Supplement to "Saddlepoint Approximations for Hawkes Jump-Diffusion Processes with an Application to Risk Management"

Yacine Aït-Sahalia
Department of Economics, Princeton University
and
Roger J. A. Laeven
Department of Quantitative Economics, University of Amsterdam

May 18, 2025

Abstract

This text serves as an Online Supplement to the paper "Saddlepoint Approximations for Hawkes Jump-Diffusion Processes with an Application to Risk Management." For context, definitions and other details see the paper. In the first section we introduce some notation, recall and expand our model specifications and discuss operator methods. In the second section we provide closed-form expressions for the corresponding moments and cumulant generating functions. In the third and fourth sections we provide detailed technical derivations of our saddlepoint approximations and Monte Carlo simulation results. The fifth section provides some preliminaries for eigenvector centrality in spectral graph theory.

A Notation, Model Specifications and Operator Methods

A.1 Notation and Preliminaries

Recall that

$$\begin{cases} \mathbb{P}\left[N_{i,t+\Delta}-N_{i,t}=0|\mathcal{F}_t\right]=1-\lambda_{i,t}\Delta+o(\Delta) \\ \mathbb{P}\left[N_{i,t+\Delta}-N_{i,t}=1|\mathcal{F}_t\right]=\lambda_{i,t}\Delta+o(\Delta) \\ \mathbb{P}\left[N_{i,t+\Delta}-N_{i,t}>1|\mathcal{F}_t\right]=o(\Delta), \end{cases} \tag{A.1}$$

with $\lambda_{i,t}$ given in (1). We assume that the constant parameters $\lambda_{i,\infty}$ in (1) are non-negative for all $i=1,\ldots,m$, and that the real-valued functions $g_{i,j}(u)$ and $\phi_{i,j}(z)$ are non-negative for all $u\geq 0$, all $|z|\geq 0$, and for all $i,j=1,\ldots,m$. The requirement that $\lambda_{i,\infty}, g_{i,j}, \phi_{i,j}$ are non-negative for all $i,j=1,\ldots,m$ guarantees that the intensity processes are non-negative with probability one. We denote by $\mathbf{\Lambda}_\infty$ the $m\times 1$ vector with components $\lambda_{i,\infty}$ and by $\mathbf{\Gamma}:=\int_0^\infty \mathbf{G}_u\,\mathrm{d}u$ the $m\times m$ matrix, where $\mathbf{G}_u$ is the matrix with elements $g_{i,j}(u)$. Furthermore, we denote by $\lambda_i:=\mathbb{E}\left[\lambda_{i,t}\right]$ the unconditional expected jump intensity, and by $\mathbf{\Lambda}$ the $m\times 1$ vector with components λ_i. Because from (A.1), $\mathbb{E}\left[\mathrm{d}N_{i,s}\right]=\lambda_i\,\mathrm{d}s$, we obtain that

$$\begin{aligned}\lambda_i &= \lambda_{i,\infty}+\sum_{j=1}^m \lambda_j\int_{-\infty}^t \mathbb{E}\left[\phi_{i,j}\left(Z_{j,s}\right)\right]g_{i,j}\left(t-s\right)\,\mathrm{d}s=\lambda_{i,\infty}+\sum_{j=1}^m\lambda_j\int_{-\infty}^t \zeta g_{i,j}\left(t-s\right)\,\mathrm{d}s \\ &= \lambda_{i,\infty}+\zeta\sum_{j=1}^m\left(\int_0^\infty g_{i,j}\left(u\right)\,\mathrm{d}u\right)\lambda_j, \end{aligned}\tag{A.2}$$

where we assume $\mathbb{E}\left[\phi_{i,j}\left(Z_{j,s}\right)\right]=\zeta,\ i,j=1,\ldots,m$, which may be normalized to unity, $\zeta\equiv 1$, since there is a multiplication with the bank-pair-specific $g_{i,j}$. Thus, in vector form, $\mathbf{\Lambda}=\mathbf{\Lambda}_\infty+\zeta\mathbf{\Gamma}\mathbf{\Lambda}$. Hence, $\mathbf{\Lambda}=\left(\mathbf{I}-\zeta\mathbf{\Gamma}\right)^{-1}\mathbf{\Lambda}_\infty$, with $\mathbf{I}$ the identity matrix. By assuming that all the elements of $\mathbf{\Lambda}$ are positive and finite we ensure stationarity of the model.

