Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Data-Guided Mark-Law Selection for Positive Mean-Reverting Jump Diffusions: Threshold Inference and Rare-Event Prediction

Halim Zeghdoudi

halim.zeghdoudi@univ-annaba.dz

Badji Mokhtar-Annaba University

Research Article

Keywords:

Posted Date: November 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11349016/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Data-Guided Mark-Law Selection for Positive Mean-Reverting Jump Diffusions: Threshold Inference and Rare-Event Prediction

Halim Zeghdoudi

Laboratory of Probability and Statistics (LaPS)
Department of Mathematics
Badji Mokhtar University, Annaba, Algeria

halim.zeghdoudi@univ-annaba.dz

Corresponding author: Halim Zeghdoudi

Email: halim.zeghdoudi@univ-annaba.dz

Abstract

This paper studies inference for a positive mean-reverting jump diffusion when jump sizes are not observed directly but inferred from increments crossing detection thresholds. Such selection changes the distribution of the retained marks and the effective observation time for jump arrivals. We combine a threshold-conditioned mark likelihood with a survival-adjusted intensity estimator and compare several positive mark laws, many expressible as finite common-rate gamma mixtures. We show that the smallest active gamma shape determines the leading order of sub-threshold jump probabilities. For Zeghdoudi marks the probability is quadratic in the threshold, and an explicit second-order expansion quantifies the downward bias of an ordinary likelihood fitted to truncated marks. These mark-specific results supplement, rather than restate, the standard affine properties of the underlying JCIR process. The paper also examines asymptotic inference under assumptions controlling jump-classification errors, parameter recovery in simulation, and threshold sensitivity in an annual Nile-flow illustration. The empirical ranking is stable over the reported threshold grid, although small information-criterion differences leave substantial mark-law uncertainty.

Keywords: jump–diffusion; CIR process; Zeghdoudi distribution; threshold inference; mark-law selection; truncated likelihood; rare-event prediction.
MSC 2020: 62M05; 62F12; 60J60; 60G51; 62P12.

1 Introduction

Many positive-valued time series are shaped by two kinds of movement: gradual fluctuations around a prevailing level and occasional upward shocks. River flows, environmental measurements, insurance intensities and financial state variables provide familiar examples. A square-root mean-reverting diffusion combined with a compound-Poisson jump component offers a convenient description of such behaviour. The resulting jump-extended Cox–Ingersoll–Ross (JCIR) process retains the positivity and affine structure of the classical model (Cox et al., 1985; Duffie and Kan, 1996; Duffie et al., 2000). These properties are well understood; the present paper does not regard their extension to another positive jump-size distribution as a new contribution.

The question examined here is statistical. In many applications the process is observed at discrete times, so the jump sizes are not available directly. Large positive increments can be used to identify candidate jump intervals, but the resulting observations are selected rather than randomly sampled from the underlying mark law. Threshold-based jump detection has an extensive literature (Mancini, 2009; Lee and Mykland, 2008; Aït-Sahalia and Jacod, 2009), as does inference for discretely observed jump diffusions (Shimizu and Yoshida, 2006; Masuda, 2007). Our concern is what happens *after* thresholding: how the selection rule alters estimation of the jump-size law, how this alteration affects estimation of jump frequency, and how uncertainty about the mark law carries over to prediction.

Consider the positive jump diffusion

$$dX_t = \kappa(\mu - X_t)\,dt + \sigma\sqrt{X_t}\,dW_t + dJ_t, \qquad J_t = \sum_{k=1}^{N_t} Y_k, \tag{1}$$

where N_t is a Poisson process with intensity $\lambda > 0$ and the positive, i.i.d. marks Y_k are independent of the Brownian motion and the jump times. The primary object of inference is the mark law F_η, rather than a modified drift or diffusion coefficient. We compare positive parametric families, including Lindley-type distributions and gamma and Weibull benchmarks. Several of the one-parameter candidates have finite gamma-mixture representations, so that the survival functions needed for threshold inference and the jump-transform calculations can be evaluated explicitly (Ghitany et al., 2008; Chouia and Zeghdoudi, 2021; Messaadia and Zeghdoudi, 2018).

Suppose a detected interval has threshold u_i and observed positive increment Z_i. On an interval containing one correctly identified jump, Z_i approximates its latent mark Y_i. The first-order likelihood must therefore condition on selection:

$$f_\eta^{\mathrm{det}}(y \mid u_i) = \frac{f_\eta(y)}{\overline{F}_\eta(u_i)}, \qquad y > u_i. \tag{2}$$

The same survival probability has a second role: it is the chance that an arriving jump is large enough to be detected. Consequently, the observed number of candidate jumps should be related to a *detection-adjusted* exposure, not automatically to the full observation horizon. Conditional mark fitting and intensity correction are thus two parts of the same observation model. Equation (2) is exact for directly observed marks subject to truncation and is used as a first-order approximation for thresholded increments; distinguishing those two settings matters for the asymptotic arguments.

The more specific mathematical contribution concerns the behaviour of the mark law near zero. For finite mixtures of common-rate gamma distributions, the smallest shape with positive mixing weight determines the leading power of the probability of a sub-threshold mark. In particular, for the Zeghdoudi law

$$f_{\mathrm{ZD}}(y;\theta) = \frac{\theta^3}{\theta+2}\, y(1+y)e^{-\theta y}, \qquad y > 0, \quad \theta > 0, \tag{3}$$

we obtain

$$\Pr_\theta(Y \le u) = \frac{\theta^3}{2(\theta+2)}u^2 + O(u^3), \qquad u \downarrow 0. \tag{4}$$

Many competing mark laws have a first-order, rather than a second-order, small-mark probability. This distinction changes the leading rate of latent jumps missed by a shrinking threshold. It also gives a measurable inferential consequence. If Zeghdoudi marks are retained only above a common oracle threshold u, but an ordinary untruncated likelihood is nevertheless fitted, the limiting rate estimate $\theta^\dagger(u)$ satisfies

$$\theta^\dagger(u) - \theta_0 = -\frac{\theta_0^4(\theta_0+3)}{2(\theta_0^2+6\theta_0+6)}\, u^2 + O(u^3). \tag{5}$$

Thus ignoring selection produces a downward bias of a known second-order magnitude. The proof, together with a general gamma-mixture detectability result and an oracle exposure calculation, is given in the additional subsection of Section 5. These results concern the threshold mechanism; they are not consequences of the standard JCIR affine transform.

For discretely observed paths, we combine threshold-conditioned mark likelihoods with a survival-adjusted jump-intensity estimator and a jump-filtered diffusion quasi-likelihood. We formulate consistency and model-selection results under increasing-horizon, high-frequency assumptions that must control the gap between idealised marked-jump observations and threshold-selected increments. In particular, diffusion noise and multiple jumps within an interval require additional error bounds; the idealised likelihood alone is not sufficient. The affine transform, invariant moments and autocovariance of the JCIR process are included as tools for interpreting the selected law and calculating predictive quantities, not as independent claims of novelty.

We assess the procedure through simulation and an illustration using annual Nile flow. The numerical study examines estimation as the observation horizon increases, while the Nile example compares candidate laws under alternative thresholds. Zeghdoudi gives the lowest reported AICc across the investigated threshold grid, but the leading one-parameter fits are close. That result supports it as a useful working specification, not as a uniquely established physical mark distribution. Moreover, the Nile series has a documented level change near 1898 (Cobb, 1978; Durbin and Koopman, 2001); its annual observations do not satisfy the high-frequency stationary

design used for the asymptotic theory. We therefore treat the example as an illustration of the fitting procedure rather than a validation of that theory.

The remainder of the paper introduces the candidate mark laws in Section 2, describes the stochastic dynamics in Section 3, and develops the inference procedure in Section 4. Section 5 gives its large-sample analysis and the threshold-specific mathematical results. Prediction, simulation and the empirical illustration follow in Sections 6–8, before the discussion and conclusion.

2 Gamma-mixture structure of the candidate mark laws

2.1 Common-rate gamma-mixture representation

For an integer $r \geq 1$ and rate $\theta > 0$, write

$$g_r(y;\theta) = \frac{\theta^r}{\Gamma(r)} y^{r-1} e^{-\theta y}, \qquad y > 0, \tag{6}$$

for the gamma density with shape r and rate θ. Most of the one-parameter mark distributions considered in this paper can be written in the form

$$f_m(y;\theta) = \sum_{r \in \mathcal{R}_m} w_{m,r}(\theta)\, g_r(y;\theta), \qquad \sum_{r \in \mathcal{R}_m} w_{m,r}(\theta) = 1, \tag{7}$$

where $\mathcal{R}_m$ is a finite set of positive integers and the weights are nonnegative for every $\theta > 0$.

Table 1 gives the corresponding representation for the one-parameter candidates used in the model comparison.

Table 1: Common-rate gamma-mixture representations of the one-parameter candidate mark distributions.

Model	Shapes	Nonzero mixture weights
Lindley	$\{1,2\}$	$w_1 = \frac{\theta}{1+\theta}$, $w_2 = \frac{1}{1+\theta}$
X-exponential	$\{1,2\}$	$w_1 = \frac{2}{3}$, $w_2 = \frac{1}{3}$
New XLindley	$\{1,2\}$	$w_1 = w_2 = \frac{1}{2}$
Q-Lindley	$\{1,2\}$	$w_1 = \frac{1}{1+\theta}$, $w_2 = \frac{\theta}{1+\theta}$
XLindley	$\{1,2\}$	$w_1 = \frac{\theta(\theta+2)}{(1+\theta)^2}$, $w_2 = \frac{1}{(1+\theta)^2}$
ZLindley	$\{1,2\}$	$w_1 = \frac{1+2\theta}{2(1+\theta)}$, $w_2 = \frac{1}{2(1+\theta)}$
Xgamma	$\{1,3\}$	$w_1 = \frac{\theta}{1+\theta}$, $w_3 = \frac{1}{1+\theta}$
Zeghdoudi	$\{2,3\}$	$w_2 = \frac{\theta}{\theta+2}$, $w_3 = \frac{2}{\theta+2}$
QGamma	$\{1,3\}$	$w_1 = \frac{\theta(\theta+4)}{(\theta+2)^2}$, $w_3 = \frac{4}{(\theta+2)^2}$

This representation is useful beyond notation. It places distributions that are usually introduced separately inside the same probabilistic framework. Their differences are determined by the gamma component shapes and by the way the mixing probabilities depend on θ. As a result, several quantities needed later for the jump–diffusion model can be derived once for the general mixture rather than separately for every candidate.

For completeness, some of the less familiar densities are

$$f_{QL}(y;\theta) = \frac{\theta(1+\theta^2 y)}{1+\theta} e^{-\theta y}, \tag{8}$$

$$f_{XL}(y;\theta) = \frac{\theta^2}{(1+\theta)^2}(\theta+2+y)e^{-\theta y}, \tag{9}$$

$$f_{ZL}(y;\theta) = \frac{\theta}{2(1+\theta)}(1+2\theta+\theta y)e^{-\theta y}, \tag{10}$$

$$f_{ZD}(y;\theta) = \frac{\theta^3}{\theta+2} y(1+y)e^{-\theta y}, \tag{11}$$

$$f_{QG}(y;\theta) = \frac{\theta^2}{(\theta+2)^2}\big(\theta+4+2\theta y^2\big)e^{-\theta y}. \tag{12}$$

The Q-Lindley specification in (8) is the model studied by Saaidia et al. (2026). Equation (9) corresponds to the XLindley distribution of Chouia and Zeghdoudi (2021), while the New XLindley density,

$$\frac{\theta}{2}(1+\theta y)e^{-\theta y},$$

is given by Khodja et al. (2023). The ZLindley model in (10) is from Saaidia et al. (2024), and (11) is the Zeghdoudi distribution introduced by Messaadia and Zeghdoudi (2018). The QGamma law in (12) is the continuous distribution used in the construction of Seghier et al. (2026). The X-exponential model is included according to Beghriche et al. (2025).

2.2 Moments, transforms, and survival functions

The common-rate representation gives the main distributional quantities in closed form.

Proposition 2.1 (Moments and Laplace transform)**.** *Let Y have density (7). If*

$$q > -\min \mathcal{R}_m,$$

then

$$\mathbb{E}(Y^q) = \theta^{-q} \sum_{r\in\mathcal{R}_m} w_{m,r}(\theta)\frac{\Gamma(r+q)}{\Gamma(r)}. \tag{13}$$

For every $s \geq 0$,

$$L_m(s;\theta) := \mathbb{E}(e^{-sY}) = \sum_{r\in\mathcal{R}_m} w_{m,r}(\theta)\left(\frac{\theta}{\theta+s}\right)^r. \tag{14}$$

Proof. By (7),

$$\mathbb{E}(Y^q) = \sum_{r\in\mathcal{R}_m} w_{m,r}(\theta)\int_0^\infty y^q g_r(y;\theta)\,dy.$$

For a gamma random variable with shape r and rate θ,

$$\int_0^\infty y^q g_r(y;\theta)\,dy = \theta^{-q}\frac{\Gamma(r+q)}{\Gamma(r)},$$

provided $r+q>0$. Since the smallest component shape is $\min\mathcal{R}_m$, the condition $q > -\min\mathcal{R}_m$ guarantees existence for every component and gives (13).

Similarly,

$$\int_0^\infty e^{-sy} g_r(y;\theta)\,dy = \left(\frac{\theta}{\theta+s}\right)^r, \qquad s\geq 0.$$

Averaging over the mixture components yields (14). □

The survival function is equally simple because all component shapes are positive integers.

Corollary 2.2 (Survival function)**.** *For $y \geq 0$,*

$$\overline{F}_m(y;\theta) = \sum_{r\in\mathcal{R}_m} w_{m,r}(\theta) e^{-\theta y} \sum_{j=0}^{r-1} \frac{(\theta y)^j}{j!}. \tag{15}$$

Proof. For an integer-shape gamma distribution,

$$\Pr\{\Gamma(r,\theta) > y\} = e^{-\theta y} \sum_{j=0}^{r-1} \frac{(\theta y)^j}{j!}.$$

Substitution into the mixture representation (7) gives (15). □

Formula (15) is particularly useful in the present setting. The threshold-adjusted likelihood involves $\overline{F}_m(u_i;\theta)$ for every detected increment, and the same quantity determines the effective exposure used later to estimate the latent jump intensity. No numerical integration is required for any of the one-parameter candidates in Table 1.

2.3 Finite sums of marks

The common-rate structure is also preserved when several jumps occur in the same observation interval.

Proposition 2.3 (Finite convolution representation)**.** *Let $Y_1,\ldots,Y_k$ be independent marks with density (7) and common parameter θ, and define*

$$S_k = \sum_{j=1}^{k} Y_j.$$

Introduce latent component labels $R_j \in \mathcal{R}_m$ satisfying

$$\mathbb{P}(R_j = r) = w_{m,r}(\theta).$$

Conditional on $(R_1,\ldots,R_k)$,

$$S_k \mid R_1,\ldots,R_k \sim \text{Gamma}\left(\sum_{j=1}^{k} R_j, \theta\right). \tag{16}$$

Consequently, S_k has a finite gamma-mixture distribution with common rate θ and possible shapes

$$r_1 + \cdots + r_k, \qquad r_j \in \mathcal{R}_m.$$

Proof. Conditional on its component label, Y_j has a gamma distribution with shape R_j and rate θ. Since independent gamma variables having the same rate add by summing their shape parameters,

$$Y_1 + \cdots + Y_k \mid (R_1,\ldots,R_k) \sim \text{Gamma}\left(R_1 + \cdots + R_k, \theta\right).$$

Averaging over the finitely many possible label combinations yields the unconditional finite-mixture representation. □

This result is useful when more than one jump occurs over an observation interval. In particular, short-time transition approximations can incorporate the k-jump contribution without resorting to repeated numerical convolution.

2.4 The Zeghdoudi mark law

The empirical analysis in Section 8 identifies the Zeghdoudi model as the most stable working specification over the threshold grid considered there. Because this distribution also has a distinctive small-mark behavior, we record the formulas needed for the stochastic and inferential developments.

