Q_k^{(\alpha)}(\beta T)]$  for  $k \geq 0$ .

**Remark 2.1.** If

$$\int_0^\infty t^{-\alpha} e^{\beta t} [g(t)]^2 < \infty \quad \text{then} \quad \frac{g(t)}{f_{\alpha,\beta}(t)} \in \mathcal{L}^2(\nu),$$

where  $\mathcal{L}^2(\nu)$  is the Hilbert space of the square-integrable functions with respect to the measure  $\nu$  having density  $f_{\alpha,\beta}(t)$ . Therefore, (2.8) represents the Fourier-Laguerre series expansion of  $g(t)/f_{\alpha,\beta}(t)$  in terms of the complete orthonormal sequences  $\{Q_k^{(\alpha)}(t)\}$ .

Some algebra allows us to write the FPT pdf  $g(t)$  in (2.8) as

$$g(t) = \beta(\beta t)^\alpha e^{-\beta t} \sum_{k \geq 0} B_k^{(\alpha)} L_k^{(\alpha)}(\beta t), \quad t > 0 \tag{2.9}$$

with coefficients  $B_0^{(\alpha)} = 1$  and

$$B_k^{(\alpha)} = (-1)^k a_k^{(\alpha)} \left( \frac{\Gamma(\alpha + 1)\Gamma(\alpha + 1 + k)}{k!} \right)^{-1/2} = \sum_{j=0}^k \binom{k}{j} \frac{(-\beta)^j \mathbb{E}(T^j)}{\Gamma(\alpha + j + 1)}, \tag{2.10}$$

depending on the moments of  $T$ . By recalling that

$$\int \tau^\alpha e^{-\tau} L_k^{(\alpha)}(\tau) d\tau = \frac{\tau^{\alpha+1} \Gamma(\alpha + k + 1)}{k!} \frac{\Phi(\alpha + k + 1; \alpha + 2; -\tau)}{\Gamma(\alpha + 2)}$$

where  $\Phi(a, b, z) = {}_1F_1(a; b; z)$  is the confluent hypergeometric function of the first kind, and using (2.9), a closed form expression of the FPT cdf is given in the following statement.

**Proposition 2.2.** The FPT cdf  $G(t)$  is

$$G(t) = \frac{(\beta t)^{\alpha+1}}{\Gamma(\alpha + 2)} \sum_{k \geq 0} \frac{\Gamma(\alpha + k + 1)}{k!} B_k^{(\alpha)} \Phi(\alpha + k + 1, \alpha + 2, -\beta t), \quad t > 0.$$

### 3. The FPT approximation

An approximation of the FPT pdf can be recovered from (2.8) by using a truncation of the series up to an order  $n$

$$\hat{g}_n(t) = \frac{\beta(\beta t)^\alpha e^{-\beta t}}{\Gamma(\alpha + 1)} \sum_{k=0}^n a_k^{(\alpha)} Q_k^{(\alpha)}(\beta t), \quad t > 0. \tag{3.1}$$

The higher is the order  $n$  the better should be the approximation. Indeed the  $\mathcal{L}^2(\nu)$ -error in replacing  $g(t)$  with its approximation  $\hat{g}_n(t)$  given in (3.1) is [47]

$$\left\| \frac{g - \hat{g}_n}{f_{\alpha,\beta}} \right\|_{\alpha,\beta} = \left[ \sum_{k \geq n+1} (a_k^{(\alpha)})^2 \right]^{1/2} \tag{3.2}$$

where  $\|\cdot\|_{\alpha,\beta}$  denotes the norm in  $\mathcal{L}^2(\nu)$ . Thus the error may be estimated by calculating the rate of decrease of  $a_k^{(\alpha)}$  when  $k \rightarrow \infty$ . The latter is given in the following proposition.

**Theorem 3.1.** Assume the FPT pdf  $g(t) \in C^2[0, +\infty)$ . If

$$\beta < \frac{2}{\mathbb{E}[T]} \quad \text{and} \quad g(t) = o(t^\delta) \quad \text{for } t \rightarrow 0, \quad \text{with } \delta > \frac{\alpha}{2} + 1$$

then in (3.2)  $a_k^{(\alpha)} = \mathbb{E}[Q_k^{(\alpha)}(\beta T)] = O(k^{-1})$  as  $k \rightarrow \infty$ .

**Proof.** Observe that  $a_k^{(\alpha)} = \mathbb{E}[Q_k^{(\alpha)}(\beta T)]$  gives

$$a_k^{(\alpha)} = \frac{1}{\beta} \int_0^\infty Q_k^{(\alpha)}(t) \tilde{g}_{\alpha,\beta}(t) t^\alpha e^{-t} dt \quad \text{where} \quad \tilde{g}_{\alpha,\beta}(t) = \frac{g(t/\beta)}{t^\alpha e^{-t}}. \tag{3.3}$$

As the generalized Laguerre polynomials  $\{L_k^{(\alpha)}(t)\}$  are eigenfunctions of a Sturm-Liouville problem [1] with associated eigenvalues  $\lambda_k = k$

$$\frac{d}{dt} (t^{\alpha+1} e^{-t} y') + k t^\alpha e^{-t} y = 0 \quad \text{with } y = y(t), k \geq 1$$

the same happens for the linearly transformed polynomials  $\{Q_k^{(\alpha)}(t)\}$  in (2.7). Therefore in (3.3), replace

$$Q_k^{(\alpha)}(t) t^\alpha e^{-t} \quad \text{with} \quad -\frac{1}{k} \frac{d}{dt} (t^{\alpha+1} e^{-t} y').$$

Integrating by parts the integral in (3.3) and neglecting the constants, the rhs of (3.3) reads

$$a_k^{(\alpha)} \approx \frac{1}{k} \int_0^\infty t^{\alpha+1} e^{-t} \frac{d}{dt} [Q_k^{(\alpha)}(t)] \frac{d}{dt} [\tilde{g}_{\alpha,\beta}(t)] dt. \tag{3.4}$$

Indeed we have

$$\lim_{t \rightarrow 0} \tilde{g}_{\alpha,\beta}(t) t^{\alpha+1} e^{-t} \frac{d}{dt} [Q_k^{(\alpha)}(t)] = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} \tilde{g}_{\alpha,\beta}(t) t^{\alpha+1} e^{-t} \frac{d}{dt} [Q_k^{(\alpha)}(t)] = 0. \tag{3.5}$$

The first limit in (3.5) results by the hypothesis  $g(t) = o(t^\delta)$  for  $t \rightarrow 0$ . The second limit in (3.5) follows by taking into account that the FPT pdf of one-dimensional diffusion processes with steady-state distribution is known to be approximately exponential for  $t \rightarrow \infty$ , with parameter  $\mathbb{E}[T]^{-1}$  [37]. Integrating by parts the integral in (3.4) and neglecting the constants, the rhs of (3.4) reads

$$a_k^{(\alpha)} \approx -\frac{1}{k} \int_0^\infty Q_k^{(\alpha)}(t) \frac{d}{dt} \left( t^{\alpha+1} e^{-t} \frac{d}{dt} [\tilde{g}_{\alpha,\beta}(t)] \right) dt, \tag{3.6}$$

where similar considerations done for (3.5) apply for recovering

$$\lim_{t \rightarrow 0} t^{\alpha+1} e^{-t} Q_k^{(\alpha)}(t) \frac{d}{dt} [\tilde{g}_{\alpha,\beta}(t)] = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} t^{\alpha+1} e^{-t} Q_k^{(\alpha)}(t) \frac{d}{dt} [\tilde{g}_{\alpha,\beta}(t)] = 0.$$

Now, in (3.6) set

$$h(t) = \frac{1}{w(t)} \frac{d}{dt} \left( t^{\alpha+1} e^{-t} \frac{d}{dt} [\tilde{g}_{\alpha,\beta}(t)] \right) \quad \text{with} \quad w(t) = t^\alpha e^{-t}.$$

Applying the Cauchy-Schwarz inequality to the rhs of (3.6), we get

$$\left| \int_0^\infty \sqrt{w(t)} Q_k^{(\alpha)}(t) h(t) \sqrt{w(t)} dt \right|^2 \leq \left( \int_0^\infty w(t) [Q_k^{(\alpha)}(t)]^2 dt \right) \left( \int_0^\infty [h(t)]^2 w(t) dt \right)$$

which is finite and not depending on the order  $k$ , if the same is true for both integrals on the lhs. Observe that the first integral corresponds to the orthonormality condition of  $Q_k^{(\alpha)}(t)$  and so it is finite and not depending on  $k$ . The second integral does not depend on  $k$  and is finite if the integrand is smooth and the limits for  $t \rightarrow 0$  and  $t \rightarrow \infty$  are finite. Note that

$$\frac{h(t)}{\sqrt{w(t)}} = \frac{(\alpha + 1 - t) \frac{d}{dt} [\tilde{g}_{\alpha,\beta}(t)] + t \frac{d^2}{dt^2} [\tilde{g}_{\alpha,\beta}(t)]}{t^{-\alpha/2} e^{t/2}}.$$

Thus for  $t \rightarrow \infty$  and  $\beta < 2/\mathbb{E}[T]$  we have

$$\lim_{t \rightarrow \infty} \frac{h(t)}{\sqrt{w(t)}} = \lim_{t \rightarrow \infty} \frac{e^{z[1-(\beta E[t])^{-1}]}}{z^{\alpha/2} e^{z/2}} + \lim_{t \rightarrow \infty} \frac{e^{z[1-(\beta E[t])^{-1}]}}{z^{\alpha/2-1} e^{z/2}} = 0$$

assuming  $g(t) = O(e^{-t/(\beta E[T])})$ . Instead for  $t \rightarrow 0$  the limit reduces to

$$\lim_{t \rightarrow 0} \frac{h(t)}{\sqrt{w(t)}} = \lim_{t \rightarrow 0} e^{t/2} t^{\delta-\alpha/2-1} + \lim_{t \rightarrow 0} e^{t/2} t^{\delta-\alpha} = 0$$

if  $\delta > 1 + \alpha/2$ .  $\square$

The request of the existence of the second derivative in Theorem 3.1 is a reasonable assumption, since the property can be seen as a consequence of the following observation.

**Remark 3.2.** Following [38], we could have alternatively asked that there exists  $\varepsilon > 0$  such that  $\sigma\sqrt{x} \geq \varepsilon$ , for all  $x$  in the state space of the process. This condition implies, in this case, the existence and boundedness of at least the first two derivatives of the FPT pdf  $g(t)$ . To investigate Pauwels’ condition, one can follow Feller’s classification of the boundaries [19]. Using the transition densities of the Feller process [35], one has to show that the flux through the value  $\varepsilon$  is zero or that the capacity of the interval  $[0, \varepsilon)$  vanishes [4]. For a fixed, small  $\varepsilon > 0$  we observe, at least numerically, that the capacity of the interval  $[0, \varepsilon)$  (see formula 19 in [4]) goes to zero as  $\mu$  increases. This means that for a “large enough” choice of  $\mu$ , the mentioned assumption in Theorem 3.1 is satisfied.