We focus specifically on the case with exponential decay (2), such that

$$\mathrm{d}\lambda_{i,t}=\alpha_i\left(\lambda_{i,\infty}-\lambda_{i,t}\right)\,\mathrm{d}t+\sum_{j=1}^m\beta_{i,j}\phi_{i,j}\left(Z_{j,t}\right)\,\mathrm{d}N_{j,t}. \tag{A.3}$$

Indeed,

$$\begin{aligned}
\mathrm{d}\lambda_{i,t} &= \sum_{j=1}^{m} \phi_{i,j}\left(Z_{j,t}\right) g_{i,j}\left(0\right) \mathrm{d}N_{j,t} + \sum_{j=1}^{m} \int_{-\infty}^{t} \phi_{i,j}\left(Z_{j,s}\right) \frac{\mathrm{d}g_{i,j}}{\mathrm{d}t}\left(t-s\right) \mathrm{d}N_{j,s}\,\mathrm{d}t \\
&= \sum_{j=1}^{m} \beta_{i,j}\phi_{i,j}\left(Z_{j,t}\right) \mathrm{d}N_{j,t} - \alpha_i \sum_{j=1}^{m} \int_{-\infty}^{t} \phi_{i,j}\left(Z_{j,s}\right) \beta_{i,j} e^{-\alpha_i (t-s)}\, \mathrm{d}N_{j,s}\,\mathrm{d}t \\
&= \sum_{j=1}^{m} \beta_{i,j}\phi_{i,j}\left(Z_{j,t}\right) \mathrm{d}N_{j,t} - \alpha_i \left(\lambda_{i,t} - \lambda_{i,\infty}\right) \mathrm{d}t. \qquad \text{(A.4)}
\end{aligned}$$

Under exponential decay, the matrix $\mathbf{\Gamma}$ is given by

$$\mathbf{\Gamma} = \begin{pmatrix} \frac{\beta_{1,1}}{\alpha_1} & \cdots & \frac{\beta_{1,m}}{\alpha_1} \\ \vdots & \ddots & \vdots \\ \frac{\beta_{m,1}}{\alpha_m} & \cdots & \frac{\beta_{m,m}}{\alpha_m} \end{pmatrix}. \qquad \text{(A.5)}$$

In an extension of the model specification in (5), we allow for stochastic volatility with an instantaneous variance process $V_{i,t}$ of the Heston type:

$$\mathrm{d}Y_{i,t} = \mu_i\,\mathrm{d}t + \sqrt{V_{i,t}}\,\mathrm{d}W_{i,t}^{Y} + Z_{i,t}\,\mathrm{d}N_{i,t} \qquad \text{(A.6)}$$

$$\mathrm{d}V_{i,t} = \kappa_i(\theta_i - V_{i,t})\,\mathrm{d}t + \eta_i\sqrt{V_{i,t}}\,\mathrm{d}W_{i,t}^{V}, \qquad \text{(A.7)}$$

where κ_i, θ_i, η_i are constant parameters satisfying $2\kappa_i\theta_i \geq \eta_i^2$. In principle, we can allow for any other stochastic volatility model, provided the moments and other quantities we compute are suitably adapted. But our main focus in this paper is on the mutual excitation phenomenon in jumps that occurs in crisis times and the implications this phenomenon has for risk management calculations. In the extended model (A.6)–(A.7), we allow not only for correlations among the Brownian motions appearing in equation (A.6) and among those appearing in (A.7), but also for correlations between the Brownian motions appearing in equations (A.6) and (A.7). The latter correlations can capture a leverage effect.[1,2]