From Table 1,

$$Y \sim \frac{\theta}{\theta+2}\text{Gamma}(2,\theta) + \frac{2}{\theta+2}\text{Gamma}(3,\theta). \tag{17}$$

The first two raw moments are therefore

$$m_1(\theta) := \mathbb{E}(Y) = \frac{2(\theta+3)}{\theta(\theta+2)}, \tag{18}$$

$$m_2(\theta) := \mathbb{E}(Y^2) = \frac{6(\theta+4)}{\theta^2(\theta+2)}, \tag{19}$$

and hence

$$\text{Var}(Y) = \frac{2(\theta^2+6\theta+6)}{\theta^2(\theta+2)^2}. \tag{20}$$

Its Laplace transform follows immediately from Proposition 2.1:

$$L_{ZD}(s;\theta) = \frac{\theta}{\theta+2}\left(\frac{\theta}{\theta+s}\right)^2 + \frac{2}{\theta+2}\left(\frac{\theta}{\theta+s}\right)^3. \tag{21}$$

Using Corollary 2.2, the survival function reduces to

$$\overline{F}_{ZD}(y;\theta) = e^{-\theta y}\left(1+\theta y+\frac{\theta^2 y^2}{\theta+2}\right), \qquad y \geq 0. \tag{22}$$

This expression will be used directly in the threshold-conditioned mark likelihood and in the detection-probability correction for the Poisson intensity.

The behavior near zero is especially relevant for the threshold procedure.

Lemma 2.4 (Small-mark probability)**.** *Let $Y \sim \text{ZD}(\theta)$. Then, as $r \downarrow 0$,*

$$\mathbb{P}(Y \leq r) = \frac{\theta^3}{2(\theta+2)}r^2 + O(r^3). \tag{23}$$

Proof. From (11),

$$f_{ZD}(y;\theta) = \frac{\theta^3}{\theta+2}y(1+y)e^{-\theta y}.$$

As $y \downarrow 0$,

$$(1+y)e^{-\theta y} = 1 + O(y),$$

and therefore

$$f_{ZD}(y;\theta) = \frac{\theta^3}{\theta+2}y + O(y^2).$$

Integrating over $(0,r)$ gives

$$\mathbb{P}(Y \leq r) = \int_0^r f_{ZD}(y;\theta)\,dy = \frac{\theta^3}{2(\theta+2)}r^2 + O(r^3),$$

which proves the result. □

Lemma 2.4 explains why the behavior of the mark density at the origin matters for threshold inference. For a distribution satisfying $f_Y(0) > 0$,

$$\mathbb{P}(Y \leq r) = f_Y(0)r + o(r), \qquad r \downarrow 0,$$

whereas the Zeghdoudi law has a second-order small-mark probability. The difference will translate directly into the missed-jump rate derived in Section 5.

3 Positive mean-reverting jump dynamics

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions. Let W be a standard Brownian motion, let N be a Poisson process with intensity $\lambda > 0$, and let $(Y_k)_{k\geq 1}$ be i.i.d. positive random variables, independent of (W, N). We consider the compound-Poisson process

$$J_t = \sum_{k=1}^{N_t} Y_k$$

and the positive mean-reverting jump–diffusion

$$\mathrm{d}X_t = \kappa(\mu - X_t)\,\mathrm{d}t + \sigma\sqrt{X_t}\,\mathrm{d}W_t + \mathrm{d}J_t, \qquad X_0 = x_0 \geq 0, \tag{24}$$

where

$$\kappa > 0, \qquad \mu > 0, \qquad \sigma > 0, \qquad \lambda > 0. \tag{25}$$

Throughout this section the mark distribution is kept generic, with Laplace transform

$$L_Y(s) = \mathbb{E}(e^{-sY}), \qquad s \geq 0.$$

This makes clear which properties are consequences of the stochastic dynamics and which depend on the chosen mark family. The JCIR–Zeghdoudi specification used later is obtained by substituting the density in (11), or equivalently the Laplace transform in (21).

3.1 Existence, uniqueness, and positivity

The finite activity of the jump component allows the process to be constructed between consecutive jump times.

Theorem 3.1 (Well-posedness and positivity)**.** *For every $x_0 \geq 0$, equation (24) admits a pathwise unique nonnegative strong solution.*

If, in addition,

$$2\kappa\mu \geq \sigma^2 \tag{26}$$

and $x_0 > 0$, then

$$\mathbb{P}(X_t > 0 \text{ for every } t > 0) = 1.$$

Proof. Let

$$0 < T_1 < T_2 < \cdots$$

denote the jump times of the Poisson process. Since N has finite intensity, only finitely many of these times occur on every bounded interval.

Before the first jump, (24) reduces to the classical CIR equation

$$\mathrm{d}X_t = \kappa(\mu - X_t)\,\mathrm{d}t + \sigma\sqrt{X_t}\,\mathrm{d}W_t, \qquad 0 \leq t < T_1.$$

The CIR equation has a pathwise unique nonnegative strong solution, despite the fact that $x \mapsto \sqrt{x}$ is not globally Lipschitz. Thus X_{T_1-} is well defined and nonnegative. At the first jump,

$$X_{T_1} = X_{T_1-} + Y_1.$$

Since $Y_1 > 0$ almost surely,

$$X_{T_1} \geq 0.$$

Starting from X_{T_1}, solve the CIR equation again on $[T_1, T_2)$ using the shifted Brownian increments. At T_2 add the second mark Y_2, and continue inductively. Because only finitely

many jumps occur on each compact time interval, this interlacing construction produces a global nonnegative strong solution.

Pathwise uniqueness follows in the same way. Two solutions driven by the same Brownian motion, Poisson process, and marks coincide before T_1 by uniqueness of the CIR equation. They therefore have the same value immediately after the first jump and coincide on $[T_1, T_2)$. Induction over the jump times proves equality on every bounded interval.

For strict positivity, recall that under (26) the boundary 0 is inaccessible to the CIR diffusion when its initial value is strictly positive. Hence, starting from $x_0 > 0$, the process remains positive until T_1. Moreover,

$$X_{T_1} = X_{T_1-} + Y_1 > 0.$$

The next CIR segment therefore also starts from a positive value and cannot reach zero. Repeating this argument over all jump intervals gives

$$X_t > 0$$

for every $t > 0$ almost surely. □

Remark 3.2. *The condition (26) is needed for strict positivity, not for existence of a nonnegative solution. The positive jumps can move the process away from the boundary, but they cannot force it toward zero.*

3.2 Generator and affine transform

For a sufficiently regular test function f, the infinitesimal generator of (24) is

$$\mathcal{A}f(x) = \kappa(\mu - x)f'(x) + \frac{1}{2}\sigma^2 x f''(x) + \lambda \int_0^\infty \{f(x+y) - f(x)\} F_Y(\,\mathrm{d}y). \tag{27}$$

The last term records the effect of the mark distribution on the dynamics. In particular, the mark law affects the transition distribution even when κ, μ, σ, and λ remain fixed.

Proposition 3.3 (Affine conditional Laplace transform)**.** *For $u \geq 0$ and $X_0 = x$,*

$$\mathbb{E}_x(e^{-uX_t}) = \exp\{-\phi(t,u) - \psi(t,u)x\}, \tag{28}$$

where

$$\psi(t,u) = \frac{ue^{-\kappa t}}{1 + \dfrac{\sigma^2 u}{2\kappa}(1 - e^{-\kappa t})} \tag{29}$$

and

$$\phi(t,u) = \int_0^t [\kappa\mu\psi(s,u) + \lambda\{1 - L_Y(\psi(s,u))\}]\,\mathrm{d}s. \tag{30}$$

Proof. Fix $u \geq 0$ and define

$$v(t,x) := \mathbb{E}_x(e^{-uX_t}).$$

The Markov property implies that v satisfies the backward equation

$$\partial_t v(t,x) = \mathcal{A}v(t,x), \qquad v(0,x) = e^{-ux}.$$

We look for a solution of exponential-affine form

$$v(t,x) = \exp\{-\phi(t) - \psi(t)x\},$$

with

$$\phi(0) = 0, \qquad \psi(0) = u.$$

The derivatives are

$$v_x = -\psi v, \qquad v_{xx} = \psi^2 v,$$

and

$$\partial_t v = -\{\phi'(t) + \psi'(t)x\}v.$$

For the jump term,

$$\begin{aligned}
&\lambda \int_0^\infty \{v(t, x+y) - v(t,x)\} F_Y(\,\mathrm{d}y) \\
&\quad = \lambda v(t,x) \int_0^\infty \{e^{-\psi(t)y} - 1\} F_Y(\,\mathrm{d}y) \\
&\quad = \lambda v(t,x)\{L_Y(\psi(t)) - 1\}.
\end{aligned}$$

Substituting these expressions into $\partial_t v = \mathcal{A}v$ and dividing by v gives

$$-\phi' - \psi' x = -\kappa\mu\psi + \kappa x\psi + \frac{1}{2}\sigma^2 x\psi^2 + \lambda\{L_Y(\psi) - 1\}.$$

Matching the coefficients of x yields the Riccati equation

$$\psi' = -\kappa\psi - \frac{1}{2}\sigma^2\psi^2, \qquad \psi(0) = u. \tag{31}$$

To solve it, set

$$z(t) = \frac{1}{\psi(t)}.$$

Then

$$z'(t) = \kappa z(t) + \frac{\sigma^2}{2}, \qquad z(0) = \frac{1}{u},$$

and hence

$$z(t) = \frac{e^{\kappa t}}{u} + \frac{\sigma^2}{2\kappa}(e^{\kappa t} - 1).$$

Taking the reciprocal gives

$$\psi(t,u) = \frac{u e^{-\kappa t}}{1 + \dfrac{\sigma^2 u}{2\kappa}(1 - e^{-\kappa t})},$$

which is (29).

The terms independent of x give

$$\phi'(t) = \kappa\mu\psi(t) + \lambda\{1 - L_Y(\psi(t))\}, \qquad \phi(0) = 0.$$

Integration gives (30). Substitution back into the backward equation verifies (28). □

For the Zeghdoudi specification,

$$L_Y(s) = L_{ZD}(s;\theta)$$

is given by (21); hence the complete conditional Laplace transform remains explicit up to the one-dimensional integral in (30).

3.3 Stationary distribution

The affine representation also identifies the limiting distribution.

Proposition 3.4 (Existence and uniqueness of the stationary law)**.** *Assume*

$$\mathbb{E}(Y) < \infty.$$

Then, for every initial state $x \geq 0$, X_t converges in distribution as $t \to \infty$ to a probability law π on $\mathbb{R}_+$ whose Laplace transform is

$$\mathcal{L}_\pi(u) = \exp\left\{-\int_0^u \frac{\kappa\mu q + \lambda\{1 - L_Y(q)\}}{q(\kappa + \sigma^2 q/2)}\,\mathrm{d}q\right\}, \qquad u \geq 0. \tag{32}$$

The distribution π is the unique invariant distribution of (24).

Proof. From (29),

$$\psi(t,u) \longrightarrow 0 \qquad (t \to \infty)$$

for every fixed $u \geq 0$. Hence the term $\psi(t,u)x$ in (28) disappears in the limit.

It remains to evaluate

$$\phi_\infty(u) := \lim_{t\to\infty} \phi(t,u).$$

The Riccati equation gives

$$\frac{\mathrm{d}\psi}{\mathrm{d}t} = -\psi\left(\kappa + \frac{\sigma^2}{2}\psi\right).$$

Therefore

$$\mathrm{d}t = -\frac{\mathrm{d}q}{q(\kappa + \sigma^2 q/2)}$$

when $q = \psi(t,u)$. Since q decreases from u to 0,

$$\begin{aligned}\phi_\infty(u) &= \int_0^\infty [\kappa\mu\psi(t,u) + \lambda\{1 - L_Y(\psi(t,u))\}]\,\mathrm{d}t \\ &= \int_0^u \frac{\kappa\mu q + \lambda\{1 - L_Y(q)\}}{q(\kappa + \sigma^2 q/2)}\,\mathrm{d}q.\end{aligned}$$

The integral is finite near $q = 0$. Indeed, if $m_1 = \mathbb{E}(Y) < \infty$, then

$$1 - L_Y(q) = m_1 q + o(q), \qquad q \downarrow 0,$$

so the integrand converges to

$$\mu + \frac{\lambda m_1}{\kappa}.$$

Consequently,

$$\lim_{t\to\infty} \mathbb{E}_x(e^{-uX_t}) = \mathcal{L}_\pi(u)$$

with $\mathcal{L}_\pi$ given by (32). The limit is independent of x and is continuous at $u = 0$, with

$$\mathcal{L}_\pi(0) = 1.$$

The continuity theorem for Laplace transforms therefore gives a probability measure π such that

$$X_t \Longrightarrow \pi.$$

To verify invariance, let P_t denote the transition semigroup. For fixed $s > 0$,

$$\delta_x P_{t+s} = (\delta_x P_t) P_s.$$

As $t \to \infty$, the left-hand side converges to π, while the right-hand side converges to πP_s by the Feller property of the process. Hence

$$\pi P_s = \pi.$$

Thus π is invariant.

Finally, if $\widetilde{\pi}$ were another invariant law, then

$$\widetilde{\pi} P_t = \widetilde{\pi}.$$

Integrating (28) with respect to $\widetilde{\pi}$ and letting $t \to \infty$ shows that $\widetilde{\pi}$ has Laplace transform (32). Uniqueness of Laplace transforms gives

$$\widetilde{\pi} = \pi.$$

□

Remark 3.5. *The limiting distribution is generally not a standard CIR stationary gamma distribution. The mark law enters (32) through L_Y, so two models with the same diffusion parameters can have different stationary distributions solely because their jump-size laws differ.*

3.4 Conditional and stationary moments

Assume from now on that

$$m_1 := \mathbb{E}(Y) < \infty, \qquad m_2 := \mathbb{E}(Y^2) < \infty.$$

Proposition 3.6 (First two moments)**.** *Define*

$$\bar{m} := \mu + \frac{\lambda m_1}{\kappa}, \qquad a := 2\kappa\mu + \sigma^2 + 2\lambda m_1. \tag{33}$$

For $X_0 = x$,

$$\mathbb{E}_x(X_t) = \bar{m} + (x - \bar{m})e^{-\kappa t}, \tag{34}$$

and

$$\begin{aligned}\mathbb{E}_x(X_t^2) = e^{-2\kappa t}x^2 + \frac{a\bar{m} + \lambda m_2}{2\kappa}(1 - e^{-2\kappa t}) \\ + \frac{a(x - \bar{m})}{\kappa}(e^{-\kappa t} - e^{-2\kappa t}).\end{aligned} \tag{35}$$

Under the invariant distribution of Proposition 3.4,

$$\mathbb{E}(X_\infty) = \bar{m}, \qquad \mathrm{Var}(X_\infty) = \frac{\sigma^2\bar{m} + \lambda m_2}{2\kappa}. \tag{36}$$

Proof. Because x and x^2 are not compactly supported, the following generator calculations can be justified by applying Dynkin's formula first to the stopped process

$$X_{t\wedge\tau_R}, \qquad \tau_R = \inf\{t : X_t \geq R\},$$

and then letting $R \to \infty$. The assumption $m_2 < \infty$ provides the required moment bounds.

For $f(x) = x$,

$$f'(x) = 1, \qquad f''(x) = 0,$$

and

$$f(x + y) - f(x) = y.$$

Hence

$$\mathcal{A}x = \kappa\mu - \kappa x + \lambda m_1.$$

Dynkin's formula gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_x(X_t) = \kappa\mu + \lambda m_1 - \kappa\mathbb{E}_x(X_t).$$

If

$$m(t) := \mathbb{E}_x(X_t),$$

then

$$m'(t) + \kappa m(t) = \kappa\mu + \lambda m_1 = \kappa\bar{m}.$$

Solving this linear equation with $m(0) = x$ yields

$$m(t) = \bar{m} + (x - \bar{m})e^{-\kappa t},$$

which proves (34).