By truncating the series in (2.9) up to the order  $n$ , the approximated  $\hat{g}_n(t)$  in (3.1) can be rewritten as

$$\hat{g}_n(t) = f_{\alpha,\beta}(t)p_n(t) \quad \text{where} \quad p_n(t) = \sum_{k=0}^n h_{n,k} \frac{(-\beta t)^k}{k!} \tag{3.7}$$

with

$$h_{n,k} = \sum_{j=k}^n B_j^{(\alpha)} \binom{\alpha+j}{j-k} \quad \text{and} \quad \binom{\alpha+j}{j-k} = \begin{cases} 1, & j = k, \\ \frac{(\alpha+j)(\alpha+j-1)\dots(\alpha+k+1)}{(j-k)!}, & j > k. \end{cases}$$

In particular  $\{h_{n+1,i}\}_{i=0}^{n+1}$  can be recovered from  $\{h_{n,i}\}_{i=0}^n$  using the following recursion formula [12]

$$h_{n+1,i} = \begin{cases} B_{n+1}^{(\alpha)}, & \text{for } i = n + 1, \\ h_{n,i} + B_{n+1}^{(\alpha)} \binom{\alpha+n+1}{n+1-i}, & \text{for } i = 0, \dots, n. \end{cases} \tag{3.8}$$

Likewise, the coefficients  $\{B_{n+1}^{(\alpha)}\}$  can be recovered from  $\{B_j^{(\alpha)}\}_{j=0}^n$  as

$$B_{n+1}^{(\alpha)} = \sum_{j=1}^{n+1} \binom{n+1}{j} (-1)^{j+1} B_{n+1-j}^{(\alpha)} + \frac{(-\beta)^{n+1} E[T^{n+1}]}{(\alpha+n+1)_{n+1}}. \tag{3.9}$$

Due to the orthogonality property of generalized Laguerre polynomials, the approximation  $\hat{g}_n(t)$  has nice properties, that are:

$$\int_0^\infty \hat{g}_n(t) dt = 1 \tag{3.10}$$

for all  $n \geq 0$ , and the first  $n$  moments of  $\hat{g}_n(t)$  are the same of  $g(t)$ . Unfortunately,  $\hat{g}_n(t)$  is not guaranteed to be a pdf since negative values can occur. Indeed the values assumed by the polynomial  $p_n(t)$  are not necessarily non-negative. However  $p_n(t)$  may hold non-negative values on the tails and in a right-handed neighborhood of the origin, depending on the sign of some coefficients in (3.7). These conditions are established in the following proposition.

**Proposition 3.3.** Suppose  $p_n(t) > 0$  for all  $t > 0$ . Then  $(-1)^n h_{n,n} \geq 0$  and  $h_{n,0} \geq 0$ . Conversely, if  $h_{n,0} > 0$  and  $(-1)^n h_{n,n} > 0$ , there exist  $t_1 > 0$  and  $t_2 > 0$  such that  $p_n(t) > 0$  in  $(0, t_1) \cup (t_2, +\infty)$ , with  $t_1$  and  $t_2$  not necessarily distinct or finite.

**Proof.** From (3.7) we have

$$p_n(0) = h_{n,0} \quad \text{and} \quad p_n(t) \sim (-1)^n h_{n,n} t^n \frac{\beta^n}{n!}, \quad \text{for } t \rightarrow \infty.$$

Since  $\beta > 0$  and  $p_n(t) \in C(0, \infty)$ , the results follow from the sign permanence theorem.  $\square$

Integrating  $\hat{g}_n(t)$  in (3.7) over  $(0, t)$ , an approximation of the FPT is

$$\hat{G}_n(t) = \frac{1}{\Gamma(\alpha + 1)} \sum_{k=0}^n \frac{(-1)^k}{k!} h_{n,k} [\Gamma(\alpha + k + 1) - \Gamma(\alpha + k + 1, \beta t)] \tag{3.11}$$

where  $\Gamma(a, t) = \int_t^\infty \tau^{a-1} e^{-\tau} d\tau$  is the incomplete Gamma function.

### 3.1. On the order $n$ of the approximation

The normalization condition (3.10) has been used to derive a first stopping criterion [12]. Indeed (3.10) is equivalent to (Proposition 4.2 [12])

$$h_{n,0} + \sum_{i=1}^n \frac{(-1)^i}{i!} h_{n,i}(\alpha + i)_i = 1. \tag{3.12}$$

As a result, in (3.7) the order  $n$  is increased as long as (3.12) is satisfied with a fixed level of tolerance. Here, with the more conservative aim of obtaining reliable approximations, we propose to join the normalization condition (3.12) with an additional condition following from Proposition 3.3. This condition guarantees an order  $n$  of approximation such that  $p_n(t)$  is positive close to the origin and as  $t \rightarrow +\infty$ . Therefore, the recursive procedure runs as long as

$$(|\hat{h}_{n+1} - 1| < \varepsilon \text{ for a fixed } \varepsilon > 0) \text{ or } (h_{n,0} > 0 \text{ and } (-1)^n h_{n,n} > 0) \tag{3.13}$$

is fulfilled.

### 3.2. On the choice of $\alpha$ and $\beta$

Denote with  $X_{\alpha,\beta}$  the rv having pdf  $f_{\alpha,\beta}$  in (2.6). In [11] we have chosen

$$\alpha = \frac{c_1^2[T]}{c_2[T]} - 1 \quad \text{and} \quad \beta = \frac{c_1[T]}{c_2[T]} \tag{3.14}$$

since with these choices we have  $\mathcal{B}_1^{(\alpha)} = \mathcal{B}_2^{(\alpha)} = 0$  in (2.10) and

$$\mathbb{E}[X_{\alpha,\beta}] = \frac{\alpha + 1}{\beta} = \mathbb{E}[T] \quad \text{and} \quad \mathbb{E}[X_{\alpha,\beta}^2] = \frac{(\alpha + 1)(\alpha + 2)}{\beta^2} = \mathbb{E}[T^2].$$

Let us underline that a range of values was investigated for  $\alpha$  and  $\beta$ . Actually the choices in (3.14) seem to be the most reliable with respect to the selection of parameters in (2.1). On the other hand, this choice falls within the classical method of moments and is also suggested in [40], where a general procedure is developed for the approximation of a pdf based on its moments. Different choices are suggested in [3] where results concerning the determination of the two parameters  $\alpha$  and  $\beta$  are presented. Unfortunately, adopting these choices requires a knowledge of the FPT pdf not depending on  $\alpha$  and  $\beta$ , which is not true in the case of CIR process. Moreover, the choices of  $\alpha$  and  $\beta$  in (3.14) return

$$c_v[X_{\alpha,\beta}] = c_v[T] = \frac{\sqrt{c_2[T]}}{c_1[T]}$$

where  $c_v$  denotes the coefficient of variation. In such a case the first equation in (3.14) reads  $\alpha + 1 = (c_v[T])^{-2}$ . Thus an higher coefficient of variation reduces  $\alpha + 1$  and increases the chance of a vertical asymptote of the gamma pdf in 0. As  $g(0) = 0$ , the occurrence of this vertical asymptote represents a numerical issue which is further worsened if the FPT pdf is flat with a large mean value and a heavy right tail with a large variance. To deal with this scenario, a successful strategy is the employment of a suitable standardization technique. The idea is to construct the approximation  $\tilde{g}_n(t)$  of the pdf  $\tilde{g}(t)$  corresponding to

$$\tilde{T} = T / \sigma_T, \tag{3.15}$$

where  $\sigma_T$  is the standard deviation of  $T$ . The approximated FPT pdf and cdf can be recovered as

$$\hat{g}_n(t) = \frac{1}{\sigma_T} \tilde{g}_n\left(\frac{t}{\sigma_T}\right) \quad \text{and} \quad \hat{G}_n(t) = \tilde{G}_n\left(\frac{t}{\sigma_T}\right)$$

respectively. As  $c_2[\tilde{T}] = \text{Var}[\tilde{T}] = 1$ , from (3.14) the parameters of the gamma pdf are

$$\tilde{\alpha} := (\mathbb{E}[\tilde{T}])^2 - 1 \quad \text{and} \quad \tilde{\beta} := \mathbb{E}[\tilde{T}]$$

so that  $\tilde{\alpha} + 1 = \tilde{\beta}^2 = (c_v[X_{\tilde{\alpha},\tilde{\beta}}])^{-2} = (c_v[\tilde{T}])^{-2}$  and  $\text{Var}[X_{\tilde{\alpha},\tilde{\beta}}] = 1$ . The advantage of this strategy is to use an initial guess (the gamma pdf) with a shape resembling the picked shape of the FPT pdf and concentrating probability mass on small time values. Moreover, the moments  $\mathbb{E}[\tilde{T}^n]$  grow slower than the moments  $\mathbb{E}[T^n]$ , leading to an observed improvement of numerical stability.

Since the pdf  $g(t)$  has support  $(0, \infty)$ , further information on the shape of the density can be recovered using some dispersion indices that work as the coefficient of variation but provide further global statistical information [29]. In particular we consider the relative entropy based dispersion coefficient

$$c_h := \frac{\sigma_h}{\mathbb{E}(T)} = \frac{1}{\mathbb{E}(T)} \exp \left\{ \int_0^\infty g(t) \ln g(t) dt - 1 \right\}. \tag{3.16}$$

The value of  $\sigma_h$  quantifies how evenly is the pdf over  $(0, \infty)$ . Moreover, in logarithmic scale,  $c_h$  is inversely proportional to the Kullback-Leibler distance of the pdf  $g(t)$  from the exponential density with mean  $\mathbb{E}[T]$ . For densities resembling the exponential distribution, the coefficients  $c_v$  and  $c_h$  are approximately equal to 1.

#### 4. Computational issues

In [12] a fast recursive procedure for implementing (3.7) is proposed, relying on nested products and taking advantage of (3.8) and (3.9). Note that, differently from the geometric Brownian motion in [12], in the case of the CIR process an additional issue arises in computing (3.9). The FPT moments are calculated from the FPT cumulants (2.3) using the recursion (2.5). Because of the series involved in (2.3), an approximation order  $m$  should be chosen before computing the moments through (2.5). Here, a standard approach has been used, which involves computing partial sums of the series as long as their difference exceeds an input tolerance.

##### 4.1. On the monotonicity of $\hat{G}_n$

For a fixed  $\Delta t > 0$ , the computation of the FPT cdf can benefit from the iterative calculation of increments

$$\Delta \hat{G}_n(t) = \frac{1}{\Gamma(\alpha + 1)} \sum_{k=0}^n \frac{(-1)^k}{k!} h_{n,k} [\Gamma(\alpha + k + 1, \beta t) - \Gamma(\alpha + k + 1, \beta(t + \Delta t))] \tag{4.1}$$

where  $\Delta \hat{G}_n(t) = \hat{G}_n(t + \Delta t) - \hat{G}_n(t), t > 0$ . Note that the increments  $\Delta \hat{G}_n(t)$  might be negative, in accordance with the values of  $t$  where  $\hat{g}_n(t) < 0$ . As a by-product, the approximated cdf may turn out to be decreasing in these intervals. Moreover, if in a right-handed neighborhood of the origin the first increments are already negative, this circumstance might determine negative values of the approximated cdf itself in the same neighborhood. A possible correction to this last drawback is: set  $\tau_0 = 0$  or  $\tau_0 = \min\{t > 0 | \Delta \hat{G}_n(t) < 0\}$ , depending if  $\Delta \hat{G}_n(\Delta t)$  is negative or not, and  $\tau_1 = \min\{t > \tau_0 | \hat{G}_n(t) > \hat{G}_n(\tau_0)\}$ . Then iteratively find the intervals  $[\tau_i, \tau_{i+1}]$  such that

$$\tau_i = \min\{t > \tau_{i-1} | \Delta \hat{G}_n(t) < 0\} \quad \text{and} \quad \tau_{i+1} = \min\{t > \tau_i | \hat{G}_n(t) > \hat{G}_n(\tau_i)\}$$

in order to replace  $\hat{G}_n(t)$  with a suitable line for  $t \in [\tau_i, \tau_{i+1}]$ . For the cases here considered the intervals  $t \in [\tau_i, \tau_{i+1}]$  have a small amplitude. The advantage of this approach is twofold. The first one is getting an approximated cdf which is positive and increasing. The second advantage is the chance to use an ad hoc numerical procedure to carry out the derivative of the resulting cdf and thereby automatically recovering an approximation of the pdf which is always positive. Since the corrected  $\hat{G}_n(t)$  is linear for  $t \in [\tau_i, \tau_{i+1}]$ , the corresponding pdf will be constant in  $[\tau_i, \tau_{i+1}]$ . Thus, this approximation is computationally simple, but it might fail to fit some properties of the FPT pdf of a CIR process. The following subsection suggests a different correction of the approximated pdf taking into account specifically these properties.