[1]Our model does not include time variation in the correlations between the Brownian motions. However, the total realized correlation structure of asset returns in our model is highly time-varying due to the presence of mutually exciting jumps: in effect, correlations measured over a period that includes one or more mutually exciting episodes will be substantially higher, in fact often close to one as is the case in the data, owing to the fact that the jumps that result from mutual excitation are primarily (but not exclusively) of the same sign. In periods where jumps occur, even though there are few of them, jumps tend to dominate all the realized statistics of asset returns; see Aït-Sahalia et al. (2017). Indeed, in the presence of systematic jumps, the Brownian correlations do not play a main role, even if they ramp up in crises times: jumps swamp all other components if and when they occur.

[2]Self- and cross-excitation induce a feedback feature through which the intensity of jumps rises in response to jumps themselves, making future jumps more probable. This feedback feature can be thought

A.2 The Univariate Model

We first consider the univariate case $m = 1$. We provide explicit expressions of the unconditional moments, as basic risk measures, and cumulant generating function up to the second order in the time interval Δ. The univariate jump-diffusion with Heston-type stochastic volatility and self-exciting jumps reads as follows:

$$\begin{cases} \mathrm{d}Y_t = \mu\,\mathrm{d}t + \sqrt{V_t}\,\mathrm{d}W_t^Y + Z_t\,\mathrm{d}N_t \\ \mathrm{d}V_t = \kappa(\theta - V_t)\,\mathrm{d}t + \eta\sqrt{V_t}\,\mathrm{d}W_t^V \\ \mathrm{d}\lambda_t = \alpha\left(\lambda_\infty - \lambda_t\right)\,\mathrm{d}t + \beta\phi(Z_t)\,\mathrm{d}N_t. \end{cases} \tag{A.8}$$

Here, $\mathbb{E}\left[\mathrm{d}W_t^Y\,\mathrm{d}W_t^V\right] =: \rho^V\,\mathrm{d}t$ and

$$\lambda := \mathbb{E}\left[\lambda_t\right] = \frac{\alpha\lambda_\infty}{\alpha - \beta\mathbb{E}\left[\phi\left[Z\right]\right]}.$$

In the univariate case, the model can only generate self-excitation in the time-series dimension, i.e., there is no cross-excitation in the (absent) cross-sectional dimension. The model (A.8) reduces to the familiar model with compound Poisson process jumps when $\beta = 0$ and $\lambda_0 = \lambda_\infty$, in which case $\lambda_t = \lambda_\infty = \lambda$ at all t.

The jump amplification function ϕ is a function selected jointly with the distribution of Z_t to maintain the positivity and stationarity of λ_t. In earlier applications of Hawkes processes, the jump intensities respond only to jumps, not to their sizes or signs. This was to preserve the autonomous nature of the couple (N, λ) facilitating the analysis of the model. As we show, it is possible, however, to introduce a level effect and an asymmetry effect by having the term $\beta\phi(Z_t)\,\mathrm{d}N_t$ instead of just a term $\beta\,\mathrm{d}N_t$ in the third equation of (A.8). This term can for instance capture a larger increase in jump intensity following a large negative jump in the profit and loss account.

We will leave the specific form of the jump amplification function ϕ and the jump magnitude distribution F_Z unrestricted at this stage, provided the respective moments are finite. We provide closed-form expressions depending on the moments of $\phi(Z_t)$ and Z_t. Throughout this and the next sections, the generic moments of the jump magnitude Z_t are denoted by $M[Z, k] := \mathbb{E}\left[Z_t^k\right]$, and similarly for $\phi(Z_t)$ and its products with Z_t. Theorem S.1 provides the univariate moments explicitly. Furthermore, we provide in Theorem S.2 an

of as playing the same role for jumps as ARCH models do for volatility. In an ARCH model, large returns induce large volatilities increasing the likelihood of observing large future returns. If no further large returns occur volatility mean reverts. Similarly, under mutual excitation, jumps induce larger intensities of jumps, increasing the likelihood of observing further jumps. If no further jumps happen to occur, the intensity mean reverts. GARCH models add past volatility into the feedback relation.

explicit expression of the autocovariance function of the squared process and provide in Theorem S.3 an explicit expression of the cumulant generating function (cgf) defined as

$$\overline{K}(\Delta, u) = \log \mathbb{E}\left[e^{u\,(Y_{t+\Delta}-Y_t)}\right], \qquad u \in \mathbb{R}, \tag{A.9}$$

up to order Δ^2, assuming state-independent volatility as in the main text. Moreover, Theorem S.4 provides the cgf explicitly up to order Δ^2 for the extended model (A.8).