For $f(x) = x^2$,

$$f'(x) = 2x, \qquad f''(x) = 2,$$

and

$$(x + y)^2 - x^2 = 2xy + y^2.$$

Therefore

$$\begin{aligned}\mathcal{A}x^2 &= 2\kappa x(\mu - x) + \sigma^2 x + \lambda(2xm_1 + m_2)\\ &= -2\kappa x^2 + \big(2\kappa\mu + \sigma^2 + 2\lambda m_1\big)x + \lambda m_2\\ &= -2\kappa x^2 + ax + \lambda m_2.\end{aligned}$$

Let

$$q(t) := \mathbb{E}_x(X_t^2).$$

Dynkin's formula and (34) give

$$q'(t) + 2\kappa q(t) = a\{\bar{m} + (x - \bar{m})e^{-\kappa t}\} + \lambda m_2.$$

Multiplying by $e^{2\kappa t}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\{e^{2\kappa t}q(t)\} = e^{2\kappa t}(a\bar{m} + \lambda m_2) + a(x - \bar{m})e^{\kappa t}.$$

Integration over $[0, t]$ gives

$$\begin{aligned}e^{2\kappa t}q(t) - x^2 = \frac{a\bar{m} + \lambda m_2}{2\kappa}(e^{2\kappa t} - 1)\\ + \frac{a(x - \bar{m})}{\kappa}(e^{\kappa t} - 1).\end{aligned}$$

Multiplying by $e^{-2\kappa t}$ yields (35).

Letting $t \to \infty$ gives

$$\mathbb{E}(X_\infty^2) = \frac{a\bar{m} + \lambda m_2}{2\kappa}.$$

Hence

$$\mathrm{Var}(X_\infty) = \frac{a\bar{m} + \lambda m_2}{2\kappa} - \bar{m}^2.$$

Using

$$a = 2\kappa\mu + \sigma^2 + 2\lambda m_1$$

and

$$\kappa\bar{m} = \kappa\mu + \lambda m_1,$$

we have

$$a = 2\kappa\bar{m} + \sigma^2.$$

Therefore

$$\begin{aligned}\mathrm{Var}(X_\infty) &= \frac{(2\kappa\bar{m} + \sigma^2)\bar{m} + \lambda m_2}{2\kappa} - \bar{m}^2 \\ &= \frac{\sigma^2\bar{m} + \lambda m_2}{2\kappa},\end{aligned}$$

which proves (36). □

The formulas make the respective roles of the model components transparent. The long-run mean is shifted above the diffusion level μ by

$$\frac{\lambda m_1}{\kappa},$$

while the stationary variance receives the additional jump contribution

$$\frac{\lambda m_2}{2\kappa}.$$

Thus both the frequency and the shape of the marks affect long-run variability.

For the JCIR–Zeghdoudi model, substitution of (18)–(19) gives

$$\bar{m}_{ZD} = \mu + \frac{2\lambda(\theta+3)}{\kappa\theta(\theta+2)}, \tag{37}$$

and

$$\mathrm{Var}(X_\infty) = \frac{1}{2\kappa}\left[\sigma^2\bar{m}_{ZD} + \frac{6\lambda(\theta+4)}{\theta^2(\theta+2)}\right]. \tag{38}$$

3.5 Stationary dependence

Proposition 3.7 (Stationary autocovariance)**.** *Suppose that X is stationary and has finite second moment. Then, for every $h \geq 0$,*

$$\mathrm{Cov}(X_{t+h}, X_t) = e^{-\kappa h}\,\mathrm{Var}(X_\infty). \tag{39}$$

If $\mathrm{Var}(X_\infty) > 0$*, then*

$$\mathbb{C}\!\!\nwarrow\!\!\searrow(X_{t+h}, X_t) = e^{-\kappa h}. \tag{40}$$

Proof. By the Markov property and (34), conditioning at time t gives

$$\mathbb{E}(X_{t+h} \mid \mathcal{F}_t) = \bar{m} + (X_t - \bar{m})e^{-\kappa h}.$$

In particular,

$$\mathbb{E}(X_{t+h} \mid X_t) = \bar{m} + (X_t - \bar{m})e^{-\kappa h}.$$

Under stationarity,

$$\mathbb{E}(X_t) = \mathbb{E}(X_{t+h}) = \bar{m}.$$

Therefore, using the tower property,

$$\begin{aligned}\mathrm{Cov}(X_{t+h}, X_t) &= \mathbb{E}\left[(X_t - \bar{m})(X_{t+h} - \bar{m})\right] \\ &= \mathbb{E}\left[(X_t - \bar{m})\mathbb{E}(X_{t+h} - \bar{m} \mid X_t)\right] \\ &= e^{-\kappa h}\mathbb{E}\left[(X_t - \bar{m})^2\right] \\ &= e^{-\kappa h}\,\mathrm{Var}(X_\infty),\end{aligned}$$

which proves (39).

Stationarity gives

$$\mathrm{Var}(X_{t+h}) = \mathrm{Var}(X_t) = \mathrm{Var}(X_\infty).$$

Dividing (39) by the common stationary variance yields (40). □

This result separates two aspects of the dynamics. The mean-reversion parameter κ determines the exponential decay of linear dependence, whereas the mark distribution changes the scale of stationary variation through m_1 and m_2. In particular, changing the mark law does not alter the decay rate in (40), but it can materially change the stationary variance.

3.6 Identifiability

Before developing estimation, it is useful to verify that the parameters are uniquely determined by the continuous-time model.

Theorem 3.8 (Generator identifiability)**.** *Let*

$$\{F_\eta : \eta \in H\}$$

be an identifiable family of probability distributions supported on $(0, \infty)$*, with Laplace transform*

$$L_\eta(s) = \int_0^\infty e^{-sy} F_\eta(\,\mathrm{d}y).$$

Suppose that two parameter vectors

$$\vartheta = (\kappa, \mu, \sigma, \lambda, \eta)$$

and

$$\widetilde{\vartheta} = (\widetilde{\kappa}, \widetilde{\mu}, \widetilde{\sigma}, \widetilde{\lambda}, \widetilde{\eta})$$

satisfying (25) *generate the same extended generator. Then*

$$\vartheta = \widetilde{\vartheta}.$$

Consequently, the JCIR–Zeghdoudi parameter vector is identifiable.

Proof. For $s > 0$, consider

$$f_s(x) = e^{-sx}.$$

Its derivatives are

$$f_s'(x) = -s f_s(x), \qquad f_s''(x) = s^2 f_s(x).$$

Moreover,

$$\begin{aligned}\int_0^\infty \{f_s(x+y) - f_s(x)\} F_\eta(\,\mathrm{d}y) &= f_s(x) \int_0^\infty (e^{-sy} - 1) F_\eta(\,\mathrm{d}y) \\ &= f_s(x)\{L_\eta(s) - 1\}.\end{aligned}$$

Substitution into (27) yields

$$\frac{\mathcal{A} f_s(x)}{f_s(x)} = -s\kappa\mu + x\left(s\kappa + \frac{\sigma^2 s^2}{2}\right) + \lambda\{L_\eta(s) - 1\}. \tag{41}$$

Assume that $\mathcal{A} = \widetilde{\mathcal{A}}$. Equality of (41) for every $x > 0$ implies equality of the coefficients of x:

$$s\kappa + \frac{\sigma^2 s^2}{2} = s\widetilde{\kappa} + \frac{\widetilde{\sigma}^2 s^2}{2}, \qquad s > 0.$$

After division by s,

$$\kappa + \frac{\sigma^2}{2}s = \widetilde{\kappa} + \frac{\widetilde{\sigma}^2}{2}s.$$

Since this identity holds for every $s > 0$, equality of the constant and linear coefficients gives

$$\kappa = \widetilde{\kappa}, \qquad \sigma^2 = \widetilde{\sigma}^2.$$

Because both diffusion coefficients are positive,

$$\sigma = \widetilde{\sigma}.$$

The remaining constant terms satisfy

$$-s\kappa\mu + \lambda\{L_\eta(s) - 1\} = -s\kappa\widetilde{\mu} + \widetilde{\lambda}\{L_{\widetilde{\eta}}(s) - 1\}. \tag{42}$$

The Laplace transforms lie in $[0, 1]$, so both jump terms in (42) are bounded uniformly in s. Dividing by s and letting $s \to \infty$ therefore gives

$$-\kappa\mu = -\kappa\widetilde{\mu}.$$

Since $\kappa > 0$,

$$\mu = \widetilde{\mu}.$$

Equation (42) now reduces to

$$\lambda\{L_\eta(s) - 1\} = \widetilde{\lambda}\{L_{\widetilde{\eta}}(s) - 1\}, \qquad s > 0. \tag{43}$$

Because each mark distribution is supported on $(0, \infty)$,

$$L_\eta(s) \longrightarrow 0, \qquad L_{\widetilde{\eta}}(s) \longrightarrow 0 \qquad (s \to \infty).$$

Taking the limit in (43) gives

$$-\lambda = -\widetilde{\lambda},$$

and hence

$$\lambda = \widetilde{\lambda}.$$

Returning to (43),

$$L_\eta(s) = L_{\widetilde{\eta}}(s), \qquad s > 0.$$

Uniqueness of Laplace transforms on $\mathbb{R}_+$ implies

$$F_\eta = F_{\widetilde{\eta}}.$$

Since the mark family is identifiable,

$$\eta = \widetilde{\eta}.$$

All components of the parameter vector are therefore equal.

For the Zeghdoudi family this last condition is particularly transparent: from

$$f_{ZD}(y; \theta) = \frac{\theta^3}{\theta + 2} y(1 + y)e^{-\theta y},$$

two parameter values can generate the same density only if their exponential decay rates agree. Thus

$$\theta = \widetilde{\theta}.$$

□

Remark 3.9. *The theorem concerns identifiability of the continuous-time model, not the finite-sample ability to estimate every parameter equally well. In practice, κ and the jump parameters can be considerably harder to recover than σ, especially when the observation horizon is short or few jumps are detected. This distinction is examined in the Monte Carlo study.*

%==========

4 Threshold-based inference from discrete observations

Suppose that the process is observed at the equally spaced times

$$t_i = i\Delta_n, \qquad i = 0, \dots, n,$$

and write

$$\Delta_i X := X_{t_i} - X_{t_{i-1}}. \tag{44}$$

The total observation horizon is

$$T_n = n\Delta_n.$$

The inferential procedure separates the increments into two groups. Large positive increments are treated as potential jump intervals and are used for mark and intensity estimation, whereas the remaining increments are used for the diffusion quasi-likelihood. Since this classification is based on a threshold, the mark likelihood must retain the same selection mechanism.

4.1 Threshold classification

An increment is classified as containing an upward jump whenever

$$\widehat{I}_{i,n} = \mathbf{1}\{\Delta_i X > u_{i,n}\} = 1, \tag{45}$$

where $u_{i,n} > 0$ is the detection threshold.

For the asymptotic theory we consider thresholds satisfying

$$r_n := \max_{1 \le i \le n} u_{i,n} \longrightarrow 0, \qquad \frac{\min_i u_{i,n}}{\sqrt{\Delta_n \log(1/\Delta_n)}} \longrightarrow \infty. \tag{46}$$

The first condition allows increasingly small jumps to become detectable, whereas the second keeps the threshold above the typical scale of the continuous martingale fluctuation.

For implementation, we use the state-dependent rule

$$u_{i,n} = c\,\widehat{\sigma}_{\mathrm{pre}} \sqrt{X_{t_{i-1}} \Delta_n \log n}, \tag{47}$$

where $\widehat{\sigma}_{\mathrm{pre}}$ is a robust preliminary volatility estimate and $c > 0$ controls the detection level. The empirical analysis does not rely on a single choice of c; the fitted mark law is examined over a grid of threshold multipliers.

The asymptotic arguments below are stated for thresholds satisfying (46). A predictable data-dependent rule such as (47) can be handled under the same arguments provided its random upper and lower bounds satisfy the corresponding conditions in probability.

4.2 Threshold-conditioned mark likelihood

Let

$$\mathcal{J}_n = \{i : \widehat{I}_{i,n} = 1\}, \qquad M_n = |\mathcal{J}_n|,$$

and put

$$Z_i = \Delta_i X, \qquad i \in \mathcal{J}_n.$$

On an interval containing a single correctly detected jump,

$$\Delta_i X = Y_i + \kappa(\mu - X_{t_{i-1}})\Delta_n + \sigma\sqrt{X_{t_{i-1}}}\,\Delta_i W + o_p(\sqrt{\Delta_n}).$$

Under (46), the diffusion contribution is of smaller order than the detection boundary. Thus, to first order,

$$Z_i \approx Y_i, \qquad Y_i > u_{i,n}.$$

This conditioning cannot be ignored. If candidate m has density $f_m(\cdot\,;\eta_m)$ and survival function $\overline{F}_m(\cdot\,;\eta_m)$, the appropriate conditional density is

$$f_m^T(z;\eta_m,u) = \frac{f_m(z;\eta_m)}{\overline{F}_m(u;\eta_m)}, \qquad z > u. \tag{48}$$

Accordingly, the threshold-adjusted log-likelihood is

$$\ell_{m,n}^T(\eta_m) = \sum_{i\in\mathcal{J}_n} \left[\log f_m(Z_i;\eta_m) - \log \overline{F}_m(u_{i,n};\eta_m)\right]. \tag{49}$$

The second term in (49) is not a minor correction. The sample itself has been created by removing increments below the detection boundary. Fitting f_m directly would therefore compare an unconditional density with a conditionally selected sample.

For a model with k_m free mark parameters, let $\widehat{\eta}_m$ maximize (49). Provided $M_n > k_m + 1$, we compute

$$\mathrm{AIC}_m = -2\ell_{m,n}^T(\widehat{\eta}_m) + 2k_m, \tag{50}$$

$$\mathrm{AICc}_m = \mathrm{AIC}_m + \frac{2k_m(k_m+1)}{M_n - k_m - 1}, \tag{51}$$

$$\mathrm{BIC}_m = -2\ell_{m,n}^T(\widehat{\eta}_m) + k_m \log M_n. \tag{52}$$

When $M_n \le k_m + 1$, AICc is undefined and the corresponding fit is not used for AICc comparison.

Because the number of detected marks may be small, AICc is used for the finite-sample empirical ranking. BIC is also reported and plays a separate role in the large-sample model-selection result below.

For interpretation, define

$$\Delta_m = \mathrm{AICc}_m - \min_{\ell\in\mathcal{M}} \mathrm{AICc}_\ell, \tag{53}$$

and the corresponding normalized Akaike weight

$$w_m = \frac{\exp(-\Delta_m/2)}{\sum_{\ell\in\mathcal{M}} \exp(-\Delta_\ell/2)}. \tag{54}$$

These weights are used descriptively to show whether the ranking is sharply concentrated or spread across several competing mark laws.

4.3 Conditional PIT diagnostics

The threshold conditioning also changes the probability integral transform. For a detected mark, define

$$U_i = \frac{F_m(Z_i;\widehat{\eta}_m) - F_m(u_{i,n};\widehat{\eta}_m)}{1 - F_m(u_{i,n};\widehat{\eta}_m)}. \tag{55}$$

If the fitted conditional model is adequate, the U_i should be close to uniform on $(0,1)$.

We therefore use the Kolmogorov–Smirnov, Cramér–von Mises and Anderson–Darling statistics computed from (55). Since the parameters are estimated from the same, often small, sample, these quantities are treated primarily as descriptive diagnostics rather than as exact distribution-free tests. Formal finite-sample p-values would require a parametric bootstrap reproducing both parameter estimation and threshold selection.