##### 4.2. On the positivity of $\hat{g}_n$

For a fixed order  $n$  of approximation, it could be of interest to construct the pdf directly, overcoming the drawbacks occurred in the numerical derivation of the cdf. In that case, although the stopping criteria in (3.13) take into account Proposition 3.3, there is no guarantee that  $\hat{g}_n(t)$  is non-negative on  $(0, \infty)$  depending on the values assumed by  $p_n(t)$  for  $t \in (t_1, t_2)$ . If this happens, an ad hoc strategy can be implemented to solve this issue.

Suppose  $(t_{1,neg}, t_{2,neg}) \subseteq (t_1, t_2)$  such that  $\hat{g}_n(t) < 0, t \in (t_{1,neg}, t_{2,neg})$ . In what follows we develop a simple numerical procedure to replace  $\hat{g}_n(t)$  with a suitable positive function  $p(t)$ , for  $t$  in a generic interval  $(t'_1, t'_2) \supseteq (t_{1,neg}, t_{2,neg})$ . It is reasonable to assume that  $\hat{g}_n(t)$  can be negative in an interval located to the right or/and to the left of the approximated mode of the FPT rv, since the FPT pdf of a diffusion process is unimodal [46]. In both cases, a classical technique would consist in selecting a fourth-degree polynomial  $p(t)$  interpolating smoothly  $(t'_1, \hat{g}_n(t'_1))$  and  $(t'_2, \hat{g}_n(t'_2))$ , fulfilling the additional constraints imposed by the conservation of probability mass

$$\int_{t'_1}^{t'_2} p(t) dt = \int_{t'_1}^{t'_2} \hat{g}_n(t) dt, \tag{4.2}$$

as well as positivity and monotonicity. Since such a polynomial is unique, the last two remaining conditions would possibly be satisfied by a computationally cumbersome choice of  $t'_1$  and  $t'_2$ . Therefore, in the following we propose a different approach both to determine numerically  $(t'_1, t'_2)$  and to correct  $\hat{g}_n(t)$ , taking into account the conditions required on the FPT pdf  $g(t)$  in Theorem 3.1.

Suppose  $m^* = \max_{t \in (0, \infty)} \hat{g}_n(t)$  be the approximated mode of the FPT rv  $T$ . In agreement with the previous observations, the following two possible scenarios might occur:

- a)  $\hat{g}_n(t) < 0$  for  $t \in (t_{1,neg}, t_{2,neg})$  with  $t_{2,neg} < m^*$
- b)  $\hat{g}_n(t) < 0$  for  $t \in (t_{1,neg}, t_{2,neg})$  with  $t_{1,neg} > m^*$ .

Two procedures were therefore developed taking into account the behavior of the FPT pdf either in a right-handed neighborhood of the origin - Case a) - and on the tail - Case b) - with the aim of minimizing the number of parameters involved while streamlining the writing and the procedure. No conservation of the probability mass (4.2) is required in this strategy. Empirically, this is justified by the circumstance that in all the observed cases the negative areas are so small that the mass involved gives almost no contribution.

Case a). Since  $g(0) = 0$ , we set  $t'_1 := 0$ . To reduce the number of parameters, we assume  $p(t) = at^\delta$  with  $a > 0$  and  $\delta > \frac{\alpha}{2} + 1$  according to Theorem 3.1. Therefore, the correction of  $\hat{g}_n(t)$  is defined as

$$\hat{g}_n^{corr}(t) = \begin{cases} at^\delta & 0 \leq t \leq t'_2, \\ \hat{g}_n(t) & t > t'_2. \end{cases} \tag{4.3}$$

In order to achieve a certain level of smoothness,  $a$  and  $\delta$  in (4.3) are chosen such that

$$p(t'_1) = \hat{g}_n(t'_1) \quad \text{and} \quad \left. \frac{d}{dt} p(t) \right|_{t=t'_1} = \left. \frac{d}{dt} \hat{g}_n(t) \right|_{t=t'_1}.$$

Finally, we set

$$t'_2 := \min \left\{ t \in (0, m^*) \left| \int_0^t \hat{g}_n(t) dt > 0 \right. \right\}$$

to avoid an excessive increment of the probability mass when  $\hat{g}_n(t)$  is replaced by  $\hat{g}_n^{corr}(t)$ . Fig. 1-a) shows an example of negative  $\hat{g}_n(t)$  in an area close to the origin together with its correction  $\hat{g}_n^{corr}(t)$ , as obtained by the procedure described above.

Case b). We assume  $p(t) = ae^{bt}$ , with  $a > 0$  and  $b < 0$ , according to the well known exponential asymptotic behavior of the FPT pdf of diffusion processes with steady-state distribution [37]. Therefore, the correction of  $\hat{g}_n(t)$  is defined as

$$\hat{g}_n^{corr}(t) = \begin{cases} \hat{g}_n(t) & t < t'_1 \\ ae^{bt} & t'_1 \leq t \leq t'_2 \\ \hat{g}_n(t) & t > t'_2. \end{cases} \tag{4.4}$$

In this case,  $a$  and  $b$  are chosen such that

$$p(t'_1) = \hat{g}_n(t'_1) \quad \text{and} \quad p(t'_2) = \hat{g}_n(t'_2).$$

In order to fit a decreasing exponential function in the interval  $(t'_1, t'_2)$ , the endpoints  $t'_1$  and  $t'_2$  are chosen such that

$$t'_2 = \min \left\{ t > t_2 \left| \frac{d}{dt} \hat{g}_n(t) < 0 \right. \right\} \quad \text{and} \quad t'_1 = \min \left\{ t < t_1 \left| \hat{g}_n(t) > \hat{g}_n(t'_2) \right. \right\}.$$

Fig. 1-b) shows an example of negative  $\hat{g}_n(t)$  for values of  $t$  larger than the mode together with its correction  $\hat{g}_n^{corr}(t)$ , as obtained by the procedure just described.

### 5. Numerical results

In this section, the functions  $\hat{G}_n(t)$  in (3.11) and  $\hat{g}_n(t)$  in (3.7) are used to approximate respectively the FPT cdf  $G(t)$  and the FPT pdf  $g(t)$  of a CIR process (2.1) with the corrections suggested in subsections 4.1 and 4.2.

Before examining the numerical results, in the following we provide some details on how the goodness of approximation was assessed.

#### 5.1. Comparisons with alternative approximation methods

As the shape of the FPT pdf and cdf for a CIR process is unknown, the proposed approximations’ validity needs to be evaluated by comparing it with alternative estimates obtained using different techniques.

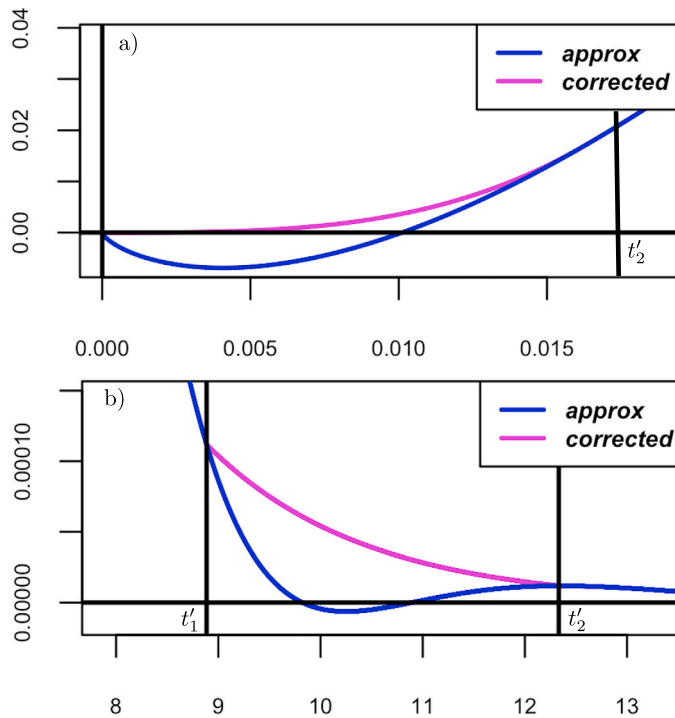


Fig. 1. In a) plots of the approximation  $\hat{g}_n$  and of its correction  $\hat{g}_n^{corr}$  (4.3) over the interval  $(0, t'_2)$  are given for  $n = 10$  and parameters  $y_0 = 0, \mu = 3, S = 10, c = -10, \sigma = 1.2, \tau = 0.2$  (see case C in Section 5.2). In b), plots of the approximation  $\hat{g}_n$  and of its correction  $\hat{g}_n^{corr}$  (4.4) over the interval  $(t'_1, t'_2)$  are given for  $n = 9$  and parameters  $y_0 = 0.2, \mu = 0.9, S = 1, c = 0, \sigma = 1.2, \tau = 2/3$  (see case A in Section 5.2).

For one-dimensional diffusion processes, the FPT pdf through a time-dependent boundary can be recovered as the solution of a Volterra integral equation of the second kind [5]. With suitable numerical methods for approximating the integral, a discrete numerical evaluation of this solution can be efficiently computed. Unfortunately, when implementing these tools, we came upon the issue that, when the coefficient of variation is large, this procedure is subject to overflow failures and could lead to completely misleading results. This circumstance is likely amplified by an unavoidable propagation of numerical errors since the new approximated values of the FPT pdf are computed using those recovered at the previous steps. However, even for FPT pdfs with a small coefficient of variation, we encountered numerous issues in its implementation. These undesirable behaviors are essentially caused by the presence of the Bessel function in the transition pdf of the process, whose derivatives cause numerical overflow issues. Similar difficulties have been encountered using the R package `fptdApprox` [45]. Ultimately, beyond these pathological cases, the numerical results are comparable to those obtained by Monte Carlo methods, which we have therefore chosen as methods for assessing the goodness of approximation.

A Monte Carlo method consists in generating sample paths of the CIR process and look for their FPTs over the given threshold. When choosing how to sample the paths of the CIR process in the Monte Carlo method, we first implemented the Milstein algorithm [39] which generates a trajectory by a suitable discretization of the stochastic differential equation (2.1). As it is well-known, this procedure is time-demanding: to get a FPT sample of size  $N$ , at least  $N$  different trajectories of the CIR process must be generated. Indeed not all the generated trajectories may reach the threshold in a reasonable time. This also implies that the FPT pdf can be underestimated if a finite time interval has been set for the simulation, as usually happens. Moreover, the fixed time step determines how accurately the dynamics can be described and the computational time increases as this time step gets smaller. Therefore, to obtain significant results, it is necessary to choose a very small step size and simulate many trajectories of the CIR process. This can be very time-consuming, especially if the coefficient of variation of  $T$  is large, as a consequence of the likelihood of a large time span length over which the trajectories must be simulated.