A.3 The Bivariate Model

Next, we proceed to the bivariate case. We assume state-independent volatilities. That is, we assume model (5) with $m = 2$, which entails for the jump intensities:

$$\begin{cases} \mathrm{d}\lambda_{1,t} = \alpha_1\left(\lambda_{1,\infty} - \lambda_{1,t}\right)\mathrm{d}t + \beta_{1,1}\phi_{1,1}\left(Z_{1,t}\right)\mathrm{d}N_{1,t} + \beta_{1,2}\phi_{1,2}\left(Z_{2,t}\right)\mathrm{d}N_{2,t}; \\ \mathrm{d}\lambda_{2,t} = \alpha_2\left(\lambda_{2,\infty} - \lambda_{2,t}\right)\mathrm{d}t + \beta_{2,1}\phi_{2,1}\left(Z_{1,t}\right)\mathrm{d}N_{1,t} + \beta_{2,2}\phi_{2,2}\left(Z_{2,t}\right)\mathrm{d}N_{2,t}. \end{cases}$$

In this case, the jump amplification function $\phi_{i,j}$, which is specified jointly with the distribution of $Z_{j,t}$ such that it maintains positivity and stationarity of the bivariate vector $\boldsymbol{\lambda}_t$, features two indexes: j for the sector initially affected and i for the sector that is excited, $i, j = 1, 2$.

We again leave the specific forms of the jump amplification functions and the distributions of the jump magnitudes essentially unrestricted at this stage, and provide expressions as functions of the moments of $Z_{j,t}$ and $\phi_{i,j}(Z_{j,t})$, for which we write $M[Z_j, k] := \mathbb{E}\left[Z_{j,t}^k\right]$ and similarly for $\phi_{i,j}(Z_{j,t})$ and its products with $Z_{j,t}$. Theorems S.5, S.6, and S.7 provide the moment functions and the auto- and cross-covariance functions of the regular and squared processes explicitly.

We also provide in Theorem S.8 an explicit expression of the marginal cgf in the bivariate model, given by

$$\overline{K}(\Delta, u_1) = \log \mathbb{E}\left[e^{u_1\,(Y_{1,t+\Delta}-Y_{1,t})}\right], \qquad u_1 \in \mathbb{R}, \tag{A.10}$$

up to order Δ^2.

Finally, we provide in Theorem S.9 an explicit expression of the bivariate cgf in the bivariate model, given by

$$\overline{K}(\Delta, u_1, u_2) = \log \mathbb{E}\left[e^{u_1\,(Y_{1,t+\Delta}-Y_{1,t}) + u_2\,(Y_{2,t+\Delta}-Y_{2,t})}\right], \qquad \mathbf{u} = (u_1, u_2) \in \mathbb{R}^2, \tag{A.11}$$

up to order Δ^2.

A.4 Jump Amplification Function

Whereas we leave the jump amplification function and jump magnitude distribution unrestricted at this stage, they need to be specified when we derive expressions for the saddlepoint approximations. Then, we will assume (6)–(7). Under specification (7), the jump magnitude moments take the following form:

$$\mathbb{E}\left[Z_j^k\right] = (-1)^k \frac{k!\, p_j}{\gamma_{j,-}^k} + \frac{k!\,(1-p_j)}{\gamma_{j,+}^k}, \qquad k = 1, 2, \ldots. \tag{A.12}$$