4.4 Zeghdoudi likelihood

For an ordinary untruncated sample $y_1, \dots, y_M$ from the Zeghdoudi density (11),

$$\ell_{ZD}(\theta) = 3M \log \theta - M \log(\theta + 2) + \sum_{j=1}^{M} \log\{y_j(1 + y_j)\} - \theta \sum_{j=1}^{M} y_j. \tag{56}$$

Proposition 4.1 (Unique ordinary MLE for the Zeghdoudi rate)**.** *For $M \geq 1$ and $y_j > 0$, the log-likelihood (56) is strictly concave on $(0, \infty)$ and has the unique maximizer*

$$\widehat{\theta} = \frac{1 - \overline{y} + \sqrt{\overline{y}^{\,2} + 4\overline{y} + 1}}{\overline{y}}, \qquad \overline{y} = \frac{1}{M} \sum_{j=1}^{M} y_j. \tag{57}$$

Proof. Differentiating (56) gives

$$\ell'_{ZD}(\theta) = \frac{3M}{\theta} - \frac{M}{\theta + 2} - \sum_{j=1}^{M} y_j, \tag{58}$$

and

$$\ell''_{ZD}(\theta) = -\frac{3M}{\theta^2} + \frac{M}{(\theta + 2)^2}. \tag{59}$$

Since

$$\frac{1}{(\theta + 2)^2} < \frac{1}{\theta^2} < \frac{3}{\theta^2},$$

we have

$$\ell''_{ZD}(\theta) < 0 \qquad \text{for every } \theta > 0.$$

Thus the log-likelihood is strictly concave.

Moreover,

$$\ell'_{ZD}(\theta) \longrightarrow +\infty \qquad (\theta \downarrow 0),$$

whereas

$$\ell'_{ZD}(\theta) \longrightarrow -\sum_{j=1}^{M} y_j < 0 \qquad (\theta \to \infty).$$

By continuity and strict monotonicity of the score, there is exactly one positive solution of the likelihood equation.

Dividing (58) by M and setting it equal to zero yields

$$\frac{3}{\theta} - \frac{1}{\theta + 2} - \overline{y} = 0.$$

Multiplication by $\theta(\theta + 2)$ gives

$$3(\theta + 2) - \theta - \overline{y}\,\theta(\theta + 2) = 0,$$

or equivalently

$$\overline{y}\,\theta^2 + (2\overline{y} - 2)\theta - 6 = 0. \tag{60}$$

The discriminant is

$$(2\overline{y} - 2)^2 + 24\overline{y} = 4(\overline{y}^{\,2} + 4\overline{y} + 1).$$

The positive root of (60) is therefore exactly (57). □

For threshold-selected data, the log-likelihood becomes

$$\ell^T_{ZD,n}(\theta) = \sum_{i\in\mathcal{J}_n} \log f_{ZD}(Z_i;\theta) - \sum_{i\in\mathcal{J}_n} \log \overline{F}_{ZD}(u_{i,n};\theta). \tag{61}$$

Using

$$\overline{F}_{ZD}(u;\theta) = e^{-\theta u} A(u,\theta),$$

where

$$A(u,\theta) = 1 + \theta u + \frac{\theta^2 u^2}{\theta+2}, \tag{62}$$

its derivative is

$$\partial_\theta A(u,\theta) = u + \frac{\theta(\theta+4)}{(\theta+2)^2} u^2. \tag{63}$$

Hence

$$\partial_\theta \log \overline{F}_{ZD}(u;\theta) = -u + \frac{u + \dfrac{\theta(\theta+4)}{(\theta+2)^2}u^2}{1 + \theta u + \dfrac{\theta^2 u^2}{\theta+2}}. \tag{64}$$

The threshold-adjusted score is therefore

$$\begin{aligned} S^T_{ZD,n}(\theta) = \frac{3M_n}{\theta} - \frac{M_n}{\theta+2} - \sum_{i\in\mathcal{J}_n} Z_i \\ - \sum_{i\in\mathcal{J}_n} \partial_\theta \log \overline{F}_{ZD}(u_{i,n};\theta). \end{aligned} \tag{65}$$

Unlike the untruncated likelihood, this equation does not reduce to a useful closed form. It is nevertheless a smooth one-dimensional optimization problem, and (57) provides a natural starting value.

4.5 Detection correction for the jump intensity

The observed number of threshold crossings is not the latent number of jumps. Suppose, for the moment, that candidate m and its parameter η_m are known. On interval i, a single jump is detected with probability approximately

$$p_{i,n} = \overline{F}_m(u_{i,n};\eta_m).$$

Since a Poisson arrival occurs with probability $\lambda\Delta_n + O(\Delta_n^2)$,

$$\mathbb{P}(\text{detected jump in interval } i) = \lambda\Delta_n p_{i,n} + O(\Delta_n^2).$$

Consequently,

$$\mathbb{E}(M_n) \approx \lambda \sum_{i=1}^n \Delta_n \overline{F}_m(u_{i,n};\eta_m). \tag{66}$$

This motivates the effective exposure

$$\widehat{E}_n = \sum_{i=1}^n \Delta_n \overline{F}_{\widehat{m}}\big(u_{i,n};\widehat{\eta}_{\widehat{m}}\big) \tag{67}$$

and the corrected estimator

$$\widehat{\lambda}_n = \frac{M_n}{\widehat{E}_n}. \tag{68}$$

When the thresholds vanish uniformly,

$$\overline{F}_{m_0}(u_{i,n};\eta_0) \to 1,$$

so that

$$E_n = \sum_{i=1}^{n} \Delta_n \overline{F}_{m_0}(u_{i,n};\eta_0) \sim T_n.$$

Thus the corrected estimator becomes asymptotically equivalent to M_n/T_n. In moderate samples, however, the distinction can be substantial when the fitted mark distribution assigns non-negligible probability below the detection boundary.

4.6 Diffusion quasi-likelihood

Let

$$\mathcal{C}_n = \{i : \widehat{I}_{i,n} = 0\}$$

denote the intervals retained for diffusion estimation. On a correctly classified jump-free interval,

$$\Delta_i X = \kappa(\mu - X_{t_{i-1}})\Delta_n + \sigma\sqrt{X_{t_{i-1}}\Delta_n}\,\varepsilon_i + R_{i,n}, \tag{69}$$

where ε_i is approximately standard normal and the discretization remainder $R_{i,n}$ is negligible at the relevant high-frequency scale.

The corresponding Gaussian quasi-log-likelihood is

$$Q_n(\kappa,\mu,\sigma) = -\frac{1}{2}\sum_{i\in\mathcal{C}_n}\Big[\log\{2\pi\sigma^2 X_{t_{i-1}}\Delta_n\} + \frac{\{\Delta_i X - \kappa(\mu - X_{t_{i-1}})\Delta_n\}^2}{\sigma^2 X_{t_{i-1}}\Delta_n}\Big]. \tag{70}$$

Numerically, weighted least squares provides stable starting values for (κ,μ), after which (70) is optimized under positivity constraints. A logarithmic parametrization is convenient for κ, μ, and σ.

5 Large-sample properties

The theoretical results are stated under an infill-and-long-span scheme. They justify the estimation strategy for increasingly dense observations over an expanding time horizon. They are not intended to imply that the annual Nile series used later satisfies a high-frequency asymptotic design.

Assumption 5.1. *As $n \to \infty$, assume the following conditions.*

(A1)

$$\Delta_n \to 0, \qquad T_n = n\Delta_n \to \infty.$$

(A2) The true dynamic parameter belongs to the interior of a compact parameter set. The process is stationary and ergodic and has the positive and inverse moments required for the likelihood and quasi-likelihood arguments below.

(A3) The thresholds satisfy (46).

(A4) The candidate family $\mathcal{M}$ is finite and contains the true mark family m_0. Each model has a compact parameter space, is identifiable, and its log-density and log-survival function satisfy the continuity and domination conditions required for a uniform law of large numbers.

(A5) For every incorrect candidate $m \neq m_0$, the limiting conditional log-likelihood is separated from that of m_0 by a strictly positive Kullback–Leibler gap.

5.1 Classification errors

The behavior of the mark density near the origin determines how many latent jumps remain below a shrinking threshold.

Lemma 5.2 (Sub-threshold Zeghdoudi marks)**.** *Suppose that the true mark distribution is Zeghdoudi and that Assumption 5.1(A1)–(A3) holds. Let*

$$r_n = \max_i u_{i,n}.$$

Then the expected number of latent jumps whose marks do not exceed their detection threshold is

$$O(T_n r_n^2). \tag{71}$$

Moreover, the expected number of observation intervals containing at least two jumps is

$$O(T_n \Delta_n). \tag{72}$$

Proof. Let N_n^{small} denote the number of latent jumps over $[0, T_n]$ whose marks lie below the relevant threshold. Since $u_{i,n} \le r_n$,

$$\mathbb{P}(Y \le u_{i,n}) \le \mathbb{P}(Y \le r_n).$$

By Lemma 2.4,

$$\mathbb{P}(Y \le r_n) = \frac{\theta_0^3}{2(\theta_0+2)} r_n^2 + O(r_n^3) = O(r_n^2).$$

The expected total number of Poisson jumps over $[0, T_n]$ is λT_n. Marking each jump according to whether its size is at most r_n gives

$$\mathbb{E}(N_n^{\text{small}}) \le \lambda T_n \mathbb{P}(Y \le r_n) = O(T_n r_n^2),$$

which proves (71).

For the second statement, on one interval,

$$\Delta_i N := N_{t_i} - N_{t_{i-1}} \sim \text{Poisson}(\lambda \Delta_n).$$

Hence

$$\begin{aligned}\mathbb{P}(\Delta_i N \ge 2) &= 1 - e^{-\lambda \Delta_n}\{1 + \lambda \Delta_n\} \\ &= \frac{\lambda^2}{2} \Delta_n^2 + O(\Delta_n^3).\end{aligned}$$

Summing over

$$n = \frac{T_n}{\Delta_n}$$

intervals gives

$$n\, O(\Delta_n^2) = O(T_n \Delta_n),$$

which proves (72). □

Remark 5.3 (Total classification discrepancy)**.** *Lemma 5.2 isolates the two errors that arise directly from the jump process. If the number of additional errors caused by the continuous component—false positive crossings and true jumps masked by a large negative diffusion increment—is $o_p(T_n)$, then the total classification discrepancy satisfies*

$$O_p(T_n r_n^2 + T_n \Delta_n) + o_p(T_n). \tag{73}$$

For a mark density satisfying $f_Y(0) > 0$,

$$\mathbb{P}(Y \le r) = f_Y(0) r + o(r),$$

and the first term is instead of order $T_n r_n$. The quadratic order for Zeghdoudi marks is therefore a genuine consequence of the vanishing density at the origin.

5.2 Small-mark order and the cost of ignoring selection

The affine property of the state process does not determine how many jumps are hidden by a detection rule. That question depends on the distribution of the marks near the origin. The following result makes this dependence precise for the finite gamma mixtures introduced in Section 2.

Proposition 5.4 (Gamma-mixture detection order)**.** *Let*

$$f_\theta(y) = \sum_{r\in\mathcal{R}} w_r(\theta)\frac{\theta^r y^{r-1}e^{-\theta y}}{\Gamma(r)}, \qquad y>0,$$

where $\mathcal{R}$ is a finite set of positive integers, $\theta>0$, $w_r(\theta)\ge 0$, and $\sum_r w_r(\theta)=1$. Fix θ and let $r_* = \min\{r\in\mathcal{R} : w_r(\theta)>0\}$. Then, as $u\downarrow 0$,

$$\Pr_\theta(Y\le u) = \frac{w_{r_*}(\theta)\theta^{r_*}}{r_*!}\,u^{r_*} + O(u^{r_*+1}). \tag{74}$$

If a compound-Poisson process of rate λ is observed over disjoint intervals of lengths Δ_i with fixed deterministic thresholds u_i, and $r_n = \max_i u_i \downarrow 0$, the expected number of latent marks not exceeding their corresponding thresholds is

$$\lambda\sum_i \Delta_i \Pr_\theta(Y\le u_i) = \lambda\frac{w_{r_*}(\theta)\theta^{r_*}}{r_*!}\sum_i \Delta_i u_i^{r_*} + O\left(\lambda\sum_i \Delta_i u_i^{r_*+1}\right). \tag{75}$$

In particular, if $T_n=\sum_i\Delta_i$, this expectation is $O(T_n r_n^{r_*})$.

Proof. For a gamma random variable with integer shape r and rate θ,

$$\Pr\{\Gamma(r,\theta)\le u\} = \frac{\theta^r}{\Gamma(r)}\int_0^u y^{r-1}e^{-\theta y}\,dy = \frac{(\theta u)^r}{r!} + O(u^{r+1}).$$

The term with the smallest active shape r_* is the only contribution of order u^{r_*}; all remaining terms are of order u^{r_*+1} or smaller. Averaging with the finite mixture weights proves (74). On interval i, the expected number of jumps is $\lambda\Delta_i$. Independence of each mark from its arrival time therefore gives $\lambda\Delta_i\Pr_\theta(Y\le u_i)$ expected sub-threshold jumps. Summing and applying the expansion yields (75). The upper bound follows because $u_i\le r_n$ and $\sum_i\Delta_i = T_n$. □

The result covers several competing mark laws at once. For Zeghdoudi marks, $r_*=2$ and $w_2(\theta)=\theta/(\theta+2)$, so the leading constant is $\theta^3/[2(\theta+2)]$. For a candidate mixture with a positive shape-one weight, $r_*=1$, and the missed-jump count is generally of first order in the maximum threshold. The comparison concerns the detection mechanism, not the affine representation of X.

A second consequence can be expressed directly as a bias calculation.

Proposition 5.5 (Second-order truncation bias for Zeghdoudi marks)**.** *Let $Y_1,\dots,Y_m$ be independent observations drawn from the Zeghdoudi law with true rate $\theta_0>0$, conditional on $Y_j>u$, where $u>0$ is a fixed common threshold. Suppose that the ordinary, untruncated Zeghdoudi likelihood is fitted to these observations. Its maximiser converges, as $m\to\infty$, to the unique parameter $\theta^\dagger(u)$ satisfying*

$$\frac{3}{\theta^\dagger(u)} - \frac{1}{\theta^\dagger(u)+2} = \mathbb{E}_{\theta_0}(Y \mid Y>u).$$

As $u\downarrow 0$,

$$\theta^\dagger(u) - \theta_0 = -\frac{\theta_0^4(\theta_0+3)}{2(\theta_0^2+6\theta_0+6)}u^2 + O(u^3). \tag{76}$$

The leading coefficient is strictly negative. By contrast, the correctly normalised conditional likelihood

$$\sum_{j=1}^{m}\{\log f_{\mathrm{ZD}}(Y_j;\theta)-\log \overline{F}_{\mathrm{ZD}}(u;\theta)\}$$

has its expected maximum at θ_0; under standard compactness and regularity conditions, its maximiser is consistent for θ_0.

Proof. Write

$$a(\theta)=\frac{3}{\theta}-\frac{1}{\theta+2},\qquad I(\theta)=\frac{3}{\theta^2}-\frac{1}{(\theta+2)^2}>0.$$

The ordinary score for one observation is $a(\theta)-Y$, and $a'(\theta)=-I(\theta)<0$. Since $a(\theta)$ decreases continuously from $+\infty$ to zero, there is a unique solution to the population likelihood equation $a(\theta)=\mathbb{E}_{\theta_0}(Y\mid Y>u)$. A law of large numbers and continuity then establish the asserted limit of the ordinary maximum likelihood estimator.