Similar problems arise when sampling positions of the CIR process using its transition pdf, a non-central chi-square distribution [18]. In such a case, starting from  $Y_0 = y_0$ , an instance of  $Y_{\Delta t}$  is generated from the conditional distribution of  $Y_{\Delta t} | Y_0 = y_0$ , an instance of  $Y_{2\Delta t}$  from the conditional distribution of  $Y_{2\Delta t} | Y_{\Delta t}$  and so on. The results obtained by the two Monte Carlo methods are comparable. However we have used the Milstein algorithm, since the computational time is lower in all the cases examined.

Once a FPT random sample has been collected, a sufficiently reliable estimate of the cdf shape can be obtained using the empirical cdf. Nonparametric methods can also be used to obtain an estimate of the pdf. The most widely used nonparametric pdf estimator is the histogram. Despite its popularity, the drawbacks of this tool are well known in the literature, as for example, the strong dependence on bandwidth choice. In the literature, KDEs are typically mentioned as simple alternatives to histograms. If the unknown pdf has its support confined to the positive half line and is not smooth at the origin, then the kernel method can perform not efficiently [26]. Indeed, the mode of the unknown pdf may be actually hidden by the KDE assigning positive mass to negative values (see Fig. 5).

**Table 1**  
 FPT dispersion indexes  $c_v$  and  $\hat{c}_h$  together with mean, standard deviation, skewness and kurtosis for the cases A, B and C.

	$c_v$	$\hat{c}_h$	$\mathbb{E}[T]$	$\sqrt{\text{Var}[T]}$	$\gamma_1$	$\kappa_1$
A	0.855	0.909	1.16	0.984	1.968	5.9862
B	1.231	0.916	2.991	13.56	2.39	8.118
C	0.765	0.855	3.937	9.084	1.905	5.572

To recover the smoothness characterizing the KDEs and still obtain an adequate estimated density with support  $(0, \infty)$ , an estimator based on an orthogonal series can be very competitive [17]. Indeed, an orthogonal series estimator is exactly what is obtained from the approximation  $\hat{g}_n(t)$  in (3.1) when the theoretical FPT moments are replaced by their corresponding sample moments. In fact, suppose a sample of iid FPTs  $\mathcal{T} = \{T_1, \dots, T_N\}$  is available, arising either from simulations or from experiments. A straightforward calculation shows that replacing FPT moments (2.4) with sample moments is equivalent to replacing  $\mathbb{E}[L_k^{(\alpha)}(\beta T)]$  in (2.10) with its sample mean estimator

$$\bar{l}_k = \frac{1}{N} \sum_{i=1}^N L_k^{(\alpha)}(\beta T_i).$$

Then, the FPT pdf  $g(t)$  can be approximated by

$$\bar{g}_n(t) = f_{\alpha,\beta}(t) \left( 1 + \sum_{k=1}^n \bar{l}_k b_k^{(\alpha)} L_k^{(\alpha)}(\beta t) \right) \quad \text{with} \quad b_k^{(\alpha)} = \frac{k! \Gamma(\alpha + 1)}{\Gamma(\alpha + 1 + k)} \tag{5.1}$$

which is the orthogonal series estimator of  $g(t)$ . This observation reveals an additional advantage of using the Laguerre-Gamma approximation. If the FPT moments/cumulants are not known but a random sample is available, the Laguerre-Gamma approach offers the opportunity to recover an approximation of the FPT pdf similarly to the orthogonal series estimators. In such a case, the estimates carried out by sample moments or by  $k$ -statistics [13] replace the occurrences of FPT moments or cumulants respectively. Under suitable hypotheses on the true pdf  $g(t)$ , estimations of the convergence order of  $\bar{g}_n(t)$  to  $g(t)$  are assessable through the mean integrated squared error [26]. Indeed, by recalling that the generalized Laguerre polynomial sequence is orthonormal with respect to the reference density  $f_{\alpha,\beta}(t)$ , the mean integrated squared error is [15]

$$\begin{aligned} J(n) &= \mathbb{E} \int_0^\infty \left| \frac{\bar{g}_n(t) - g(t)}{f_{\alpha,\beta}(t)} \right|^2 f_{\alpha,\beta}(t) dt \\ &= \frac{1}{N} \sum_{k=1}^n b_k^{(\alpha)} \text{Var}[L_k^{(\alpha)}(\beta T)] + \sum_{k>n} b_k^{(\alpha)} \mathbb{E}[L_k^{(\alpha)}(\beta T)]^2 \end{aligned} \tag{5.2}$$

with  $b_k^{(\alpha)}$  as in (5.1). Thus increasing the sample size leads to a reduction of the error  $J(n)$  as expected from the estimation of moments with sample moments. There are various strategies in the literature to choose the degree of the polynomial approximation  $n$  in (5.2), see for example [15]. A discussion on which strategy is the most effective for the CIR process goes beyond the scope of this paper.

In all the cases examined, the results estimated by the orthogonal series method on a collected FPT sample overlap with the Laguerre-Gamma approximations (3.7), when the stopping criteria addressed in Section 3.1 are used. This explains why the subsequent section does not include these results. Instead, to assess the goodness of these approximations, the corresponding histograms have been used despite their well-known flaws.

### 5.2. Numerical examples

To analyze the efficiency and the usefulness of the proposed method, in the following we consider three scenarios:

- case A:  $y_0 = 0.2, \mu = 0.9, S = 1, c = 0, \sigma = 1.2$  and  $\tau = 2/3$ ,
- case B:  $y_0 = 0.01, \mu = 0.005, S = 0.02, c = 0, \sigma = 0.1$  and  $\tau = 0.25$ ,
- case C:  $y_0 = 0, \mu = 3, S = 10, c = -10, \sigma = 1.2$  and  $\tau = 0.2$ .

Each case results in FPT pdfs and cdfs with different forms and statistical properties, as shown in Fig. 2, where plots of empirical FPT cdfs are given.

According to Section 5.1, the empirical cdfs have been constructed after using the Milstein method to simulate a sample of  $10^4$  FPTs for each case. In the following, these three samples are denoted by  $\mathcal{T}_A, \mathcal{T}_B$  and  $\mathcal{T}_C$  respectively.

For each case, we have computed the FPT dispersion coefficients as given in Section 3.2. The coefficient of variation is computed using the theoretical FPT mean and variance. The estimated relative entropy based dispersion coefficient  $\hat{c}_h$  is computed with the Vasicek estimator [30] using the samples  $\mathcal{T}_A, \mathcal{T}_B$  and  $\mathcal{T}_C$  respectively. The results are given in Table 1.

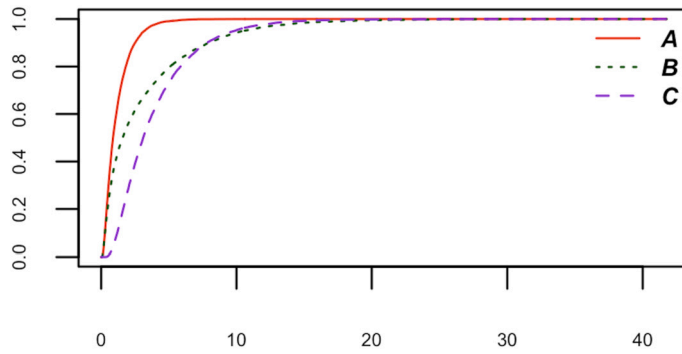


Fig. 2. Plots of the empirical (not standardized) FPT cdfs for cases A (in solid red), B (in dash green), and C (in dashed purple).

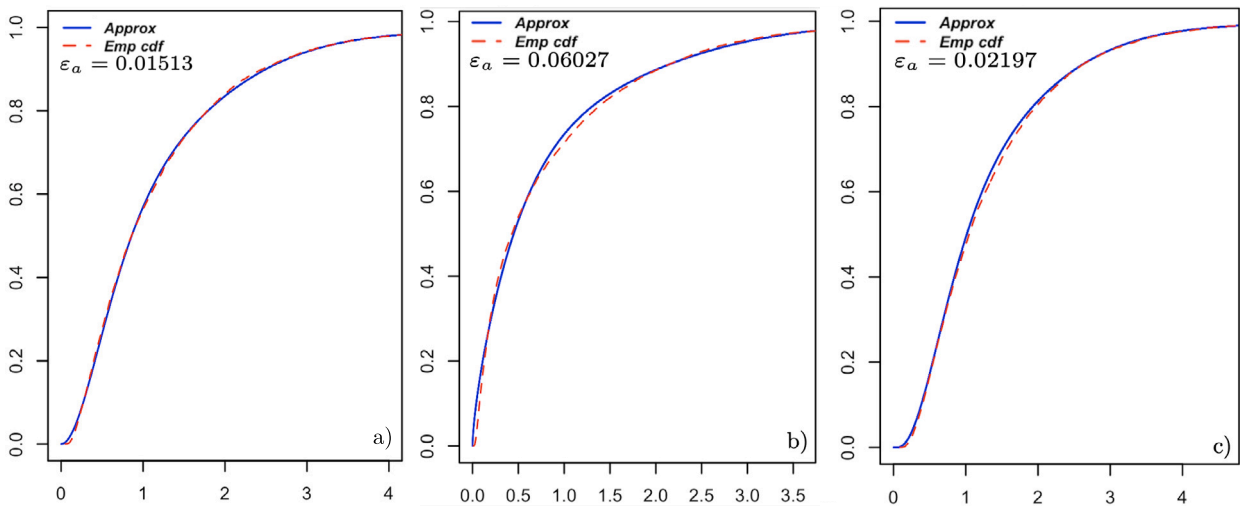


Fig. 3. Plots of the approximated  $\tilde{G}_n(t)$  (in solid blue) and of the empirical cdf (in dashed red) together with the corresponding maximum absolute error  $\epsilon_\alpha$ . The plots refer to case A in a) with  $n = 10$ ,  $\alpha = 0.367$  and  $\beta = 1.17$ ; to case B in b) with  $n = 10$ ,  $\alpha = -0.34$  and  $\beta = 0.812$ ; to case C in c) with  $n = 9$ ,  $\alpha = 0.7$  and  $\beta = 1.306$ . Note that  $\tilde{G}_n(t)$  is obtained using the stopping criteria (3.13) and corrected according to Section 4.1 while the empirical cdf is obtained from the standardized samples  $\mathcal{T}_A$ ,  $\mathcal{T}_B$  and  $\mathcal{T}_C$  respectively.

Figs. 3, 4, 5 and 6 refer to the standardized FPT rv  $\tilde{T}$  in (3.15). From now onward, whenever mentioned, the stopping criteria in (3.13) is employed with  $\epsilon = 10^{-14}$ . In Fig. 3 we have plotted the empirical cdfs  $G_e(t)$ , corresponding to the samples  $\mathcal{T}_A$ ,  $\mathcal{T}_B$  and  $\mathcal{T}_C$ , together with the approximated cdfs  $\tilde{G}_n(t)$ , as obtained using (4.1) and the corrections described in Section 4.1, for  $\Delta t = 10^{-5}$ , normalized by the corresponding standard deviations (see Table 1). Moreover each figure displays the maximum absolute error defined as  $\epsilon_\alpha = \max_{t \geq 0} |\tilde{G}_n(t) - G_e(t)|$ . Figs. 4, 5 and 6, correspond to cases A, B and C respectively. To emphasize the differences in density estimations, as discussed in Section 5.1, we have plotted in these figures a classical KDE,<sup>3</sup> a histogram<sup>4</sup> and the standardized approximated pdf  $\tilde{g}_n(t)$ , corrected according to Section 4.2. Since the three considered instances correspond to FPT pdfs with different shapes, these comparisons should provide a comprehensive picture of the strengths and weaknesses of the proposed method, which are discussed in the following.