Hence, assuming ϕ_j to be doubly exponential as in (6),

$$\begin{aligned}\mathbb{E}\left[\phi_j(Z_j)\right] &= p_j c_{j,-}(1 - \mathbb{E}\left[\exp(-\xi_{j,-}(-Z_{j,-}))\right]) + (1-p_j)c_{j,+}(1 - \mathbb{E}\left[\exp(-\xi_{j,+}Z_{j,+})\right]) \\ &= p_j c_{j,-}\left(1 - \frac{\gamma_{j,-}}{\gamma_{j,-}+\xi_{j,-}}\right) + (1-p_j)c_{j,+}\left(1 - \frac{\gamma_{j,+}}{\gamma_{j,+}+\xi_{j,+}}\right).\end{aligned}$$

By equating $\mathbb{E}\left[\phi_j(Z_j)\right]$ to 1, normalizing $\frac{c_{j,+}}{c_{j,-}} =: \chi_j$, and solving for $c_{j,-}$ (depending on the new parameter χ_j) we obtain:

$$c_{j,-} = \frac{1}{p_j(1 - \frac{\gamma_{j,-}}{\gamma_{j,-}+\xi_{j,-}}) + (1-p_j)\chi_j(1 - \frac{\gamma_{j,+}}{\gamma_{j,+}+\xi_{j,+}})}. \tag{A.13}$$

Furthermore, one easily verifies that

$$\begin{aligned}\mathbb{E}\left[Z_j^k \phi_j(Z_j)\right] = p_j c_{j,-}&\left((-1)^k \frac{k!}{\gamma_{j,-}^k} + (-1)^{k+1}\frac{\gamma_{j,-}}{\gamma_{j,-}+\xi_{j,-}}\frac{k!}{(\gamma_{j,-}+\xi_{j,-})^k}\right) \\ &+ (1-p_j)c_{j,+}\left(\frac{k!}{\gamma_{j,+}^k} - \frac{\gamma_{j,+}}{\gamma_{j,+}+\xi_{j,+}}\frac{k!}{(\gamma_{j,+}+\xi_{j,+})^k}\right),\end{aligned}$$

which reduces to the above expression for $\mathbb{E}\left[\phi_j(Z_j)\right]$ when $k = 0$ (realizing that $0! = 1$), and that

$$\begin{aligned}\mathbb{E}\left[\phi_j(Z_j)^2\right] = p_j c_{j,-}^2&\left(1 - 2\frac{\gamma_{j,-}}{\gamma_{j,-}+\xi_{j,-}} + \frac{\gamma_{j,-}}{\gamma_{j,-}+2\xi_{j,-}}\right) \\ &+ (1-p_j)c_{j,+}^2\left(1 - 2\frac{\gamma_{j,+}}{\gamma_{j,+}+\xi_{j,+}} + \frac{\gamma_{j,+}}{\gamma_{j,+}+2\xi_{j,+}}\right).\end{aligned}$$

A.5 Operator Methods

The computation of the moments and cumulants of the P&L model proceeds using operator methods (e.g., Aït-Sahalia et al. 2010). In Aït-Sahalia et al. (2015) these methods were employed on a simpler model, without the jump amplification function, for a simpler set of moments, and without consideration of the cumulant generating function. We first compute the one-step ahead conditional moments using the full Markovian state vector. Next, we take expected values, integrating out the latent state variables (volatilities if stochastic, and jump intensities). Doing so, we obtain expressions that depend only upon observable state variables. These expressions, however, contain all the parameters of the model, including those related to the latent state variables.