For the true law the survival function is

$$\overline{F}_{\mathrm{ZD}}(u;\theta)=e^{-\theta u}\left(1+\theta u+\frac{\theta^2u^2}{\theta+2}\right).$$

Uniformly for θ in any compact subset of $(0,\infty)$,

$$\log \overline{F}_{\mathrm{ZD}}(u;\theta)=-c(\theta)u^2+O(u^3),\qquad c(\theta)=\frac{\theta^3}{2(\theta+2)},$$

and differentiation gives

$$\partial_\theta \log \overline{F}_{\mathrm{ZD}}(u;\theta)=-c'(\theta)u^2+O(u^3),\qquad c'(\theta)=\frac{\theta^2(\theta+3)}{(\theta+2)^2}.$$

Differentiating the integral defining the survival probability yields

$$\partial_\theta \log \overline{F}_{\mathrm{ZD}}(u;\theta_0)=\mathbb{E}_{\theta_0}\left\{a(\theta_0)-Y\mid Y>u\right\}.$$

Consequently,

$$\mathbb{E}_{\theta_0}(Y\mid Y>u)=a(\theta_0)+c'(\theta_0)u^2+O(u^3).$$

A Taylor expansion of a at θ_0 now gives

$$-I(\theta_0)\{\theta^\dagger(u)-\theta_0\}=c'(\theta_0)u^2+O(u^3),$$

where the squared Taylor remainder is $O(u^4)$. Substituting

$$\frac{c'(\theta_0)}{I(\theta_0)}=\frac{\theta_0^4(\theta_0+3)}{2(\theta_0^2+6\theta_0+6)}$$

proves (76). Finally, the conditional likelihood uses the actual sampling density $f_{\mathrm{ZD}}(y;\theta)/\overline{F}_{\mathrm{ZD}}(u;\theta)$ on (u,∞). Its expected log-likelihood is uniquely maximised at the generating parameter by the Kullback–Leibler inequality and identifiability of the conditional family. Standard likelihood consistency then applies under the stated conditions. □

Remark 5.6 (Oracle count correction and its limitation)**.** *With deterministic interval thresholds u_i and directly observable jump marks, Poisson thinning gives the number K_n^{oracle} of detected* individual jumps*, with*

$$K_n^{\mathrm{oracle}}\sim \mathrm{Poisson}(\lambda E_n),\qquad E_n=\sum_i \Delta_i \overline{F}_\theta(u_i).$$

Thus, for known θ,

$$\mathbb{E}\left(\frac{K_n^{\text{oracle}}}{E_n}\right) = \lambda, \qquad \operatorname{Var}\left(\frac{K_n^{\text{oracle}}}{E_n}\right) = \frac{\lambda}{E_n}.$$

For the uncorrected estimator M_n/T_n,

$$\mathbb{E}\left(\frac{K_n^{\text{oracle}}}{T_n}\right) - \lambda = -\frac{\lambda}{T_n}\sum_i \Delta_i \Pr_\theta(Y \le u_i).$$

In particular the leading bias is of order r_n^2 for Zeghdoudi marks and usually of order r_n for positive-density-at-zero marks. These identities are exact for ideal thinning and deterministic thresholds. In the main text, however, M_n counts detected intervals, *not individual jumps, so $M_n = K_n^{\text{oracle}}$ need not hold when an interval contains more than one jump. A state-dependent threshold is predictable but random, and observed increments contain diffusion noise; the feasible estimator therefore requires a separate compensator/martingale argument and control of classification error. It would be incorrect to claim exact unbiasedness for that estimator on the basis of the oracle calculation alone.*

5.3 Consistency of mark estimation and model selection

Theorem 5.7 (Consistency of threshold-adjusted mark estimation)**.** *Suppose Assumption 5.1 holds,*

$$M_n \longrightarrow \infty \qquad \text{in probability,}$$

and the number of incorrectly classified observations is $o_p(M_n)$. Then, for the true mark family m_0,

$$\widehat{\eta}_{m_0} \xrightarrow{p} \eta_0. \tag{77}$$

Proof. We first consider an oracle sample consisting only of correctly detected single jumps. Conditional on the threshold $u_{i,n}$, the density of the corresponding mark is

$$y \mapsto \frac{f_{m_0}(y;\eta_0)}{\overline{F}_{m_0}(u_{i,n};\eta_0)}, \qquad y > u_{i,n}.$$

For a candidate parameter η, define its conditional contribution by

$$\ell_i^T(\eta) = \log f_{m_0}(Y_i;\eta) - \log \overline{F}_{m_0}(u_{i,n};\eta).$$

Because

$$r_n = \max_i u_{i,n} \to 0,$$

we have, uniformly in i,

$$\overline{F}_{m_0}(u_{i,n};\eta) \longrightarrow 1$$

for parameters in a compact set. Under the domination condition in Assumption 5.1(A4), the expected conditional log-likelihood therefore converges uniformly to

$$Q(\eta) = \mathbb{E}_{\eta_0}\{\log f_{m_0}(Y;\eta)\}.$$

By the Kullback–Leibler identity,

$$Q(\eta_0) - Q(\eta) = \mathbb{E}_{\eta_0}\left[\log \frac{f_{m_0}(Y;\eta_0)}{f_{m_0}(Y;\eta)}\right] \ge 0,$$

with equality only at $\eta = \eta_0$ by identifiability. Thus η_0 is the unique maximizer of the limiting criterion.

The uniform law of large numbers gives

$$\sup_{\eta} \left| \frac{1}{M_n} \ell^{T,\mathrm{oracle}}_{m_0,n}(\eta) - Q(\eta) \right| \xrightarrow{p} 0.$$

The argmax theorem therefore yields consistency of the oracle threshold-adjusted estimator.

It remains to replace the oracle marks by detected increments. On a correctly classified single-jump interval,

$$Z_i - Y_i = O_p(\sqrt{\Delta_n}) + O(\Delta_n).$$

The threshold condition implies

$$\frac{\sqrt{\Delta_n}}{\min_i u_{i,n}} \longrightarrow 0.$$

Consequently, the continuous perturbation is asymptotically small relative to the scale at which marks are retained. Under the regularity assumptions on the log-density and log-survival function, replacing Y_i by Z_i changes the normalized criterion by $o_p(1)$ uniformly in η.

Finally, only $o_p(M_n)$ observations are incorrectly classified. Their contribution to the normalized likelihood is therefore also $o_p(1)$. The feasible and oracle criteria have the same limiting objective, and the argmax theorem gives

$$\widehat{\eta}_{m_0} \xrightarrow{p} \eta_0.$$

□

Theorem 5.8 (BIC consistency over the finite candidate family)**.** *Under the conditions of Theorem 5.7 and Assumption 5.1(A5),*

$$\mathbb{P}(\widehat{m}_{\mathrm{BIC}} = m_0) \longrightarrow 1. \tag{78}$$

Proof. Let $m \neq m_0$ be an incorrect candidate. By the assumed positive Kullback–Leibler separation, there exists $K_m > 0$ such that

$$\frac{1}{M_n} \left\{ \ell^T_{m_0,n}(\widehat{\eta}_{m_0}) - \ell^T_{m,n}(\widehat{\eta}_m) \right\} \xrightarrow{p} K_m.$$

Hence

$$\ell^T_{m_0,n}(\widehat{\eta}_{m_0}) - \ell^T_{m,n}(\widehat{\eta}_m) = M_n K_m + o_p(M_n).$$

The BIC difference is

$$\begin{aligned} \mathrm{BIC}_m - \mathrm{BIC}_{m_0} &= 2 \left\{ \ell^T_{m_0,n}(\widehat{\eta}_{m_0}) - \ell^T_{m,n}(\widehat{\eta}_m) \right\} \\ &\quad + (k_m - k_{m_0}) \log M_n \\ &= 2M_n K_m + o_p(M_n) + O(\log M_n). \end{aligned}$$

Since

$$\frac{\log M_n}{M_n} \to 0,$$

the positive likelihood-separation term dominates the BIC penalty difference. Thus

$$\mathbb{P}(\mathrm{BIC}_m > \mathrm{BIC}_{m_0}) \longrightarrow 1.$$

The candidate set is finite, so the result holds simultaneously for all $m \neq m_0$, proving (78). □

Remark 5.9 (AICc and equal-dimensional candidates)**.** *The preceding theorem concerns BIC consistency. AICc is used in the empirical study because the number of detected marks is small.*

For candidates having the same number k of free parameters and fitted to the same M_n observations, the AIC, AICc, and BIC penalties are common across those candidates. Hence minimizing any of the three criteria gives exactly the same ordering as maximizing the threshold-adjusted log-likelihood. This is an ordering statement, not an additional model-selection consistency theorem.

5.4 Asymptotic normality for the Zeghdoudi mark parameter

Let

$$s_\theta(y) = \partial_\theta \log f_{ZD}(y;\theta) = \frac{3}{\theta} - \frac{1}{\theta+2} - y.$$

For one untruncated Zeghdoudi mark,

$$\mathcal{I}_{ZD}(\theta) = -\mathbb{E}_\theta\left[\frac{\partial^2}{\partial\theta^2}\log f_{ZD}(Y;\theta)\right] = \frac{3}{\theta^2} - \frac{1}{(\theta+2)^2}. \tag{79}$$

Clearly,

$$\mathcal{I}_{ZD}(\theta) > 0 \qquad (\theta > 0).$$

Theorem 5.10 (Asymptotic normality of the threshold-adjusted estimator)**.** *Suppose the true mark law is Zeghdoudi with parameter θ_0. Assume the conditions of Theorem 5.7, together with the smoothness conditions needed for differentiation under the integral sign and a Lindeberg condition for the triangular array of threshold-conditioned scores. Suppose also that the feasible score based on the detected increments is asymptotically equivalent, at the $\sqrt{M_n}$ scale, to the score based on correctly detected single-jump marks. Then*

$$\sqrt{M_n}(\widehat{\theta}_n - \theta_0) \xrightarrow{d} N\left(0, \mathcal{I}_{ZD}(\theta_0)^{-1}\right). \tag{80}$$

Proof. For an oracle detected mark with threshold $u_{i,n}$, the conditional score is

$$s^T_{i,n}(\theta) = s_\theta(Y_i) - \partial_\theta \log \overline{F}_{ZD}(u_{i,n};\theta). \tag{81}$$

Because this is the score of a correctly normalized conditional density,

$$\mathbb{E}_{\theta_0}\left[s^T_{i,n}(\theta_0) \mid Y_i > u_{i,n}\right] = 0.$$

Thus no additional truncation bias is introduced by the conditional likelihood.

As $r_n \to 0$,

$$\overline{F}_{ZD}(u_{i,n};\theta_0) \longrightarrow 1$$

uniformly in i. Hence the conditional distribution of the retained mark converges to the untruncated Zeghdoudi distribution. The corresponding conditional information therefore satisfies

$$\mathcal{I}^T_{i,n}(\theta_0) \longrightarrow \mathcal{I}_{ZD}(\theta_0)$$

uniformly in i.

The Zeghdoudi distribution has an exponential tail and consequently has moments of every positive order. Under the stated regularity condition, the Lindeberg condition for the triangular array of conditional scores holds, giving

$$\frac{1}{\sqrt{M_n}}\sum_{i\in\mathcal{J}_n} s^T_{i,n}(\theta_0) \xrightarrow{d} N\left(0, \mathcal{I}_{ZD}(\theta_0)\right). \tag{82}$$

Similarly, a uniform law of large numbers for the conditional Hessian gives

$$-\frac{1}{M_n}\partial^2_\theta \ell^T_{ZD,n}(\theta^*_n) \xrightarrow{p} \mathcal{I}_{ZD}(\theta_0) \tag{83}$$

for any sequence $\theta^*_n \xrightarrow{p} \theta_0$.

Since

$$\partial_\theta \ell^T_{ZD,n}(\widehat{\theta}_n) = 0,$$

a Taylor expansion about θ_0 yields

$$0 = \partial_\theta \ell^T_{ZD,n}(\theta_0) + \partial^2_\theta \ell^T_{ZD,n}(\theta^*_n)(\widehat{\theta}_n - \theta_0).$$

Therefore

$$\sqrt{M_n}(\widehat{\theta}_n - \theta_0) = \left[-\frac{1}{M_n}\partial^2_\theta \ell^T_{ZD,n}(\theta^*_n)\right]^{-1} \times \frac{1}{\sqrt{M_n}}\partial_\theta \ell^T_{ZD,n}(\theta_0).$$

Combining (82), (83), and the assumed feasible–oracle score equivalence gives (80) by Slutsky's theorem. □

The small-mark expansion gives an additional useful comparison. From Lemma 2.4,

$$\overline{F}_{ZD}(r;\theta) = 1 - \frac{\theta^3}{2(\theta+2)}r^2 + O(r^3). \tag{84}$$

Differentiating uniformly on compact subsets of $(0,\infty)$ gives

$$\partial_\theta \log \overline{F}_{ZD}(r;\theta) = O(r^2). \tag{85}$$

Remark 5.11 (Equivalence with the ordinary score)**.** *If one wishes to replace the conditional score by the ordinary untruncated Zeghdoudi score in a first-order expansion, a sufficient condition is*

$$\sqrt{M_n}\, r_n^2 \longrightarrow 0, \tag{86}$$

because

$$\sum_{i\in\mathcal{J}_n} \partial_\theta \log \overline{F}_{ZD}(u_{i,n};\theta_0) = O_p(M_n r_n^2).$$

After division by $\sqrt{M_n}$, this is negligible under (86).

For a mark density with $f_Y(0) > 0$, the corresponding survival correction is generally of order r_n, and the analogous sufficient condition becomes $\sqrt{M_n}\, r_n \to 0$. Condition (86) is therefore a useful consequence of the quadratic small-mark behavior of the Zeghdoudi law. It is a sufficient condition for equivalence with the ordinary score; it is not required simply to define or maximize the correctly specified threshold-conditioned likelihood.

5.5 Consistency of the corrected jump-intensity estimator

Define the true effective exposure

$$E_n := \sum_{i=1}^{n} \Delta_n \overline{F}_{m_0}(u_{i,n};\eta_0). \tag{87}$$

Theorem 5.12 (Consistency of the detection-corrected intensity estimator)**.** *Assume*

$$E_n \longrightarrow \infty,$$

the selected mark model is consistent,

$$\mathbb{P}(\widehat{m} = m_0) \longrightarrow 1,$$

and

$$\widehat{\eta}_{\widehat{m}} \xrightarrow{p} \eta_0.$$

If the contributions of false detections, masked jumps, and multiple-jump intervals are $o_p(E_n)$, then

$$\widehat{\lambda}_n \xrightarrow{p} \lambda. \tag{88}$$

Proof. Consider first the oracle count of correctly detected single-jump intervals. For interval i,

$$p_{i,n} = \overline{F}_{m_0}(u_{i,n}; \eta_0).$$

The probability of one jump whose mark exceeds the threshold is

$$\begin{aligned}\mathbb{P}(\Delta_i N = 1,\, Y > u_{i,n}) &= e^{-\lambda\Delta_n}\lambda\Delta_n p_{i,n} \\ &= \lambda\Delta_n p_{i,n} + O(\Delta_n^2).\end{aligned}$$

Therefore

$$\mathbb{E}(M_n^{\text{oracle}}) = \lambda E_n + O(T_n\Delta_n).$$

Under the stated contamination condition, this remainder is negligible at the scale E_n.

The interval indicators are independent for the compound-Poisson component, and their variances are bounded by their expectations. Hence

$$\operatorname{Var}(M_n^{\text{oracle}}) = O(E_n).$$

It follows from Chebyshev's inequality that

$$\frac{M_n^{\text{oracle}}}{E_n} \xrightarrow{p} \lambda.$$

The assumed contamination bound gives the same limit for the feasible count:

$$\frac{M_n}{E_n} \xrightarrow{p} \lambda. \tag{89}$$

Next consider the estimated denominator

$$\widehat{E}_n = \sum_{i=1}^{n} \Delta_n \overline{F}_{\widehat{m}}(u_{i,n}; \widehat{\eta}_{\widehat{m}}).$$

Model-selection and parameter consistency, together with continuity of the survival functions and the assumed uniform regularity, imply

$$\frac{\widehat{E}_n}{E_n} \xrightarrow{p} 1. \tag{90}$$

Combining (89) and (90),

$$\widehat{\lambda}_n = \frac{M_n}{\widehat{E}_n} = \frac{M_n}{E_n}\frac{E_n}{\widehat{E}_n} \xrightarrow{p} \lambda$$

by Slutsky's theorem. □

5.6 Consistency of the diffusion quasi-likelihood

Proposition 5.13 (Consistency of the diffusion QMLE)**.** *Assume Assumption 5.1(A1)–(A3), together with the inverse-moment and local discretization bounds required for the square-root diffusion quasi-likelihood. Suppose that the fraction of retained intervals still containing jumps tends to zero in probability. Then any maximizer of (70) satisfies*

$$(\widehat{\kappa}_n, \widehat{\mu}_n, \widehat{\sigma}_n) \xrightarrow{p} (\kappa, \mu, \sigma). \tag{91}$$

Proof. Write the true drift as

$$b_0(x) = \kappa_0(\mu_0 - x),$$

and the candidate drift as

$$b(x; \kappa, \mu) = \kappa(\mu - x).$$

On a correctly retained jump-free interval,

$$\Delta_i X = b_0(X_{t_{i-1}})\Delta_n + \sigma_0\sqrt{X_{t_{i-1}}}\Delta_i W + R_{i,n}, \tag{92}$$

where the accumulated contribution of $R_{i,n}$ is negligible under the standard high-frequency moment assumptions.