5.2.1. Case A

Among the three instances taken into consideration, case A has the lightest tails (see Fig. 2). This produces an accurate approximation, even with a small value of  $n$ . Indeed there is a low maximum absolute error between the empirical and approximated cdf (see Fig. 3-a)), which is assumed at  $t = 0.134$ . For the pdf, the suggested approach combined with the stopping criteria (3.13) yield an approximation  $\tilde{g}_n(t)$  with  $n = 10$ ,  $\alpha = 0.367$  and  $\beta = 1.17$ . In this case,  $\tilde{g}_n(t)$  is negative on a small interval after the mode, as shown in Fig. 1. Therefore, a suitable correction of  $\tilde{g}_n(t)$  has been implemented, as outlined in Section 4.2. This corrected approximation is plotted in Fig. 4. The same figure shows also the estimated pdf obtained with a classical KDE and with a histogram, both computed on the standardized sample  $\mathcal{T}_A$ .

<sup>3</sup> The KDE has been generated by the R function `density()` [42].

<sup>4</sup> The histogram has been generated by the R function `hist()` [42].

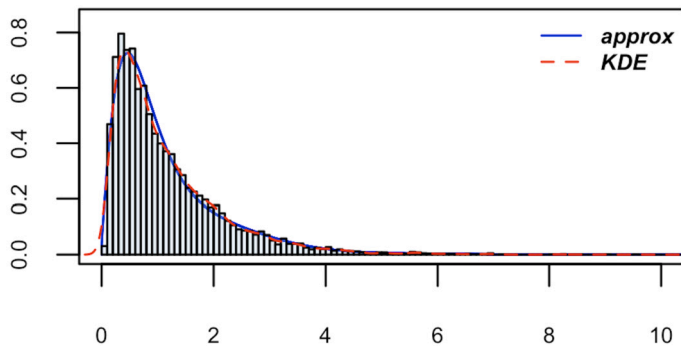


Fig. 4. Plot of  $\tilde{g}_n(t)$  (in solid blue) in case A with  $n = 10$ ,  $\alpha = 0.367$  and  $\beta = 1.17$ , obtained with the stopping criteria (3.13) and corrected to ensure positivity as in (4.4), together with a KDE (in dashed red) and a histogram both computed with the standardized sample  $\mathcal{T}_A$ .

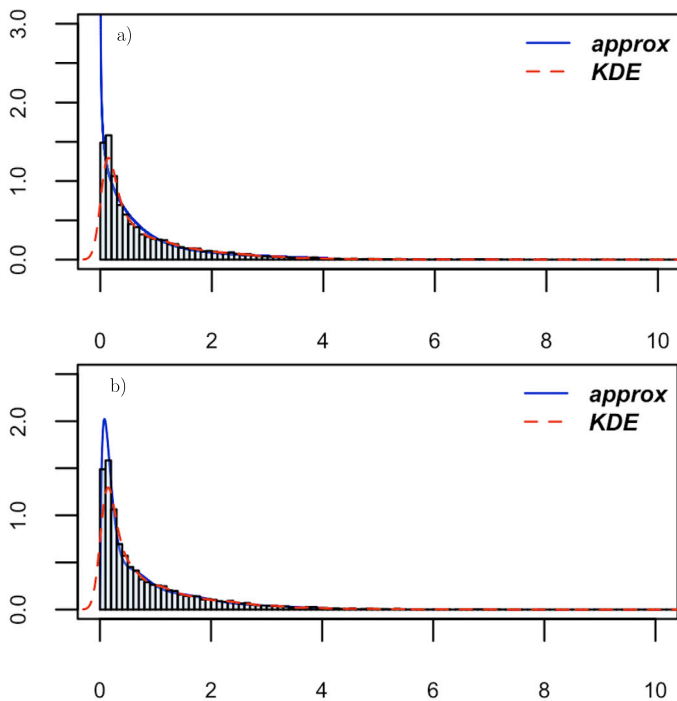


Fig. 5. In a), plot of  $\tilde{g}_n(t)$  (in solid blue) in case B with  $n = 10$ ,  $\alpha = -0.34$  and  $\beta = 0.812$ , obtained with the stopping criteria (3.13) together with a KDE (in dashed red) and a histogram both computed with the standardized sample  $\mathcal{T}_B$ . In b), a plot of  $\tilde{g}_n(t)$  (in solid blue) in case B with  $n = 55$ ,  $\alpha = -0.34$  and  $\beta = 0.812$ , obtained without any stopping criterion, increasing the numerical precision.

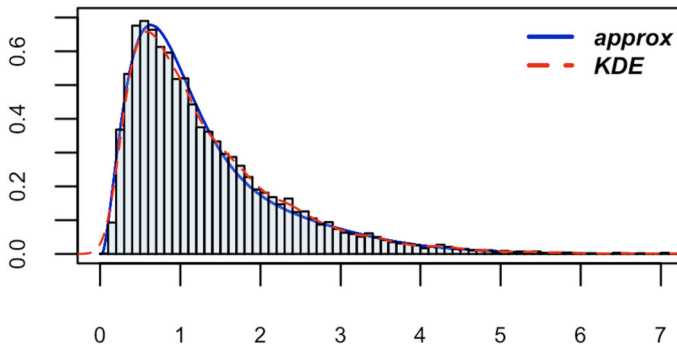


Fig. 6. Plot of  $\tilde{g}_n(t)$  (in solid blue) for case C with  $n = 9$ ,  $\alpha = 0.7$  and  $\beta = 1.306$ , obtained with the stopping criteria (3.13) and corrected to ensure positivity as in (4.3), together with a KDE (in dashed red) and a histogram both computed with the sample  $\mathcal{T}_C$ , after the latter has been standardized.

5.2.2. Case B

In this case, different considerations are required. From Table 1, the FPT rv has a coefficient of variation larger than 1. This makes the approximation more challenging because the distribution has a significantly heavy tail (see Fig. 2). The stopping criteria (3.13) yield an approximated FPT pdf  $\tilde{g}_n(t)$  with  $n = 10$ ,  $\alpha = -0.34$  and  $\beta = 0.812$ , not adequately reproducing  $\tilde{g}(t)$  when compared with a KDE and a histogram, as shown in Fig. 5 a). This is a result of the first conservative stopping criterion in (3.13), which was initially proposed to avoid numerical instability caused by an excessively high order of approximation. To underline that the behavior of  $\tilde{g}_n$  is significantly influenced by the numerical precision, Fig. 5-b) shows the remarkably good approximation obtained when computing  $\tilde{g}_n(t)$  with a very high numerical precision, allowing to push the recursion up to  $n = 55$ . The numerical precision has been raised using the R-package `bignum` [25]. Still, it is worth to mention that the approximation  $\tilde{g}_n(t)$  in Fig. 5-a) yields a satisfactory result for the tail of  $\tilde{g}(t)$ . Note that the maximum absolute error at  $t = 0.042$  between empirical and approximated cdf is higher compared to case A (see Fig. 3-b)). This results from overestimating  $\tilde{g}(t)$  in the right-handed neighborhood of the origin, see Fig. 5-b).

5.2.3. Case C

In this case, from Fig. 2 and Table 1, the FPT pdf has a coefficient of variation less than 1 along with a tail whose heaviness lies between cases A and B. As in case A, this value of the coefficient of variation should intuitively ensure a good approximation. Indeed there is a low maximum absolute error between the empirical and approximated cdfs (see Fig. 3-c)), which is assumed at  $t = 1.365$ . For the pdf, the suggested approach yields an approximated FPT pdf  $\tilde{g}_n(t)$  with  $n = 9$ ,  $\alpha = 0.7$  and  $\beta = 1.306$ . In this case,  $\tilde{g}_n(t)$  is negative on a small interval before the mode and close to the origin (see Fig. 1). Therefore, also in this case, a suitable correction of  $\tilde{g}_n(t)$  has been implemented, as outlined in Section 4.2. The resulting approximation is shown in Fig. 6.

6. Application: an acceptance-rejection type algorithm

In this section we propose a possible application of the polynomial FPT pdf approximation consisting in sampling FPTs using a method similar to the acceptance-rejection one.

The acceptance-rejection method is a classical technique for sampling from a distribution that is unknown or difficult to simulate through an inverse transformation [44]. Under such circumstances, samples are collected from an auxiliary density if a suitable probability of acceptance is known [6,8].

More in details, let  $Z$  be a rv with pdf  $\pi$ . Suppose there exists a constant  $M > 0$  and a pdf  $q(x)$  such that

$$q(x) > 0 \text{ whenever } \pi(x) > 0 \text{ and } \frac{\pi(x)}{q(x)} \leq M, \quad \forall x \in \text{supp}(Z). \tag{6.1}$$

The acceptance-rejection method exploits the condition in (6.1) to sample from the support of  $Z$ . Therefore, if the FPT pdf  $g(t)$  were known and an upper bound for the right hand side of

$$\frac{g(t)}{f_{\alpha,\beta}(t)} = \sum_{k \geq 0} a_k^\alpha Q_k^{(\alpha)}(\beta t), \tag{6.2}$$

were available, this method could have been used to sample from the FPT rv. However,  $g(t)$  is generally unknown. Therefore,

$$\frac{\hat{g}_n(t)}{f_{\alpha,\beta}(t)} = p_n(t) \tag{6.3}$$

can be used in place of (6.2), assuming  $\hat{g}_n(t)$  as given in (3.7) and non-negative for all  $t > 0$ . Unfortunately, since  $p_n(t)$  is a polynomial for any  $n > 0$ , the right hand side of (6.3) is clearly unbounded on  $(0, \infty)$ .

We provide a suitable modification of the standard acceptance-rejection method with the aim of sampling from the FPT rv using (6.3). In the following, suppose  $T_n$  be the rv with non-negative pdf  $\hat{g}_n(t)$  over  $(0, \infty)$ . The main steps of the method here proposed can be summarized as follows:

- i) find a constant  $C$  such that  $\mathbb{P}(T > C) \leq \varepsilon$ , for a fixed, small  $\varepsilon > 0$ ;
- ii) for  $t \leq C$  apply the classical acceptance-rejection method using the ratio

$$\frac{\hat{g}_{T_n|C}(t)}{\tilde{f}_{\alpha,\beta|C}(t)} \leq M \text{ where } M = \frac{\mathbb{P}(T_n \leq C)}{\mathbb{P}(X_{\alpha,\beta} \leq C)} \max_{t \in (0,C]} p_n(t) \tag{6.4}$$

and

$$\tilde{f}_{\alpha,\beta|C}(t) = \frac{f_{\alpha,\beta}(t)}{\mathbb{P}(X_{\alpha,\beta} \leq C)} \mathbb{1}_{(0,C]}(t), \quad \hat{g}_{T_n|C}(t) = \frac{\hat{g}_n(t)}{\mathbb{P}(T_n \leq C)} \mathbb{1}_{(0,C]}(t); \tag{6.5}$$

- iii) for  $t > C$  sample from a truncated exponential rv  $\tilde{T}$  with pdf

$$g_{\tilde{T}}(t) = \frac{1}{\mathbb{E}[T]} \exp\left(-\frac{t-C}{\mathbb{E}[T]}\right) \mathbb{1}_{(C,+\infty)}(t). \tag{6.6}$$

The last step takes into account the FPT pdf’s exponential asymptotic behavior for one-dimensional diffusion processes with steady-state distribution [37]. Algorithm 1 outlines the proposed method for constructing an “approximated” sample  $S$  of size  $N$  from  $T$ .