The dynamics of our P&L model depend upon additional latent variables $\boldsymbol{\varsigma}_t = (\mathbf{V}_t, \boldsymbol{\Lambda}_t)$, where $\mathbf{V}_t$ is the P&L's stochastic variance and $\boldsymbol{\Lambda}_t$ is the stochastic jump intensity. The computation of the corresponding moments and cumulants requires the evaluation of expectations of functions of the form $\psi(\Delta, \mathbf{y}_1, \mathbf{y}_0, \boldsymbol{\varsigma}_1, \boldsymbol{\varsigma}_0)$, where the subscripts 1 and 0 refer to two dates separated by a time interval Δ. Recall from the main text that $Y_{i,t}$ denotes bank i's market value of the trading book at time t. Ultimately, we need to evaluate expressions of the form $\mathbb{E}\left[\psi(\Delta, \mathbf{Y}_1, \mathbf{Y}_0, \boldsymbol{\varsigma}_1, \boldsymbol{\varsigma}_0)\right]$, for specific choices of ψ. To compute any conditional moment or cumulant, we use the explicit expression of the infinitesimal Markov generator $\mathcal{A}$ of our P&L model and expand it as a series in the time interval Δ as follows:

$$\begin{aligned}\mathbb{E}_{\mathbf{Y}_1, \boldsymbol{\varsigma}_1}\left[\psi(\Delta, \mathbf{Y}_1, \mathbf{Y}_0, \boldsymbol{\varsigma}_1, \boldsymbol{\varsigma}_0) | \mathbf{Y}_0, \boldsymbol{\varsigma}_0\right] &= \exp\left(\Delta \mathcal{A}\right) \cdot \psi(0, \mathbf{Y}_0, \mathbf{Y}_0, \boldsymbol{\varsigma}_0, \boldsymbol{\varsigma}_0) \\ &= \sum_{j=0}^{J} \frac{\Delta^j}{j!} \left(\mathcal{A}^j \cdot \psi\right)(0, \mathbf{Y}_0, \mathbf{Y}_0, \boldsymbol{\varsigma}_0, \boldsymbol{\varsigma}_0) + O_p\left(\Delta^{J+1}\right), \qquad \text{(A.14)}\end{aligned}$$

where subscripts in $\mathbb{E}_{\mathbf{Y}_1, \boldsymbol{\varsigma}_1}$ indicate the random variables that the expected value operates on, and $\mathcal{A}^j \cdot \psi$ is defined recursively by $\mathcal{A}^j \cdot \psi = \mathcal{A} \cdot (\mathcal{A}^{j-1} \cdot \psi)$ for all $j \geq 1$. Remarkably, the iterates $\mathcal{A}^j \cdot \psi$, hence the terms in (A.14), can be evaluated in closed form for the moment functions of interest.

Hence, we can obtain the conditional expectation of ψ, using the full state vector including its unobservable components, $\mathbb{E}_{\mathbf{Y}_1, \boldsymbol{\varsigma}_1}\left[\psi(\Delta, \mathbf{Y}_1, \mathbf{Y}_0, \boldsymbol{\varsigma}_1, \boldsymbol{\varsigma}_0) | \mathbf{Y}_0, \boldsymbol{\varsigma}_0\right]$. We then need to condition down by integrating out the unobservable state variables. All the expectations are taken with respect to the law of the process at the true parameter values. From the law of iterated expectations,

$$\begin{aligned}\mathbb{E}\left[\psi(\Delta, \mathbf{Y}_1, \mathbf{Y}_0, \boldsymbol{\varsigma}_1, \boldsymbol{\varsigma}_0)\right] &= \mathbb{E}_{\mathbf{Y}_1, \mathbf{Y}_0, \boldsymbol{\varsigma}_1, \boldsymbol{\varsigma}_0}\left[\mathbb{E}_{\mathbf{Y}_1, \boldsymbol{\varsigma}_1}\left[\psi(\Delta, \mathbf{Y}_1, \mathbf{Y}_0, \boldsymbol{\varsigma}_1, \boldsymbol{\varsigma}_0) | \mathbf{Y}_0, \boldsymbol{\varsigma}_0\right]\right] \\ &= \sum_{j=0}^{J} \frac{\Delta^j}{j!} \mathbb{E}_{\mathbf{Y}_0, \boldsymbol{\varsigma}_0}\left[\left(\mathcal{A}^j \cdot \psi\right)(0, \mathbf{Y}_0, \mathbf{Y}_0, \boldsymbol{\varsigma}_0, \boldsymbol{\varsigma}_0)\right] + O\left(\Delta^{J+1}\right), \qquad \text{(A.15)}\end{aligned}$$

such that the last step in the necessary calculations involves computing unconditional expectations with respect to the stationary law of the state variables. In what follows, we provide the expressions for these moments and cumulants.