We first consider the volatility parameter. At the leading $\sqrt{\Delta_n}$ scale, the drift term is negligible and

$$\frac{(\Delta_i X)^2}{X_{t_{i-1}}\Delta_n}$$

has conditional mean close to σ_0^2. A martingale law of large numbers and ergodicity therefore imply that the volatility component of the normalized Gaussian contrast has its unique limit at

$$\sigma^2 = \sigma_0^2.$$

Since $\sigma > 0$,

$$\widehat{\sigma}_n \xrightarrow{p} \sigma_0.$$

For the drift parameters, substitute the consistent volatility estimator and expand the squared residual:

$$\begin{aligned}
&\{\Delta_i X - b(X_{t_{i-1}}; \kappa, \mu)\Delta_n\}^2 \\
&\quad = \{\Delta_i X - b_0(X_{t_{i-1}})\Delta_n\}^2 \\
&\qquad + \{b(X_{t_{i-1}}; \kappa, \mu) - b_0(X_{t_{i-1}})\}^2\Delta_n^2 \\
&\qquad - 2\{\Delta_i X - b_0(X_{t_{i-1}})\Delta_n\}\{b(X_{t_{i-1}}; \kappa, \mu) - b_0(X_{t_{i-1}})\}\Delta_n.
\end{aligned}$$

After normalization by T_n, the last term is a martingale term and vanishes in probability. Ergodicity converts the deterministic quadratic term into

$$\int \frac{\{b(x; \kappa, \mu) - b_0(x)\}^2}{\sigma_0^2 x}\,\pi(\mathrm{d}x), \tag{93}$$

where π is the stationary distribution.

The quantity in (93) is nonnegative and equals zero only when

$$b(x; \kappa, \mu) = b_0(x)$$

for π-almost every x. Since both sides are affine functions of x and the stationary law is nondegenerate, this implies

$$\kappa = \kappa_0, \qquad \kappa\mu = \kappa_0\mu_0.$$

Because $\kappa_0 > 0$,

$$\mu = \mu_0.$$

Finally, removing a vanishing fraction of contaminated intervals does not alter either limiting contrast. Uniform convergence on the compact parameter set and uniqueness of the limiting minimizer therefore yield (91). This is the usual long-span ergodic quasi-likelihood argument for discretely observed diffusions (Kessler, 1997), with threshold filtering contributing an asymptotically negligible deletion of increments. □

6 Prediction and rare-event quantities

The fitted model provides closed-form summaries for both the state process and the jump mechanism. These quantities are particularly useful because they preserve the effect of the selected mark distribution rather than reducing prediction to the fitted diffusion alone.

6.1 Conditional mean and variance

For a forecast horizon $h > 0$, Proposition 3.6 gives

$$\mathbb{E}(X_{t+h} \mid X_t) = \bar{m} + (X_t - \bar{m})e^{-\kappa h}, \tag{94}$$

where

$$\bar{m} = \mu + \frac{\lambda m_1}{\kappa}.$$

The corresponding plug-in predictor is

$$\widehat{X}_{t+h|t} = \widehat{\bar{m}} + (X_t - \widehat{\bar{m}})e^{-\widehat{\kappa} h}, \tag{95}$$

with

$$\widehat{\bar{m}} = \widehat{\mu} + \frac{\widehat{\lambda}\,\widehat{m}_1}{\widehat{\kappa}}. \tag{96}$$

The conditional variance also has a convenient form. Let

$$V_\infty = \frac{\sigma^2 \bar{m} + \lambda m_2}{2\kappa}.$$

Combining the first and second conditional moments gives

$$\mathrm{Var}(X_{t+h} \mid X_t) = V_\infty(1 - e^{-2\kappa h}) + \frac{\sigma^2}{\kappa}(X_t - \bar{m})(e^{-\kappa h} - e^{-2\kappa h}). \tag{97}$$

Thus a plug-in forecast variance is obtained by replacing $\kappa, \sigma, \bar{m}, V_\infty$ and the mark moments by their fitted values.

For Zeghdoudi marks,

$$m_1(\theta) = \frac{2(\theta+3)}{\theta(\theta+2)}, \qquad m_2(\theta) = \frac{6(\theta+4)}{\theta^2(\theta+2)},$$

so both the predictive mean level and variance remain explicit.

6.2 Future large-jump probabilities

Let $y > 0$ be a mark-size threshold. The jumps whose marks exceed y form a thinned Poisson process with rate

$$\lambda \overline{F}_Y(y).$$

Consequently, the number of such jumps in $(t, t+h]$ is

$$N_{>y}(t, t+h] \sim \text{Poisson}\left(\lambda h \overline{F}_Y(y)\right).$$

It follows that

$$\mathbb{P}\left(\max_{t<s\le t+h} \Delta J_s > y\right) = 1 - \exp\left\{-\lambda h \overline{F}_Y(y)\right\}. \tag{98}$$

The expected number of marks exceeding y over the same horizon is

$$\mathbb{E}\{N_{>y}(t, t+h]\} = \lambda h \overline{F}_Y(y). \tag{99}$$

For Zeghdoudi marks,

$$\overline{F}_{ZD}(y;\theta) = e^{-\theta y}\left(1+\theta y+\frac{\theta^2 y^2}{\theta+2}\right),$$

and therefore

$$\mathbb{P}\left(\max_{t<s\le t+h}\Delta J_s > y\right) = 1-\exp\left[-\lambda h e^{-\theta y}\left(1+\theta y+\frac{\theta^2 y^2}{\theta+2}\right)\right]. \tag{100}$$

This quantity refers specifically to the occurrence of a large jump mark. It should not be confused with the probability that the state variable X_{t+h} itself exceeds a prescribed level, since the latter also depends on the diffusion evolution and on the timing and accumulation of all jumps.

6.3 Predictive use of the conditional PIT

The conditional PIT values in (55) provide a direct diagnostic for the fitted mark component. If the mark law and threshold correction are adequate, these transformed exceedances should show no systematic departure from uniformity.

This diagnostic is preferable to applying an ordinary goodness-of-fit test to the detected marks because the latter ignores the event through which the sample was selected. With very few detected jumps, however, graphical PIT assessment and descriptive statistics should be interpreted cautiously. A parametric bootstrap that repeats threshold detection, estimation, and PIT calculation would provide a more reliable finite-sample calibration.

7 Monte Carlo study

7.1 Simulation design

The simulation examines the finite-sample behavior of the proposed estimators when the data are generated from the JCIR–Zeghdoudi model. The purpose of this experiment is to study recovery of the diffusion, jump-intensity, and mark parameters under the principal working specification developed in the previous sections. It should therefore be viewed primarily as an estimator performance study rather than as an independent validation of the empirical model ranking.

The data-generating parameters are

$$\kappa = 1.2, \qquad \mu = 0.8, \qquad \sigma = 0.25, \qquad \lambda = 0.8, \qquad \theta = 6. \tag{101}$$

The observation interval is fixed at

$$\Delta = 0.02,$$

and three sample sizes are considered:

$$n \in \{500, 1000, 2000\}.$$

Hence the corresponding observation horizons are

$$T_n = n\Delta \in \{10, 20, 40\}.$$

Increasing n therefore lengthens the observation horizon while preserving the same sampling interval. This design is useful for examining the long-span behavior emphasized in the asymptotic theory.

The Feller condition is satisfied since

$$2\kappa\mu = 2(1.2)(0.8) = 1.92 > \sigma^2 = 0.0625.$$

Thus the diffusion component remains well inside the strictly positive regime.

Sample paths are generated using a full-truncation Euler scheme with three internal substeps within each observation interval. At every internal substep, the number of jumps is generated from the corresponding Poisson distribution, and the jump magnitudes are simulated using the Zeghdoudi gamma-mixture representation

$$Y \sim \frac{\theta}{\theta+2}\text{Gamma}(2,\theta) + \frac{2}{\theta+2}\text{Gamma}(3,\theta).$$

The internal subdivision is used only to improve path generation; inference is carried out from the observations recorded at spacing Δ.

For each simulated trajectory, candidate jump intervals are identified by the adaptive rule (47), with threshold multiplier

$$c = 1.4.$$

The retained jump marks are fitted using the threshold-conditioned likelihood (49). The jump intensity is estimated with the detection-corrected exposure estimator (68), and the diffusion parameters are obtained from the jump-filtered Gaussian quasi-likelihood (70).

A total of $R = 200$ Monte Carlo replications is used for each sample size. The implementation also fits all candidate mark distributions to each simulated data set and records the corresponding information criteria. The results reported below focus first on parameter recovery under the true Zeghdoudi specification.

7.2 Performance measures

Let ξ_0 denote the true value of a generic parameter and let $\widehat{\xi}^{(1)}, \ldots, \widehat{\xi}^{(R_v)}$ denote the estimates from the R_v valid replications. We report the Monte Carlo mean

$$\overline{\widehat{\xi}} = \frac{1}{R_v}\sum_{r=1}^{R_v}\widehat{\xi}^{(r)}, \tag{102}$$

the empirical bias

$$\text{Bias}(\widehat{\xi}) = \overline{\widehat{\xi}} - \xi_0, \tag{103}$$

the root mean squared error

$$\text{RMSE}(\widehat{\xi}) = \left[\frac{1}{R_v}\sum_{r=1}^{R_v}\left(\widehat{\xi}^{(r)} - \xi_0\right)^2\right]^{1/2}, \tag{104}$$

and the Monte Carlo standard deviation

$$\text{SD}(\widehat{\xi}) = \left[\frac{1}{R_v - 1}\sum_{r=1}^{R_v}\left(\widehat{\xi}^{(r)} - \overline{\widehat{\xi}}\right)^2\right]^{1/2}. \tag{105}$$

A replication is regarded as valid for the reported joint summary when the threshold procedure detects enough marks for stable mark estimation. At $n = 500$, 192 of the 200 replications satisfy this requirement; all replications are retained for $n = 1000$ and $n = 2000$.

7.3 Parameter recovery

Table 2 summarizes the Monte Carlo results.

Table 2: Monte Carlo performance of the estimators under the JCIR–Zeghdoudi model. The true parameter vector is $(\kappa, \mu, \sigma, \lambda, \theta) = (1.2, 0.8, 0.25, 0.8, 6)$.

n	Parameter	Valid	Mean	Bias	RMSE	SD
500	κ	192	1.3698	0.1698	0.4248	0.3904
500	μ	192	0.8197	0.0197	0.0747	0.0722
500	σ	192	0.2489	-0.0011	0.0080	0.0079
500	λ	192	0.8771	0.0771	0.3340	0.3259
500	θ	192	6.5818	0.5818	2.4003	2.3348
1000	κ	200	1.2606	0.0606	0.2428	0.2357
1000	μ	200	0.8147	0.0147	0.0618	0.0601
1000	σ	200	0.2493	-0.0007	0.0060	0.0060
1000	λ	200	0.8102	0.0102	0.2386	0.2390
1000	θ	200	6.3401	0.3401	1.3682	1.3286
2000	κ	200	1.1984	-0.0016	0.1293	0.1296
2000	μ	200	0.8117	0.0117	0.0408	0.0392
2000	σ	200	0.2504	0.0004	0.0046	0.0046
2000	λ	200	0.8007	0.0007	0.1530	0.1534
2000	θ	200	6.1585	0.1585	0.8617	0.8492

Several features of Table 2 are worth noting. The volatility parameter σ is estimated accurately throughout. Even for $n = 500$, its empirical bias is only -0.0011 and its RMSE is 0.0080. At $n = 2000$, the mean estimate is 0.2504, essentially equal to the true value 0.25. This behavior is consistent with the fact that σ is identified mainly through the high-frequency scale of the continuous increments.

The mean-reversion level μ is also estimated with relatively small bias. Its Monte Carlo mean decreases from 0.8197 at $n = 500$ to 0.8117 at $n = 2000$, while the RMSE falls from 0.0747 to 0.0408.

The mean-reversion speed κ is more difficult to estimate over shorter horizons. At $n = 500$, the mean estimate is 1.3698, giving a positive bias of 0.1698. Both bias and variability decline as the observation horizon increases. By $n = 2000$, the mean estimate is 1.1984 and the bias is -0.0016, while the RMSE has fallen to 0.1293. This pattern is expected for mean-reverting diffusions because drift parameters are learned primarily from the long-span behavior of the process rather than from the local variation alone.

The jump-intensity estimator shows a similar improvement. At the shortest horizon,

$$\overline{\widehat{\lambda}} = 0.8771,$$

compared with the true value 0.8. The bias falls to 0.0102 for $n = 1000$ and to 0.0007 for $n = 2000$. The RMSE decreases from 0.3340 to 0.1530. These results suggest that the detection-probability correction is increasingly effective once a sufficient number of jump opportunities is observed.

The Zeghdoudi rate parameter θ is the most variable quantity in the smallest sample. This is not surprising: unlike the diffusion parameters, it is estimated only from increments classified as jumps. At $n = 500$, its mean estimate is 6.5818 and the RMSE is 2.4003. With the longer observation horizons, the effective mark sample becomes larger and the dispersion decreases markedly. At $n = 2000$,

$$\overline{\widehat{\theta}} = 6.1585, \qquad \mathrm{RMSE}(\widehat{\theta}) = 0.8617.$$

The remaining positive bias is small relative to the uncertainty observed at the shortest horizon.

Overall, the simulation shows the distinction between parameters learned from different parts of the trajectory. The volatility parameter is recovered accurately from local increments, whereas κ, λ, and especially θ require a longer observation horizon. The monotone reduction in RMSE across the three sample sizes is consistent with the large-sample arguments developed in Section 5.

7.4 Mark-law selection in the simulations

The same Monte Carlo implementation also fits the complete candidate family to the detected marks in every valid replication and records the selected model under AICc and BIC. These frequencies provide a different diagnostic from Table 2. Parameter recovery asks whether the estimators approach the correct numerical values when the Zeghdoudi model is true; selection frequency asks how often the finite-family comparison identifies that mark law from the threshold-selected sample.

This distinction is important because several of the one-parameter candidates have very similar shapes after conditioning above the detection threshold. A low selection frequency at short horizons would therefore not necessarily indicate poor parameter estimation. It may instead reflect the limited information available for distinguishing closely related mark laws.

For this reason, model-selection frequencies should be reported together with the corresponding information-criterion gaps rather than interpreted as a binary success-or-failure measure. In particular, when only a small number of jumps are detected, a narrow AICc difference indicates substantial competition among the candidate models.

The present parameter-recovery table does not by itself establish the ability of the procedure to distinguish Zeghdoudi marks from alternative families. A stronger assessment of the selection step is obtained by repeating the simulation under several competing data-generating mark laws, for example Zeghdoudi, Q-Lindley, XLindley, and QGamma, and recording

$$\mathbb{P}(\widehat{m} = m_0)$$

for each sample size. Such a multi-DGP experiment separates estimator accuracy from finite-sample model discrimination and provides a direct numerical counterpart to the finite-candidate selection theory.

8 Application to annual Nile flow

8.1 Data and role of the application

The empirical illustration uses the `Nile` data set distributed with base R. It contains 100 annual measurements of Nile flow at Aswan from 1871 to 1970, expressed in units of $10^8\,\mathrm{m}^3$. The series is a classical change-point example, with a well-known level shift around 1898 (Cobb, 1978; Durbin and Koopman, 2001). For numerical conditioning, the observations are divided by 1000, and the sampling interval is taken as $\Delta = 1$ year.

The Nile data are not used here as a literal realization of the high-frequency stationary model considered in Section 5. Rather, the purpose is to illustrate the complete inferential workflow on a well-known positive time series: threshold detection, conditional mark fitting, comparison of candidate mark laws, correction of the latent jump intensity, and propagation of the selected mark distribution into the fitted jump–diffusion.