---

**Algorithm 1:** Modified acceptance-rejection method.

---

Set the parameters  $\alpha, \beta > 0, n > 0$  and  $\epsilon > 0$ .

**Initialize**

Find a constant  $C$  such that  $\mathbb{P}(T > C) \leq \epsilon$ .  
 Set  $M$  as in (6.4).  
 $j \leftarrow 1$ .  
 $S \leftarrow \{ \}$ .

**While**  $j < N$

**With probability**  $\epsilon$

Generate  $\bar{T}$  from the truncated exponential pdf in (6.6).  
 $S \leftarrow S \cup \bar{T}$ .  
 $j \leftarrow j + 1$ .

**With probability**  $1 - \epsilon$

Generate  $G$  from the truncated gamma pdf  $\tilde{f}_{\alpha,\beta|C}$  in (6.5).  
 Generate  $U \sim U(0, 1)$ .

**While**  $U > \frac{\hat{g}_{r|C}(G)}{M \tilde{f}_{\alpha,\beta|C}(G)}$

Generate  $G$  from the truncated gamma pdf  $\tilde{f}_{\alpha,\beta|C}$  in (6.5).  
 Generate  $U \sim U(0, 1)$ .

$S \leftarrow S \cup G$ .

$j \leftarrow j + 1$ .

**Return**  $S$ .

---

In addition to the user-specified input parameters, the constant  $C$  in *i*) must be chosen for the algorithm initialization. The so-called *Vysochanskij-Petunin* inequality for one-sided tail bounds [41] is used to achieve this goal.

**Theorem 6.1** (*Vysochanskij-Petunin inequality*). *If  $r \geq 0$  and  $X$  is a rv with unimodal density, finite mean  $\mu$  and finite variance  $\sigma^2$ , then*

$$\mathbb{P}(X - \mu \geq r) \leq \begin{cases} \frac{4}{9} \frac{\sigma^2}{\sigma^2 + r^2} & \text{if } 3r^2 \geq 5\sigma^2, \\ \frac{4}{3} \frac{\sigma^2}{\sigma^2 + r^2} - \frac{1}{3} & \text{otherwise.} \end{cases} \tag{6.7}$$

Since the FPT rv of diffusion processes has a unimodal pdf [46], the *Vysochanskij-Petunin* inequality can be applied. Setting  $4\sigma^2/[9(\sigma^2 + r^2)] = \epsilon$  in the first inequality (6.7), we recover  $r = r(\epsilon)$  as a function of  $\epsilon$  and get the condition  $\epsilon \leq 1/6$  from  $3r(\epsilon)^2 \geq 5\sigma^2$ . Then set  $C = \mu + r(\epsilon)$  where

$$r(\epsilon) = \begin{cases} \sqrt{\frac{4\sigma^2}{9\epsilon} - \sigma^2} & \text{if } \epsilon \leq 1/6, \\ \sqrt{\frac{4\sigma^2}{1+3\epsilon} - \sigma^2} & \text{if } 1/6 < \epsilon \leq 1. \end{cases} \tag{6.8}$$

**Remark 6.2.** From (6.8),  $C$  may become arbitrarily large by decreasing  $\epsilon$ . However, as  $C$  increases, so does  $M$ , increasing the chance of rejection and subsequently the required number of iterations.

The quality of the outcome of Algorithm 1 relies on the approximation  $\hat{g}_n(t)$  and the selection of an exponential distribution for  $t > C$ .

A theoretical justification for Algorithm 1 is provided in the following. We first calculate the cdf of the rv  $Y$ , whose observations are generated by Algorithm 1, and then prove that such a distribution turns out to be a good approximation of the FPT cdf.

**Lemma 6.3.** *If  $Y$  denotes the rv sampled at the end of each cycle of Algorithm 1, then*

$$\mathbb{P}(Y \leq t) = \epsilon \left[ 1 - \exp\left(-\frac{t-C}{\mathbb{E}(T)}\right) \right] + (1 - \epsilon)\mathbb{P}(T_n \leq t \mid T_n \leq C), \quad t > 0$$

where  $C, n$  and  $\epsilon$  are given in Algorithm 1, and  $T_n$  is the rv with pdf  $\hat{g}_n(t)$ .

**Proof.** According to Algorithm 1, we have

$$\mathbb{P}(Y \leq t) = \mathbb{P}(X = 1)\mathbb{P}(\tilde{T} \leq t) + \mathbb{P}(X = 0)\mathbb{P}(G \leq t \mid G \text{ accepted}) \tag{6.9}$$

where  $X$  is a Bernoulli rv of parameter  $\varepsilon \in (0, 1)$  independent from the rv  $\tilde{T}$ , with truncated exponential pdf  $g_{\tilde{T}}(t)$  in (6.6), and the rv  $G$  with truncated gamma pdf  $\tilde{f}_{\alpha,\beta|C}(t)$  in (6.5). Thus from (6.6) and (6.9), we have

$$\mathbb{P}(Y \leq t) = \varepsilon \left[ 1 - \exp\left(-\frac{t-C}{\mathbb{E}(T)}\right) \right] + (1 - \varepsilon)\mathbb{P}(G \leq t \mid G \text{ accepted}). \tag{6.10}$$

For the latter term in (6.10) observe that

$$\mathbb{P}(G \leq t \mid G \text{ accepted}) = M \mathbb{P}(G \leq t, G \text{ accepted}) \tag{6.11}$$

since  $\mathbb{P}(G \text{ accepted}) = 1/M$  with  $M$  given in (6.4). Moreover

$$\begin{aligned} \mathbb{P}(G \leq t, G \text{ accepted}) &= \int_0^\infty \mathbb{P}(G \leq t, G \text{ accepted} \mid G = x) \tilde{f}_{\alpha,\beta|C}(x) dx \\ &= \int_0^\infty \mathbb{P}(G \leq t \mid G = x) \mathbb{P}(G \text{ accepted} \mid G = x) \tilde{f}_{\alpha,\beta|C}(x) dx \end{aligned} \tag{6.12}$$

since  $(G \leq t)$  and  $(G \text{ accepted})$  are conditionally independent events. Observe that  $\mathbb{P}(G \leq t \mid G = x) = \mathbb{1}_{x \leq t}$  and

$$\begin{aligned} \mathbb{P}(G \text{ accepted} \mid G = x) &= \mathbb{P}\left(U \leq \frac{\hat{g}_{T_n|C}(G)}{M \tilde{f}_{\alpha,\beta|C}(G)} \mid G = x\right) \\ &= \mathbb{P}\left(U \leq \frac{\hat{g}_{T_n|C}(x)}{M \tilde{f}_{\alpha,\beta|C}(x)}\right) = \frac{\hat{g}_{T_n|C}(x)}{M \tilde{f}_{\alpha,\beta|C}(x)} \end{aligned} \tag{6.13}$$

since  $U$  is a rv with uniform distribution over  $(0, 1)$ . Plugging (6.13) in (6.12) and the resulting integral in (6.11), we get

$$\mathbb{P}(G \leq t \mid G \text{ accepted}) = M \int_0^\infty \mathbb{1}_{x \leq t} \frac{\hat{g}_{T_n|C}(x)}{M \tilde{f}_{\alpha,\beta|C}(x)} \tilde{f}_{\alpha,\beta|C}(x) dx = \int_0^t \hat{g}_{T_n|C}(x) dx,$$

and the result follows from (6.5).  $\square$

The following is a technical lemma necessary for the subsequent proposition. The lemma gives an upper bound of the error in approximating the FPT pdf  $g(t)$  with the Laguerre-Gamma expansion  $\hat{g}_n(t)$  in (3.7) for  $t \leq C$ .

**Lemma 6.4.** Under the same hypotheses as Lemma 6.3, we have

$$\left| \mathbb{P}(T_n \leq t \mid T_n \leq C) - \mathbb{P}(T \leq t \mid T \leq C) \right| \leq \frac{1}{\mathbb{P}(T_n \leq C)} \left[ \sum_{k \geq n+1} (a_k^{(\alpha)})^2 \right]^{\frac{1}{2}} \left( 1 + \frac{1}{\mathbb{P}(T \leq C)} \right)$$

with  $a_k^{(\alpha)}$  given in (3.1).

**Proof.** If  $g_{T|C}(t) = \mathbb{1}_{(0,C)}(t)g(t)/\mathbb{P}(T \leq C)$  is the truncated FPT pdf for  $t \leq C$ , and  $\hat{g}_{T_n|C}(t)$  is given in (6.5), we have

$$\begin{aligned} \left| \hat{g}_{T_n|C}(t) - g_{T|C}(t) \right| &= \left| \frac{\hat{g}_n(t)}{\mathbb{P}(T_n \leq C)} - \frac{g(t)}{\mathbb{P}(T \leq C)} \right| \\ &\leq \frac{1}{\mathbb{P}(T_n \leq C)} \left| \hat{g}_n(t) - g(t) \right| + g(t) \left| \frac{1}{\mathbb{P}(T_n \leq C)} - \frac{1}{\mathbb{P}(T \leq C)} \right|. \end{aligned}$$

Using the previous inequality, the difference between the truncated FPT cdf and its approximation by means of the Laguerre-Gamma expansion may be bounded as follows:

$$\begin{aligned} \left| \mathbb{P}(T_n \leq t \mid T_n \leq C) - \mathbb{P}(T \leq t \mid T \leq C) \right| &= \left| \int_0^t [\hat{g}_{T_n|C}(s) - g_{T|C}(s)] ds \right| \\ &\leq \int_0^\infty |\hat{g}_{T_n|C}(t) - g_{T|C}(t)| dt \leq \frac{1}{\mathbb{P}(T_n \leq C)} \int_0^\infty \frac{|\hat{g}_n(s) - g(s)|}{\varphi_{\alpha,\beta}(s)} \varphi_{\alpha,\beta}(s) ds \end{aligned}$$

$$\begin{aligned}
 & + \left| \frac{1}{\mathbb{P}(T_n \leq C)} - \frac{1}{\mathbb{P}(T \leq C)} \right| \underbrace{\int_0^\infty g(s) ds}_{=1} \\
 & = \frac{1}{\mathbb{P}(T_n \leq C)} \left\| \frac{\hat{g}_n - g}{f_{\alpha,\beta}} \right\|_{\mathcal{L}^1(\nu)} + \left| \frac{1}{\mathbb{P}(T_n \leq C)} - \frac{1}{\mathbb{P}(T \leq C)} \right|
 \end{aligned} \tag{6.14}$$

where  $\mathcal{L}^1(\nu)$  is the Hilbert space of the integrable functions with respect to the measure  $\nu$  having density  $f_{\alpha,\beta}(t)$ . Observe that

$$\left\| \frac{\hat{g}_n - g}{f_{\alpha,\beta}} \right\|_{\mathcal{L}^1(\nu)} \leq \left\| \frac{\hat{g}_n - g}{f_{\alpha,\beta}} \right\|_{\mathcal{L}^2(\nu)} = \left[ \sum_{k \geq n+1} (a_k^{(\alpha)})^2 \right]^{1/2} \tag{6.15}$$

where the last equality follows from (3.2). Finally, since

$$\left| \frac{1}{\mathbb{P}(T_n \leq C)} - \frac{1}{\mathbb{P}(T \leq C)} \right| \leq \frac{1}{\mathbb{P}(T_n \leq C)} \left( 1 + \frac{\mathbb{P}(T_n \leq C)}{\mathbb{P}(T \leq C)} \right) \tag{6.16}$$

with  $\mathbb{P}(T_n \leq C) \leq 1$ , plugging (6.16) and (6.15) in (6.14) the result follows.  $\square$

The following proposition gives the theoretical justification of the proposed acceptance-rejection Algorithm 1.