B Explicit Expressions for the Moments and Cumulant Generating Functions

Throughout this section, we use the shorthand notation $\Delta Y_{i,t} := Y_{i,t+\Delta} - Y_{i,t}$, $i = 1, 2$, and omit the index i when $m = 1$.

B.1 The Univariate Case

Theorem S.1. *For the univariate model (A.8), the moments are given in closed form up to order Δ^2 by the following expressions:*

$$\begin{aligned}
\mathbb{E}\left[\Delta Y_t\right] &= (\mu + \lambda M\left[Z,1\right])\Delta + o(\Delta^2) \\
\mathbb{E}\left[(\Delta Y_t - \mathbb{E}\left[\Delta Y_t\right])^2\right] &= (\theta + \lambda M\left[Z,2\right])\Delta \\
&\quad + \beta\lambda\left(M\left[Z\phi(Z),1\right] + \frac{\beta M\left[Z,1\right] M\left[\phi(Z),2\right]}{2(\alpha - \beta M\left[\phi(Z),1\right])}\right) M\left[Z,1\right]\Delta^2 + o(\Delta^2) \\
\mathbb{E}\left[(\Delta Y_t - \mathbb{E}\left[\Delta Y_t\right])^3\right] &= \lambda M\left[Z,3\right]\Delta \\
&\quad + \frac{3}{2}\left(\eta\theta\rho^V + \beta\lambda\left(M\left[Z,2\right] M\left[Z\phi(Z),1\right] + M\left[Z,1\right] M\left[Z^2\phi(Z),1\right]\right.\right. \\
&\quad \left.\left. + \frac{\beta M\left[Z,1\right] M\left[Z,2\right] M\left[\phi(Z),2\right]}{(\alpha - \beta M\left[\phi(Z),1\right])}\right)\right)\Delta^2 + o(\Delta^2) \\
\mathbb{E}\left[(\Delta Y_t - \mathbb{E}\left[\Delta Y_t\right])^4\right] &= \lambda M\left[Z,4\right]\Delta + \left(\frac{3\theta\eta^2}{2\kappa} + 3\theta^2 + 6\theta\lambda M\left[Z,2\right]\right. \\
&\quad + \beta\lambda\left(2M\left[Z,3\right] M\left[Z\phi(Z),1\right] + 3M\left[Z,2\right] M\left[Z^2\phi(Z),1\right] + 2M\left[Z,1\right] M\left[Z^3\phi(Z),1\right]\right) \\
&\quad + 3\lambda\left(\lambda + \frac{\beta^2 M\left[\phi(Z),2\right]}{2(\alpha - \beta M\left[\phi(Z),1\right])}\right) M\left[Z,2\right]^2 \\
&\quad \left. + \frac{4\beta^2\lambda M\left[Z,1\right] M\left[Z,3\right] M\left[\phi(Z),2\right]}{(\alpha - \beta M\left[\phi(Z),1\right])}\right)\Delta^2 + o(\Delta^2).
\end{aligned}$$

Furthermore, the autocovariance function of the process is given by

$$\begin{aligned}
\mathbb{E}\left[(\Delta Y_t - \mathbb{E}\left[\Delta Y_t\right])(\Delta Y_{t+\Delta} - \mathbb{E}\left[\Delta Y_{t+\Delta}\right])\right] &= \frac{1}{2(\alpha - \beta M\left[\phi(Z),1\right])}\beta\lambda M\left[Z,1\right] \\
&\quad \times (2\alpha M\left[Z\phi(Z),1\right] - 2\beta M\left[\phi(Z),1\right] M\left[Z\phi(Z),1\right] + \beta M\left[Z,1\right] M\left[\phi(Z),2\right])\Delta^2 + o(\Delta^2).
\end{aligned}$$