For the reference analysis we take the threshold multiplier

$$c = 0.7.$$

The rule (47) identifies ten positive increments as jump candidates. They end in the years

$$1874,\ 1878,\ 1903,\ 1906,\ 1908,\ 1914,\ 1916,\ 1929,\ 1946,\ 1964,$$

with scaled increments

$$0.247,\ 0.417,\ 0.246,\ 0.215,\ 0.328,\ 0.368,\ 0.418,\ 0.244,\ 0.239,\ 0.269.$$

Because these observations enter the mark sample only after crossing their interval-specific thresholds, all model comparisons use the threshold-conditioned likelihood (49). An ordinary untruncated likelihood would ignore the mechanism by which the sample was formed.

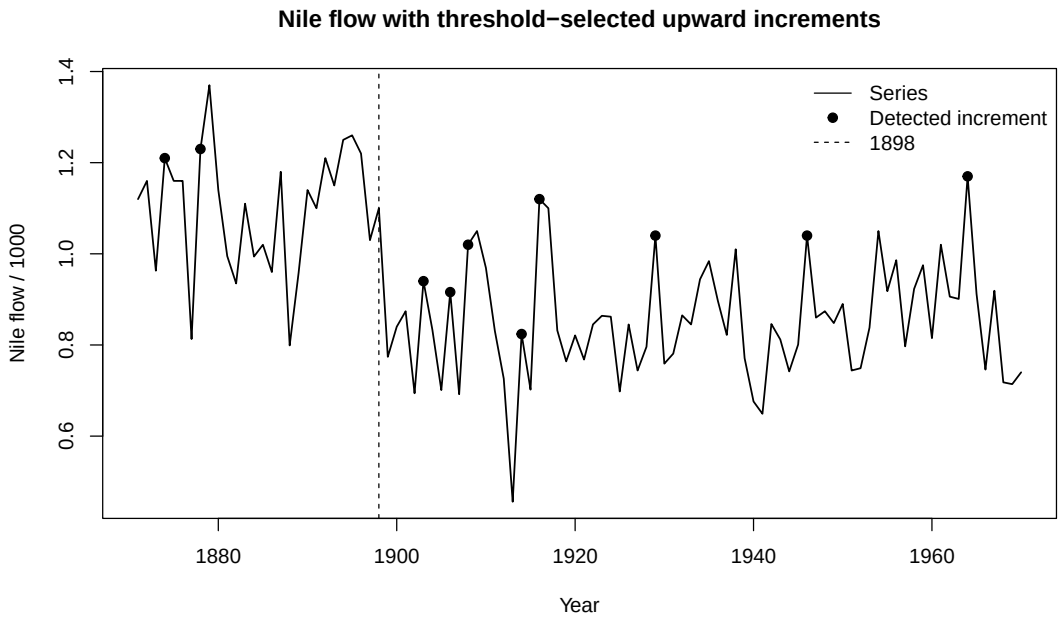


Figure 1: Annual Nile flow, scaled by 1000, with the positive increments selected by the threshold rule at $c = 0.7$. The vertical dashed line marks the well-known level shift around 1898. The figure is therefore interpreted as an illustration of the estimation procedure rather than as evidence that the entire series follows a stationary jump–diffusion.

8.2 Comparison of candidate mark laws

Table 3 reports the threshold-adjusted maximum likelihood fits at the reference multiplier $c = 0.7$. Gamma and Weibull contain two free mark parameters, whereas the remaining candidates contain one. Since only ten increments are retained, AICc is used for the primary finite-sample comparison; AIC and BIC are reported alongside it. The conditional PIT statistics are used descriptively because the sample is too small for their usual asymptotic calibration to be especially informative.

Table 3: Threshold-adjusted comparison of candidate mark laws for the Nile illustration at $c = 0.7$, ordered by AICc. The Akaike weights are normalized over the eleven reported candidates.

Model	$\widehat{\eta}_1$	$\widehat{\eta}_2$	ℓ_T	AIC	AICc	BIC	ΔAICc	w_{AICc}	KS_{PIT}	AD_{PIT}
Zeghdoudi	15.0700	–	14.4841	-26.9682	-26.4682	-26.6656	0.0000	0.1110	0.2361	0.6662
Q-Lindley	14.2985	–	14.4777	-26.9554	-26.4554	-26.6528	0.0128	0.1103	0.2348	0.6629
New XLindley	13.8345	–	14.4593	-26.9187	-26.4187	-26.6161	0.0495	0.1083	0.2304	0.6491
X-exponential	13.4707	–	14.4461	-26.8923	-26.3923	-26.5897	0.0759	0.1069	0.2276	0.6391
Lindley	12.2082	–	14.4134	-26.8268	-26.3268	-26.5242	0.1414	0.1034	0.2222	0.6186
Xgamma	14.0651	–	14.4107	-26.8214	-26.3214	-26.5188	0.1468	0.1031	0.2196	0.6120
ZLindley	11.8872	–	14.4091	-26.8183	-26.3183	-26.5157	0.1499	0.1030	0.2215	0.6165
XLindley	11.5528	–	14.4072	-26.8143	-26.3143	-26.5118	0.1539	0.1028	0.2212	0.6156
QGamma	12.3968	–	14.3894	-26.7789	-26.2789	-26.4763	0.1893	0.1010	0.2183	0.6061
Weibull	2.0090	0.2233	14.6324	-25.2648	-23.5505	-24.6596	2.9177	0.0258	0.2863	0.8655
Gamma	5.6459	25.1187	14.5832	-25.1663	-23.4520	-24.5611	3.0162	0.0246	0.2781	0.8219

At the reference threshold, Zeghdoudi has the smallest AIC, AICc, and BIC. The numerical advantage over the other one-parameter models is nevertheless very small. Q-Lindley is only

$$0.0128$$

AICc units behind, and all nine one-parameter candidates satisfy

$$\Delta\text{AICc} < 0.20.$$

The Akaike weights reflect the same lack of sharp separation:

$$w_{ZD} = 0.1110, \qquad w_{QL} = 0.1103,$$

while the remaining one-parameter weights are also close to 0.10.

The appropriate interpretation is therefore not that the reference sample strongly identifies a unique mark distribution. Rather, several one-parameter laws describe the threshold-selected increments almost equally well. Zeghdoudi provides the smallest information criterion at the chosen cutoff, but the uncertainty among the leading candidates remains substantial.

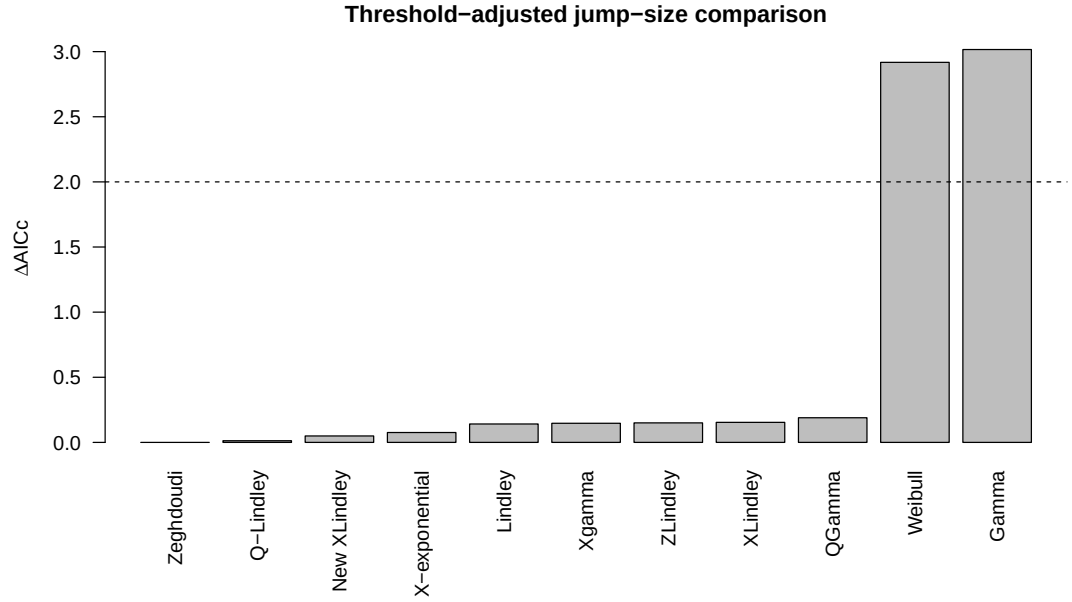


Figure 2: Threshold-adjusted ΔAICc values at the reference multiplier $c = 0.7$. The dashed line at $\Delta\text{AICc} = 2$ is shown as a conventional descriptive reference, not as a formal decision boundary.

The comparison also illustrates why conditioning on the threshold matters. All detected marks are bounded away from zero by construction. Fitting an ordinary density to this selected sample would ignore the probability mass removed below each threshold and could change the relative likelihoods of models that behave differently near the origin. The conditional likelihood instead compares the candidate distributions on the event that actually generated the observed mark sample.

8.3 Sensitivity to the detection threshold

A single threshold choice should not determine the substantive conclusion. We therefore repeat the complete mark-law comparison over

$$c \in \{0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}.$$

The results are reported in Table 4.

Table 4: Sensitivity of the mark-law comparison to the threshold multiplier. The table reports the Zeghdoudi estimate, its AICc, its difference from XLindley, the detection-corrected intensity, and the minimum-AICc model at each cutoff. The comparisons use the same Nile series and should therefore be read as a stability diagnostic rather than as independent replications.

c	Detected	$\widehat{\theta}_{ZD}$	AICc_{ZD}	Δ_{XL}	$\widehat{\lambda}_{ZD}$	Minimum-AICc model
0.5	20	15.6753	-55.4803	0.2063	0.6882	Zeghdoudi
0.6	14	15.0496	-37.6177	0.2936	0.6289	Zeghdoudi
0.7	10	15.0699	-26.4682	0.1538	0.6420	Zeghdoudi
0.8	6	12.6501	-12.2098	0.2569	0.3405	Zeghdoudi
0.9	4	10.7948	-4.8828	0.3440	0.2024	Zeghdoudi
1.0	4	13.1417	-6.7898	0.2123	0.4736	Zeghdoudi

Zeghdoudi attains the minimum AICc at every reported cutoff. This repeated ranking is useful because it shows that the reference choice $c = 0.7$ is not responsible for the identity of the leading model. It should not, however, be interpreted as six independent confirmations: the fitted samples are nested and come from the same time series.

The parameter estimates are most stable over the lower part of the grid. For example,

$$\widehat{\theta}_{ZD} = 15.6753,\ 15.0496,\ 15.0699$$

for $c = 0.5, 0.6, 0.7$, respectively. Once the threshold is raised to $c \geq 0.8$, only six or four increments remain and both $\widehat{\theta}_{ZD}$ and the corrected intensity become more variable. This loss of stability is expected from the small effective mark sample and supports retaining $c = 0.7$ for the reference fit.

Taken together, Tables 3 and 4 motivate Zeghdoudi as a stable working mark law for the fitted dynamics. The choice is based on the persistence of the minimum-AICc ranking under threshold perturbation, not on a large AICc advantage at any single cutoff.

8.4 Fitted JCIR–Zeghdoudi dynamics

Using the reference multiplier $c = 0.7$, the fitted dynamic parameters are approximately

$$\widehat{\kappa} = 0.3544, \qquad \widehat{\mu} = 0.8333, \qquad \widehat{\sigma} = 0.1301, \qquad \widehat{\theta} = 15.0699. \tag{106}$$

The ten detected increments correspond to the naive observed rate

$$\frac{10}{99} \approx 0.1010 \qquad \text{per year.}$$

This rate counts only increments that cross the threshold. After correcting for the fitted interval-specific detection probabilities using (68), the estimated latent compound-Poisson intensity is

$$\widehat{\lambda} \approx 0.6420 \qquad \text{per year.} \tag{107}$$

The difference between 0.1010 and 0.6420 is substantial. Under the fitted model, it arises because many latent Zeghdoudi marks are expected to lie below the annual detection boundaries. Consequently, $\widehat{\lambda}$ should not be interpreted as an observed frequency of large annual increases. It is the estimated intensity of the latent jump component implied by the fitted threshold model.

For

$$\widehat{\theta} = 15.0699,$$

the fitted first two mark moments are

$$\widehat{m}_1 \approx 0.1405, \qquad \widehat{m}_2 \approx 0.02952.$$

The fitted long-run mean is therefore

$$\widehat{\bar{m}} = \widehat{\mu} + \frac{\widehat{\lambda}\widehat{m}_1}{\widehat{\kappa}} \approx 1.088. \tag{108}$$

Using (36),

$$\widehat{\mathrm{Var}}(X_\infty) = \frac{\widehat{\sigma}^2\widehat{\bar{m}} + \widehat{\lambda}\widehat{m}_2}{2\widehat{\kappa}}, \tag{109}$$

which gives a stationary standard deviation of approximately

$$0.230$$

on the scaled flow unit.

The fitted mean-reversion coefficient also implies the model-based autocorrelation function

$$\widehat{\rho}(h) = e^{-0.3544h}. \tag{110}$$

This quantity is reported as a feature of the fitted JCIR model and should not be confused with evidence that the full Nile series is stationary.

8.5 Structural change and interpretation

The structural break near 1898 is central to the interpretation of this example. A persistent change in level is fundamentally different from an isolated jump. In a richer dynamic model, a sustained shift would normally be represented through the drift, a change-point component, or a latent regime rather than through a sequence of independent compound-Poisson marks.

This distinction matters because absorbing structural change into the jump component can inflate the fitted jump intensity and alter long-horizon predictive quantities. The present application therefore has a deliberately limited role. It demonstrates how the proposed threshold likelihood, mark-law comparison, and intensity correction operate on observed data; it does not claim that the annual Nile record satisfies the stationary high-frequency assumptions used in Section 5.

For more densely observed applications, the same inferential structure can be combined with a time-varying mean level, explicit change points, regime-switching dynamics, or covariate-dependent jump intensity. Such extensions would allow persistent changes and isolated shocks to be modeled separately.

9 Discussion

The main statistical issue addressed in this paper is not the construction of another positive distribution. The mark law enters the stochastic dynamics through the jump measure and therefore influences the conditional transform, stationary moments, latent intensity correction, and rare-event probabilities. Treating it as an estimable component of the dynamic model is therefore more informative than fixing it before the jump–diffusion is fitted.

A second point concerns the way the mark sample is obtained. Once large increments have been selected by a threshold, those observations are no longer an ordinary sample from the original mark distribution. The conditional likelihood in (49) accounts for this selection explicitly, while (68) uses the same detection probabilities to estimate the latent Poisson exposure. These two corrections are naturally linked: one concerns the distribution of the observed marks, the other the number of shocks that remain unobserved.

The gamma-mixture representation makes this framework particularly tractable. For most of the one-parameter candidates, moments, Laplace transforms, survival probabilities, and finite-jump convolutions remain explicit. This allows the mark distributions to be compared inside the same stochastic model rather than as unrelated marginal fits.

The Zeghdoudi law has an additional feature that matters for threshold inference. Since its density vanishes linearly at the origin,

$$\mathbb{P}(Y \leq r) = O(r^2),$$

whereas a mark law with positive density at zero generally gives an $O(r)$ small-mark probability. This difference propagates directly to the rate of missed sub-threshold jumps and to the first-order comparison between the conditional and untruncated score functions. The role of Zeghdoudi in the paper is therefore both empirical and inferential.

The Nile application also illustrates an important limitation. At the reference threshold the information-criterion differences among the one-parameter candidates are extremely small. The data do not support a claim of sharp separation. What they do show is that Zeghdoudi remains the minimum-AICc specification throughout the reported threshold grid. This supports its use as a stable working model while leaving substantial uncertainty about the precise one-parameter mark family.