**Proposition 6.5.** *Under the same hypotheses as Lemma 6.3, for every  $\delta > 0$ , there exist a finite constant  $C_\delta$  and an integer  $n_\delta \in \mathbb{N}$  such that, for  $n > n_\delta$  and  $C > C_\delta$  in Algorithm 1, we have*

$$|\mathbb{P}(Y \leq t) - \mathbb{P}(T \leq t)| < \delta, \quad t > 0.$$

**Proof.** Fix  $\delta, n > 0$  and  $\varepsilon > 0$  and let  $C > 0$  be the corresponding quantity calculated in Algorithm 1. By plugging  $\mathbb{P}(T \leq t) = \mathbb{P}(T \leq C)\mathbb{P}(T \leq t | T \leq C) + \mathbb{P}(T > C)\mathbb{P}(T \leq t | T > C)$  in  $|\mathbb{P}(Y \leq t) - \mathbb{P}(T \leq t)|$  and using Lemma 6.3, the following bound can be recovered:

$$|\mathbb{P}(Y \leq t) - \mathbb{P}(T \leq t)| \leq P_1 + P_2$$

where

$$\begin{aligned}
 P_1 & := |\varepsilon \mathbb{P}(\bar{T} \leq t | \bar{T} > C) - \mathbb{P}(T > C)\mathbb{P}(T \leq t | T > C)| \\
 P_2 & := |(1 - \varepsilon)\mathbb{P}(T_n \leq t | T_n \leq C) - \mathbb{P}(T \leq C)\mathbb{P}(T \leq t | T \leq C)|,
 \end{aligned}$$

$\bar{T}$  is the truncated exponential rv with pdf in (6.6) and  $T_n$  is the rv with pdf  $\hat{g}_n(t)$ . Now, let us bound  $P_1$ . Adding and subtracting the quantity  $\varepsilon \mathbb{P}(T \leq t | T > C)$  in  $P_1$ , we get

$$P_1 \leq P_{1,1} + P_{1,2} \quad \text{with} \quad \begin{cases} P_{1,1} & := \varepsilon |\mathbb{P}(\bar{T} \leq t | \bar{T} > C) - \mathbb{P}(T \leq t | T > C)|, \\ P_{1,2} & := \mathbb{P}(T \leq t | T > C) |\varepsilon - \mathbb{P}(T > C)|. \end{cases}$$

Since  $\mathbb{P}(T > C) \leq \varepsilon$ , thanks to Remark 6.2 and to the exponential behavior of the tails of the FPT pdf, it is always possible to find  $C_{\delta,1}$  (big enough) such that  $P_{1,1} < \frac{\delta}{4}$  and  $P_{1,2} < \frac{\delta}{4}$ . Let us apply the same strategy to  $P_2$ . Adding and subtracting the quantity  $(1 - \varepsilon)\mathbb{P}(T \leq t | T \leq C)$  in  $P_2$ , we get

$$P_2 \leq P_{2,1} + P_{2,2} \quad \text{with} \quad \begin{cases} P_{2,1} & := (1 - \varepsilon) |\mathbb{P}(T_n \leq t | T_n \leq C) - \mathbb{P}(T \leq t | T \leq C)|, \\ P_{2,2} & := \mathbb{P}(T \leq t | T \leq C) |\varepsilon - \mathbb{P}(T > C)|. \end{cases}$$

From Lemma 6.4, the quantity  $P_{2,1}$  can be made sufficiently small for a suitable  $n$ , since the remainder (3.2) goes to zero as  $n$  increases. Therefore, thanks to Remark 6.2 and the exponential behavior of the FPT pdf tail, we can always find  $C_{\delta,2}$  and  $n_\delta$  such that  $P_{2,1} < \frac{\delta}{4}$  and  $P_{2,2} < \frac{\delta}{4}$ . Setting  $C_\delta = \max(C_{\delta,1}, C_{\delta,2})$  concludes the proof.  $\square$

**Remark 6.6.** Note that although Proposition 6.5 guarantees that we can always choose a constant  $C$  and an order  $n$  such that the cdf of  $Y$  sampled in each cycle of Algorithm 1 is sufficiently close to the cdf of  $T$ , as  $C$  increases,  $M = \max_{t \in [0,C]} p_n(t)$  will also increase or at most remain constant. However, the probability of acceptance in Algorithm 1 decreases if  $M$  increases, as we have already noted in Remark 6.2. Therefore, there is a trade-off between increasing  $C$ , to achieve better accuracy, and the running time of the proposed algorithm.

The ability of Algorithm 1 to generate a satisfactory “approximated” sample from the FPT rv is shown in Figs. 7, 8, and 9. We consider cases A, B and C with the standardization procedure outlined in Section 3.2. Hence in (6.5), instead of  $\hat{g}_n(t)$ , we use the approximation  $\tilde{g}_n(t)$  of the pdf  $\tilde{g}(t)$  corresponding to  $\bar{T} = T/\sigma_T$ , where  $\sigma_T$  is the standard deviation of  $T$ . Case B requires a special

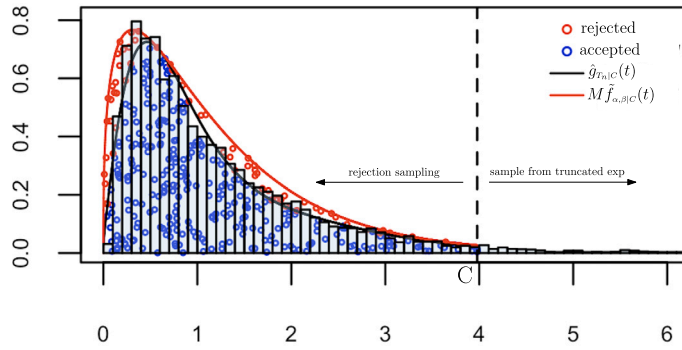


Fig. 7. Plot of rejected draws (in red) and accepted draws (in blue) for the acceptance-rejection part of Algorithm 1 with  $N = 3000$ ,  $\epsilon = 0.05$  and  $n = 10$  applied to case A, together with the corresponding  $M\tilde{f}_{\alpha,\beta|C}(t)$  and  $\tilde{g}_{n|C}(t) = \frac{\tilde{g}_n(t)}{\mathbb{P}(T_n \leq C)} \mathbb{1}_{(0,C]}(t)$  as in (6.5), with  $C = 3.97$ ,  $\alpha = 0.367$  and  $\beta = 1.17$ . Only a portion (10%) of rejected and accepted draws has been plotted, for easier viewing. On the same plot an histogram computed on the sample of size  $N = 3000$  arising from Algorithm 1 with the aforementioned parameters is shown.

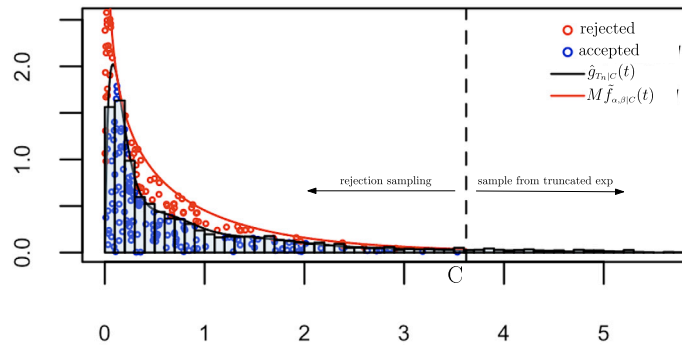


Fig. 8. Plot of rejected draws (in red) and accepted draws (in blue) for the acceptance-rejection part of Algorithm 1 with  $N = 3000$ ,  $\epsilon = 0.05$  and  $n = 55$  applied to case B, together with the corresponding  $M\tilde{f}_{\alpha,\beta|C}(t)$  and  $\tilde{g}_{n|C}(t) = \frac{\tilde{g}_n(t)}{\mathbb{P}(T_n \leq C)} \mathbb{1}_{(0,C]}(t)$  as in (6.5), with  $C = 3.621$ ,  $\alpha = -0.34$  and  $\beta = 0.812$ . Only a portion (10%) of rejected and accepted draws has been plotted, for easier viewing. On the same plot an histogram computed on the sample of size  $N = 3000$  arising from Algorithm 1 with the aforementioned parameters is shown.

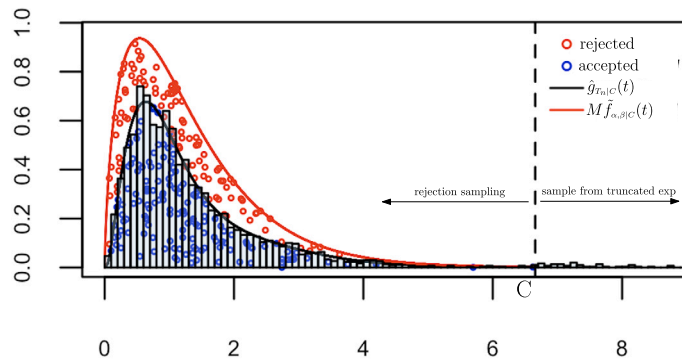


Fig. 9. Plot of rejected draws (in red) and accepted draws (in blue) for the acceptance-rejection part of Algorithm 1 with  $N = 3000$ ,  $\epsilon = 0.015$  and  $n = 9$  applied to case C, together with the corresponding  $M\tilde{f}_{\alpha,\beta|C}(t)$  and  $\tilde{g}_{n|C}(t) = \frac{\tilde{g}_n(t)}{\mathbb{P}(T_n \leq C)} \mathbb{1}_{(0,C]}(t)$  as in (6.5), with  $C = 6.65$ ,  $\alpha = 0.7$  and  $\beta = 1.306$ . Only a portion (10%) of rejected and accepted draws has been plotted, for easier viewing. On the same plot an histogram computed on the sample of size  $N = 3000$  arising from Algorithm 1 with the aforementioned parameters is shown.

attention for two main reasons. Firstly, the choice  $\alpha = -0.34$  and  $\beta = 0.812$  leads to a peaked reference distribution, which increases the probability of rejection. Secondly, to obtain a satisfactory  $\tilde{g}_n(t)$ , as already mentioned, a high numerical precision is required. However, the numerical investigations have shown that the main source of computational time rise is the accuracy of the computed  $\tilde{g}_n(t)$  rather than the acceptance-rejection method itself. Despite this, the overall computational time remains competitive in all the considered cases.

### 7. Conclusions

In this paper we develop a method for approximating the FPT pdf and cdf of a CIR process that relies on a series expansion involving the generalized Laguerre polynomials and the gamma pdf. The significant improvements of the method proposed recently in [11] follow two main directions. The first one is theoretical: we detail a study of the approximation error along with considerations on the choice of the reference pdf parameters as well as some sufficient conditions for the non-negativity of the approximant in a right hand side of the origin and in the tail. The second direction is more of a numerical nature. We propose computational strategies to overcome some numerical issues like the possible negativity of the resulting function, a feature that is undesirable in the approximation of a pdf. Methods of standardization according to considerations on dispersion measures, stopping criteria and iterative mechanisms for the choice of the best order of truncation of the series are investigated and applied to three main classes of target pdfs here considered. Moreover, we investigate the performance of the approximation of the FPT cdf that constitutes a novelty in this framework. The developed iterative procedure is of general nature and can be applied both starting from the model or from FPT sample data. The method, in fact, works well both using theoretical cumulants or their values estimated from the data. The discussion can be easily extended to the case of other diffusion processes as long as it is possible to calculate their FPT moments and the reference pdfs are moment determined [50].