Several extensions are natural. A constant jump intensity could be replaced by

$$\lambda_t = \exp(\beta_0 + \beta^\top Z_t),$$

allowing observed covariates to affect jump occurrence. Persistent structural changes could be handled through change-point or regime-switching drift models. Model uncertainty could be propagated through model averaging instead of selecting a single candidate. For genuinely high-frequency data, bootstrap or resampling procedures could be used to calibrate the threshold, PIT diagnostics, and finite-sample uncertainty of the corrected intensity. Finally, the finite candidate set could be replaced by a semiparametric gamma-mixture sieve when no named mark law provides an adequate description of the detected jumps.

A further numerical extension would be useful for evaluating the model-selection step itself. The present Monte Carlo study examines parameter recovery under a JCIR–Zeghdoudi data-generating process. Repeating the experiment under several competing mark laws would allow the empirical selection probabilities

$$\mathbb{P}(\widehat{m} = m_0)$$

to be studied directly and would complement the BIC consistency result of Section 5.

10 Conclusion

This paper develops a threshold-based framework for positive mean-reverting jump–diffusions in which the jump-size distribution is treated as part of the inferential problem. Most of the one-parameter candidates considered here share a finite common-rate gamma-mixture structure, making their moments, Laplace transforms, survival functions, and finite-jump convolutions available in closed form.

The main inferential step is to account for the fact that detected marks have already crossed interval-specific thresholds. The resulting conditional likelihood avoids fitting an unconditional density to a truncated sample, while the same survival probabilities provide an effective-exposure correction for the latent jump intensity. Within the JCIR model we establish positivity, affine transform formulas, conditional and stationary moments, stationary dependence, generator identifiability, consistency of the mark, intensity, and diffusion estimators, and BIC consistency over a finite candidate family.

For Zeghdoudi marks, the ordinary likelihood has a unique closed-form maximizer, and the threshold-conditioned estimator admits a first-order normal limit under the stated regularity conditions. More importantly for the threshold procedure, the density vanishes at the origin, leading to

$$\mathbb{P}(Y \leq r) = O(r^2)$$

rather than the first-order behavior associated with positive density at zero. This local property has a direct effect on the frequency of latent jumps that remain below the detection boundary.

The Nile analysis should be read with the same caution as the theoretical claims. Zeghdoudi has the smallest AICc at the reference threshold and remains the minimum-AICc model for every multiplier between 0.5 and 1.0, but the information-criterion differences among the leading one-parameter models are very small. The evidence therefore supports Zeghdoudi as a stable working specification, not as an empirically dominant alternative.

The broader implication is that jump detection, mark-law estimation, intensity correction, and stochastic-dynamic inference are closely connected. Treating them within a common framework makes the effect of the jump-size law visible both statistically and dynamically, while also making model uncertainty easier to assess.

Data and Code Availability

The empirical analysis uses the publicly available `Nile` dataset included in base R. The R code used for parameter estimation, Monte Carlo simulations, model comparison, and prediction is available from the corresponding author upon reasonable request.

Ethics Approval

This study is based on mathematical modelling, numerical simulations, and publicly available data. It does not involve human participants, personal data, or animal experiments. Therefore, approval from an Institutional Review Board or Research Ethics Committee was not required.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interests

The author declares that there are no competing financial or non-financial interests.

A Additional derivations

This appendix collects calculations that support the formulas used in the main text but are not needed for the flow of the statistical development.

A.1 Survival functions from the gamma-mixture representation

Let

$$G_r(y;\theta)$$

denote the distribution function of a gamma random variable with integer shape $r \geq 1$ and rate $\theta > 0$. Its survival function is

$$\overline{G}_r(y;\theta) = e^{-\theta y} \sum_{j=0}^{r-1} \frac{(\theta y)^j}{j!}, \qquad y \geq 0. \tag{111}$$

If the mark density has the common-rate mixture representation

$$f_m(y;\theta) = \sum_{r \in \mathcal{R}_m} w_{m,r}(\theta) g_r(y;\theta),$$

then integration component by component gives

$$\overline{F}_m(y;\theta) = \sum_{r \in \mathcal{R}_m} w_{m,r}(\theta) e^{-\theta y} \sum_{j=0}^{r-1} \frac{(\theta y)^j}{j!}. \tag{112}$$

Thus every one-parameter candidate in Table 1 has an explicit survival function. This is useful for the threshold-conditioned likelihood and the detection-exposure correction because no numerical integration is required.

For the Zeghdoudi distribution,

$$Y \sim \frac{\theta}{\theta+2}\text{Gamma}(2,\theta) + \frac{2}{\theta+2}\text{Gamma}(3,\theta).$$

Using (111),

$$\begin{aligned}
\overline{F}_{ZD}(y;\theta) &= \frac{\theta}{\theta+2} e^{-\theta y}(1+\theta y) \\
&\quad + \frac{2}{\theta+2} e^{-\theta y}\left(1 + \theta y + \frac{\theta^2 y^2}{2}\right) \\
&= e^{-\theta y}\left[\frac{\theta(1+\theta y) + 2(1+\theta y) + \theta^2 y^2}{\theta+2}\right] \\
&= e^{-\theta y}\left[1 + \theta y + \frac{\theta^2 y^2}{\theta+2}\right].
\end{aligned} \tag{113}$$

This agrees with (22). As a direct check,

$$f_{ZD}(y;\theta) = -\frac{\partial}{\partial y}\overline{F}_{ZD}(y;\theta).$$

Indeed,

$$\begin{aligned}
-\frac{\partial}{\partial y}\overline{F}_{ZD}(y;\theta) &= e^{-\theta y}\left[\theta\left(1 + \theta y + \frac{\theta^2 y^2}{\theta+2}\right) - \left(\theta + \frac{2\theta^2 y}{\theta+2}\right)\right] \\
&= \frac{\theta^3}{\theta+2} y(1+y) e^{-\theta y},
\end{aligned}$$

which recovers (11).

A.2 Short-time transition approximation

For a short observation interval of length Δ, condition on

$$X_t = x.$$

Write

$$b(x) = \kappa(\mu - x).$$

Over $(t, t+\Delta]$, the increment admits the local representation

$$X_{t+\Delta} - X_t = b(x)\Delta + \sigma\sqrt{x}\,(W_{t+\Delta} - W_t) + \sum_{j=1}^{K_\Delta} Y_j + R_\Delta, \tag{114}$$

where

$$K_\Delta \sim \text{Poisson}(\lambda\Delta)$$

and, under the usual local regularity conditions,

$$R_\Delta = o_p(\sqrt{\Delta}).$$

Since

$$W_{t+\Delta} - W_t \overset{d}{=} \sqrt{\Delta}\,\varepsilon, \qquad \varepsilon \sim N(0,1),$$

the leading approximation is

$$\Delta X \approx b(x)\Delta + \sigma\sqrt{x\Delta}\,\varepsilon + S_{K_\Delta}, \tag{115}$$

where

$$S_k = \sum_{j=1}^{k} Y_j, \qquad S_0 = 0.$$

Let

$$\varphi_\Delta(z; x) = \frac{1}{\sqrt{2\pi\sigma^2 x\Delta}} \exp\left[-\frac{\{z - b(x)\Delta\}^2}{2\sigma^2 x\Delta}\right]$$

denote the local Gaussian density of the diffusion increment.

For $k \geq 1$, let h_k denote the density of the k-jump mark sum S_k. By Proposition 2.3, h_k is a finite mixture of gamma densities. Conditional on $K_\Delta = k$, the approximate increment density is therefore

$$p_k(z \mid x) = \int_0^\infty \varphi_\Delta(z - s; x)\, h_k(s)\, ds, \qquad k \geq 1, \tag{116}$$

whereas

$$p_0(z \mid x) = \varphi_\Delta(z; x). \tag{117}$$

The full short-time mixture is

$$p_\Delta(z \mid x) \approx \sum_{k=0}^{\infty} e^{-\lambda\Delta} \frac{(\lambda\Delta)^k}{k!} p_k(z \mid x). \tag{118}$$

Retaining at most two jumps gives

$$\begin{aligned} p_\Delta(z \mid x) \approx\ & e^{-\lambda\Delta} p_0(z \mid x) + e^{-\lambda\Delta}(\lambda\Delta) p_1(z \mid x) \\ & + e^{-\lambda\Delta} \frac{(\lambda\Delta)^2}{2} p_2(z \mid x). \end{aligned} \tag{119}$$

The omitted Poisson probability is

$$\mathbb{P}(K_\Delta \geq 3) = O(\Delta^3).$$

Thus, apart from the local diffusion approximation, truncation at two jumps introduces an error of Poisson order $O(\Delta^3)$.

The gamma-mixture structure is especially convenient here. If

$$\mathbb{P}(R_j = r) = w_{m,r}(\theta), \qquad r \in \mathcal{R}_m,$$

then

$$h_k(s) = \sum_{(r_1,\dots,r_k)\in\mathcal{R}_m^k} \left\{ \prod_{j=1}^{k} w_{m,r_j}(\theta) \right\} g_{r_1+\cdots+r_k}(s;\theta). \tag{120}$$

Hence the mark part of each p_k is available analytically as a finite gamma mixture. The remaining operation in (116) is a one-dimensional Gaussian–gamma convolution. This provides a practical route to an approximate transition likelihood when one prefers not to classify increments by a hard threshold.

For Zeghdoudi marks, for example,

$$\mathcal{R}_{ZD} = \{2, 3\}, \qquad w_2 = \frac{\theta}{\theta+2}, \qquad w_3 = \frac{2}{\theta+2}.$$

Thus for two jumps,

$$\begin{aligned} h_2(s) &= w_2^2 g_4(s;\theta) + 2w_2w_3 g_5(s;\theta) + w_3^2 g_6(s;\theta) \\ &= \frac{\theta^2}{(\theta+2)^2} g_4(s;\theta) + \frac{4\theta}{(\theta+2)^2} g_5(s;\theta) + \frac{4}{(\theta+2)^2} g_6(s;\theta). \end{aligned} \tag{121}$$

This explicit representation avoids numerical convolution of the marks themselves.

A.3 Derivation of the stationary variance

Proposition 3.6 gives

$$q_\infty := \mathbb{E}(X_\infty^2) = \frac{(2\kappa\mu + \sigma^2 + 2\lambda m_1)\bar{m} + \lambda m_2}{2\kappa}. \tag{122}$$

Since

$$\operatorname{Var}(X_\infty) = q_\infty - \bar{m}^2,$$

multiplication by 2κ yields

$$\begin{aligned} 2\kappa \operatorname{Var}(X_\infty) &= (2\kappa\mu + \sigma^2 + 2\lambda m_1)\bar{m} + \lambda m_2 - 2\kappa\bar{m}^2 \\ &= \sigma^2\bar{m} + \lambda m_2 + 2\bar{m}\{\kappa\mu + \lambda m_1 - \kappa\bar{m}\}. \end{aligned} \tag{123}$$

From

$$\bar{m} = \mu + \frac{\lambda m_1}{\kappa},$$

we have

$$\kappa\bar{m} = \kappa\mu + \lambda m_1.$$

The last term in (123) therefore vanishes, leaving

$$2\kappa \operatorname{Var}(X_\infty) = \sigma^2\bar{m} + \lambda m_2.$$

Hence

$$\operatorname{Var}(X_\infty) = \frac{\sigma^2\bar{m} + \lambda m_2}{2\kappa}, \tag{124}$$

which is the variance formula in (36).

For the JCIR–Zeghdoudi specification,

$$m_1 = \frac{2(\theta+3)}{\theta(\theta+2)}, \qquad m_2 = \frac{6(\theta+4)}{\theta^2(\theta+2)},$$

so

$$\operatorname{Var}(X_\infty) = \frac{1}{2\kappa}\left[\sigma^2\left\{\mu + \frac{2\lambda(\theta+3)}{\kappa\theta(\theta+2)}\right\} + \frac{6\lambda(\theta+4)}{\theta^2(\theta+2)}\right]. \tag{125}$$

References

Aït-Sahalia, Y. and Jacod, J. (2009). Testing for jumps in a discretely observed process. *The Annals of Statistics*, 37(1), 184–222. https://doi.org/10.1214/07-AOS568.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1), 1–37. https://doi.org/10.1093/jjfinec/nbh001.

Chouia, S. and Zeghdoudi, H. (2021). The XLindley distribution: properties and application. *Journal of Statistical Theory and Applications*, 20(2), 318–327. https://doi.org/10.2991/jsta.d.210607.001.

Cobb, G. W. (1978). The problem of the Nile: conditional solution to a change-point problem. *Biometrika*, 65(2), 243–251. https://doi.org/10.1093/biomet/65.2.243.

Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman & Hall/CRC, Boca Raton.

Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2), 385–407. https://doi.org/10.2307/1911242.

Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes, Volume I: Elementary Theory and Methods*. 2nd ed., Springer, New York. https://doi.org/10.1007/b97277.

Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, 6(4), 379–406. https://doi.org/10.1111/j.1467-9965.1996.tb00123.x.

Duffie, D., Pan, J., and Singleton, K. J. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6), 1343–1376. https://doi.org/10.1111/1468-0262.00164.

Durbin, J. and Koopman, S. J. (2001). *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford.

Fan, J. and Wang, Y. (2007). Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of the American Statistical Association*, 102(480), 1349–1362. https://doi.org/10.1198/016214507000001067.

Ghitany, M. E., Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78(4), 493–506. https://doi.org/10.1016/j.matcom.2007.06.007.

Kessler, M. (1997). Estimation of an ergodic diffusion from discrete observations. *Scandinavian Journal of Statistics*, 24(2), 211–229. https://doi.org/10.1111/1467-9469.00059.

Khodja, N., Gemeay, A. M., Zeghdoudi, H., Karakaya, K., Alshangiti, A. M., Bakr, M. E., Balogun, O. S., Muse, A. H., and Hussam, E. (2023). Modeling voltage real data set by a new version of Lindley distribution. *IEEE Access*, 11, 67220–67229. https://doi.org/10.1109/ACCESS.2023.3287926.

Lee, S. S. and Mykland, P. A. (2008). Jumps in financial markets: a new nonparametric test and jump dynamics. *The Review of Financial Studies*, 21(6), 2535–2563. https://doi.org/10.1093/rfs/hhm056.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36(2), 270–296. https://doi.org/10.1111/j.1467-9469.2008.00622.x.

Masuda, H. (2007). Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stochastic Processes and their Applications*, 117(1), 35–56. https://doi.org/10.1016/j.spa.2006.04.010.

Messaadia, H. and Zeghdoudi, H. (2018). Zeghdoudi distribution and its applications. *International Journal of Computing Science and Mathematics*, 9(1), 58–65. https://doi.org/10.1504/IJCSM.2018.090722.

Saaidia, N., Belhamra, T., and Zeghdoudi, H. (2024). On ZLindley distribution: statistical properties and applications. *Studies in Engineering and Exact Sciences*, 5(1), 3078–3097. https://doi.org/10.54021/seesv5n1-153.

Saaidia, N., Pakyari, R., and Zeghdoudi, H. (2026). The Q-Lindley distribution: goodness-of-fit tests, modeling, inference, and applications. *Journal of Statistical Theory and Applications*, 25, Article 6. https://doi.org/10.1007/s44199-025-00148-5.

Seghier, F. Z., Zeghdoudi, H., Ameeq, M., and Kanwal, S. (2026). The Poisson–QGamma distribution: properties, estimation methods, regression modeling, and applications in engineering count data. *Stats*, 9(3), Article 52. https://doi.org/10.3390/stats9030052.

Sen, S., Maiti, S. S., and Chandra, N. (2016). The Xgamma distribution: statistical properties and application. *Journal of Modern Applied Statistical Methods*, 15(1), 774–788. https://doi.org/10.22237/jmasm/1462077420.

Shimizu, Y. and Yoshida, N. (2006). Estimation of parameters for diffusion processes with jumps from discrete observations. *Statistical Inference for Stochastic Processes*, 9(3), 227–277. https://doi.org/10.1007/s11203-005-8114-x.

Beghriche, A., Azouz, Z., and Zeghdoudi, H. (2025). Entropy-based properties and Bayesian modeling of the Xexponential distribution with applications in astronomy. *Astronomy and Computing*, 101054. https://doi.org/10.1016/j.ascom.2025.101054.