As a side result an acceptance-rejection-like method, that makes use of the approximation, is proposed. It allows the generation of FPT data, although its distribution is unknown. Its usefulness is increased by the lack of existing exact methods for simulating the sample paths of the CIR process and consequently its FPTs. Furthermore, note that the suggested acceptance-rejection like procedure could be directly extended to a wider class of pdfs (not necessarily of FPT rv) which, although unknown, may admit a series expansion of the type discussed in this paper.

### Data availability

No data was used for the research described in the article.

### Acknowledgements

The authors would like to thank Andrea Marafante for his decisive contribution in developing and implementing the code necessary for the contents of this paper. A special thank goes to Antonio Di Crescenzo for some inspiring discussions.

*Funding:* The authors E.D. and G.D. acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2.2.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union – NextGenerationEU– Project Title “Non–Markovian Dynamics and Non-local Equations” – 202277N5H9 - CUP: D53D23005670006 - Grant Assignment Decree No. 973 adopted on June 30, 2023, by the Italian Ministry of University and Research (MUR). The authors participate in the INdAM - GNAMPA Project, CUP E53C23001670001.

### Appendix A

The logarithmic (partition) polynomials  $\{P_k\}$  are [7]

$$P_k(x_1, \dots, x_k) = \sum_{j=1}^k (-1)^{j-1} (j-1)! B_{k,j}(x_1, \dots, x_{k-j+1})$$

where  $\{B_{k,j}\}$  are the partial exponential Bell polynomials. For a fixed positive integer  $k$  and  $j = 1, \dots, k$ , the  $j$ -th partial exponential Bell polynomial in the variables  $x_1, x_2, \dots, x_{k-j+1}$  is an homogeneous polynomial of degree  $j$  given by

$$B_{k,j}(x_1, \dots, x_{k-j+1}) = \sum \frac{k!}{\lambda_1! \lambda_2! \dots \lambda_{k-j+1}!} \prod_{i=1}^{k-j+1} \left(\frac{x_i}{i!}\right)^{\lambda_i} \tag{A.1}$$

where the sum is taken over all sequences  $\lambda_1, \lambda_2, \dots, \lambda_{k-j+1}$  of non-negative integers such that

$$\lambda_1 + 2\lambda_2 + \dots + (k-j+1)\lambda_{k-j+1} = k, \quad \lambda_1 + \lambda_2 + \dots + \lambda_{k-j+1} = j.$$

The complete Bell (exponential) polynomials  $\{Y_k\}$  are [7]

$$Y_k(x_1, \dots, x_k) = \sum_{j=1}^k B_{k,j}(x_1, \dots, x_{k-j+1}), \quad k \geq 1$$

where  $\{B_{k,j}\}$  are the partial exponential Bell polynomials (A.1) and  $Y_0 = 1$ .

### References

[1] A. Aminataei, S. Ahmadi-Asl, Z. KalatehBojdid, Rational Laguerre functions and their applications, *Entropy* 14 (2015) 124–142.

- [2] G. Ascione, M. Bufalo, G. Orlando, On the ergodicity of a three-factor CIR model, arXiv preprint arXiv:2307.11443, 2023.
- [3] H.J. Belt, A.C. den Brinker, Optimal parametrization of truncated generalized Laguerre series, in: 1997 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 5, IEEE, 1997, pp. 3805–3808.
- [4] L. Bertini, L. Passalacqua, Modelling interest rates by correlated multi-factor CIR-like processes, arXiv preprint arXiv:0807.3898, 2008.
- [5] A. Buonocore, A.G. Nobile, L.M. Ricciardi, A new integral equation for the evaluation of first-passage-time probability densities, *Adv. Appl. Probab.* 19 (4) (1987) 784–800.
- [6] G. Casella, C.P. Robert, M.T. Wells, Generalized accept-reject sampling schemes, *Lect. Notes Monogr. Ser.* (2004) 342–347.
- [7] C.A. Charalambides, Enumerative combinatorics, in: CRC Press Series on Discrete Mathematics and Its Applications, Chapman & Hall/CRC, Boca Raton, FL, 2002.
- [8] S. Chib, E. Greenberg, Understanding the Metropolis-Hastings algorithm, *Am. Stat.* 49 (4) (1995) 327–335.
- [9] J.C. Cox, J.E. Ingersoll, S.A. Ross, A theory of the term structure of interest rates, *Econometrica* 53 (2) (1985) 385–407.
- [10] M. Deaconu, S. Herrmann, Hitting Time for Bessel Processes—Walk on Moving Spheres Algorithm (WoMS), 2013.
- [11] E. Di Nardo, G. D'Onofrio, A cumulant approach for the first-passage-time problem of the Feller square-root process, *Appl. Math. Comput.* 391 (2021) 125707.
- [12] E. Di Nardo, G. D'Onofrio, T. Martini, Approximating the first passage time density from data using generalized Laguerre polynomials, *Commun. Nonlinear Sci. Numer. Simul.* 118 (2023) 106991.
- [13] E. Di Nardo, G. Guarino, kStatistics: unbiased estimates of joint cumulant products from the multivariate Faà Di Bruno's formula, *R J.* 14 (2022) 208–228.
- [14] E. Di Nardo, D. Senato, An umbral setting for cumulants and factorial moments, *Eur. J. Comb.* 27 (3) (2006) 394–413.
- [15] P.J. Diggle, P. Hall, The selection of terms in an orthogonal series density estimator, *J. Am. Stat. Assoc.* 81 (393) (1986) 230–233.
- [16] G. D'Onofrio, P. Lansky, E. Prozzi, On two diffusion neuronal models with multiplicative noise: the mean first-passage time properties, *Chaos, Interdiscip. J. Nonlinear Sci.* 28 (4) (2018).
- [17] S. Efromovich, Orthogonal series density estimation, *Wiley Interdiscip. Rev.: Comput. Stat.* 2 (4) (2010) 467–476.
- [18] W. Feller, Two singular diffusion problems, *Ann. Math.* 54 (1) (1951) 173–182.
- [19] W. Feller, The parabolic differential equations and the associated semi-groups of transformations, *Ann. Math.* (1952) 468–519.
- [20] J. Forman, M. Sørensen, The Pearson diffusions: a class of statistically tractable diffusion processes, *Scand. J. Stat.* 35 (3) (2008) 438–465.
- [21] S. Gerhold, F. Hubalek, R.B. Paris, The running maximum of the Cox-Ingersoll-Ross process with some properties of the Kummer function, *J. Inequal. Spec. Funct.* 13 (2) (2022) 1–18.
- [22] V. Giorno, P. Lánský, A. Nobile, L. Ricciardi, Diffusion approximation and first-passage-time problem for a model neuron: III. A birth-and-death process approach, *Biol. Cybern.* 58 (1988) 387–404.
- [23] V. Giorno, A.G. Nobile, On the first-passage time problem for a Feller-type diffusion process, *Mathematics* 9 (19) (2021) 2470.
- [24] A. Göing-Jaeschke, M. Yor, A survey and some generalizations of Bessel processes, *Bernoulli* 9 (2) (2003) 313–349.
- [25] D. Hall, bignum: Arbitrary-Precision Integer and Floating-Point Mathematics, R package version 0.3.2, 2023.
- [26] P. Hall, Estimating a density on the positive half line by the method of orthogonal series, *Ann. Inst. Stat. Math.* 32 (1980) 351–362.
- [27] S. Herrmann, C. Zucca, Exact simulation of the first-passage time of diffusions, *J. Sci. Comput.* 79 (2019) 1477–1504.
- [28] S. Karlin, H.M. Taylor, A Second Course in Stochastic Processes, Academic Press, Inc., New York-London, 1981.
- [29] L. Kostal, P. Lansky, O. Pokora, Variability measures of positive random variables, *PLoS ONE* 6 (7) (2011) e21998.
- [30] L. Kostal, O. Pokora, Nonparametric estimation of information-based measures of statistical dispersion, *J. Math. Comput. Sci.* 14 (7) (2012) 1221–1233.
- [31] W. Lin, J.E. Zhang, The valid regions of Gram–Charlier densities with high-order cumulants, *J. Comput. Appl. Math.* 407 (2022) 113945.
- [32] V. Linetsky, Computing hitting time densities for CIR and OU diffusions: applications to mean-reverting models, *J. Comput. Finance* 7 (2004) 1–22.
- [33] T.-H. Lung, Approximations for Skewed Probability Densities Based on Laguerre Series and Biological Applications, North Carolina State University, 1998.
- [34] E. Martin, U. Behn, G. Germano, First-passage and first-exit times of a Bessel-like stochastic process, *Phys. Rev. E* 83 (5) (2011) 051115.
- [35] J. Masoliver, J. Perelló, First-passage and escape problems in the Feller process, *Phys. Rev. E* 86 (4) (2012) 041116.
- [36] A. Mijatović, M.R. Pistorius, J. Stolte, Randomisation and recursion methods for mixed-exponential Lévy models, with financial applications, *J. Appl. Probab.* 52 (4) (2015) 1076–1096.
- [37] A.G. Nobile, L.M. Ricciardi, L. Sacerdote, Exponential trends of first-passage-time densities for a class of diffusion processes with steady-state distribution, *J. Appl. Probab.* 22 (3) (1985) 611–618.
- [38] E. Pauwels, Smooth first-passage densities for one-dimensional diffusions, *J. Appl. Probab.* 24 (2) (1987) 370–377.
- [39] E. Platen, P.E. Kloeden, Numerical Solution of Stochastic Differential Equations, Springer-Verlag, 1992.
- [40] S.B. Provost, H.-T. Ha, Distribution approximation and modelling via orthogonal polynomial sequences, *Statistics* 50 (2) (2016) 454–470.
- [41] F. Pukelsheim, The three sigma rule, *Am. Stat.* 48 (2) (1994) 88–91.
- [42] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2022.
- [43] S. Redner, A Guide to First-Passage Processes, Cambridge University Press, 2001.
- [44] C.P. Robert, G. Casella, Monte Carlo Statistical Methods, vol. 2, Springer, 1999.
- [45] P. Román-Román, J. Serrano-Pérez, F. Torres-Ruiz, fptdapprox: approximation of first-passage-time densities for diffusion processes, R package version, 2015, p. 2.5.
- [46] U. Rosler, Unimodality of passage times for one-dimensional strong Markov processes, *Ann. Probab.* 8 (4) (1980) 853.
- [47] J.A. Shohat, On the development of functions in series of orthogonal polynomials, *Bull. Am. Math. Soc.* 41 (2) (1935) 49–82.
- [48] C.E. Smith, A Laguerre series approximation for the probability density of the first passage time of the Ornstein-Uhlenbeck process, in: Noise in Physical Systems and 1/f Fluctuations, Ohmsha, Tokyo, 1991, pp. 389–392.
- [49] S. Song, G. Xu, Y. Wang, On first hitting times for skew CIR processes, *Methodol. Comput. Appl. Probab.* 18 (2016) 169–180.
- [50] J.M. Stoyanov, G.D. Lin, P. Kopanov, New checkable conditions for moment determinacy of probability distributions, *Theory Probab. Appl.* 65 (3) (2020) 497–509.
- [51] G.A. Wilson, A. Wragg, Numerical methods for approximating continuous probability density functions, over  $[0, \infty)$ , using moments, *J. Inst. Math. Appl.* 12 (1973) 165–173.