Remark S.1. *The model reduces to a Poissonian jump-diffusion when $\beta = 0$. Then,*

$$\begin{aligned}
\mathbb{E}\left[\Delta Y_t\right] &= (\mu + \lambda M\left[Z,1\right])\Delta + o(\Delta^2) \\
\mathbb{E}\left[(\Delta Y_t - \mathbb{E}\left[\Delta Y_t\right])^2\right] &= (\theta + \lambda M\left[Z,2\right])\Delta + o(\Delta^2) \\
\mathbb{E}\left[(\Delta Y_t - \mathbb{E}\left[\Delta Y_t\right])^3\right] &= \lambda M\left[Z,3\right]\Delta + \frac{3}{2}\eta\theta\rho^V\Delta^2 + o(\Delta^2) \\
\mathbb{E}\left[(\Delta Y_t - \mathbb{E}\left[\Delta Y_t\right])^4\right] &= \lambda M\left[Z,4\right]\Delta + 3\left(\frac{\theta\eta^2}{2\kappa} + (\theta + \lambda M\left[Z,2\right])^2\right)\Delta^2 + o(\Delta^2),
\end{aligned}$$

and the autocorrelation is identical zero.

Theorem S.2. *For the model* (5) *with* $m = 1$*, the autocovariance function of the squared process is given in closed form up to order* Δ^2 *by the following expression:*

$$\mathbb{E}\left[(\Delta Y_t - \mathbb{E}\left[\Delta Y_t\right])^2(\Delta Y_{t+\Delta} - \mathbb{E}\left[\Delta Y_{t+\Delta}\right])^2\right] = \frac{1}{2(\alpha - \beta M\left[\phi(Z),1\right])}\beta\lambda M\left[Z,2\right]$$
$$\times(2\alpha M\left[Z^2\phi(Z),1\right] - 2\beta M\left[\phi(Z),1\right] M\left[Z^2\phi(Z),1\right] + \beta M\left[Z,2\right] M\left[\phi(Z),2\right])\Delta^2 + o(\Delta^2).$$

Theorem S.3. *For the model* (5) *with* $m = 1$*, the cgf is given in closed form up to order* Δ^2 *by the following expression:*

$$\begin{aligned}
&\overline{K}(\Delta,u)\\
&= \left(\mu u + \frac{\sigma^2 u^2}{2} + \lambda\left(L(u,Z)-1\right)\right)\Delta\\
&+ \frac{\beta\lambda(L(u,Z)-1)\left(2L(u,\phi(Z))(\alpha - \beta M\left[\phi(Z),1\right]) + \beta M\left[\phi(Z),2\right]\left(L(u,Z)-1\right) - 2\alpha M\left[\phi(Z),1\right] + 2\beta M\left[\phi(Z),1\right]^2\right)}{4(\alpha - \beta M\left[\phi(Z),1\right])}\Delta^2\\
&+ o(\Delta^2),
\end{aligned}$$

where $L(u,Z) = \mathbb{E}\left[e^{uZ}\right]$ *and* $L(u,\phi(Z)) = \mathbb{E}\left[\phi(Z)e^{uZ}\right]$.

Theorem S.4. *For the extended univariate model* (A.8)*, the cgf is given in closed form up to order* Δ^2 *by the following expression:*

$$\begin{aligned}
&\overline{K}(\Delta,u)\\
&= \left(\mu u + \frac{\theta u^2}{2} + \lambda\left(L(u,Z)-1\right)\right)\Delta + \frac{\eta\theta u^3(4\kappa\rho^V + \eta u)}{16\kappa}\Delta^2\\
&+ \frac{\beta\lambda(L(u,Z)-1)\left(2L(u,\phi(Z))(\alpha - \beta M\left[\phi(Z),1\right]) + \beta M\left[\phi(Z),2\right]\left(L(u,Z)-1\right) - 2\alpha M\left[\phi(Z),1\right] + 2\beta M\left[\phi(Z),1\right]^2\right)}{4(\alpha - \beta M\left[\phi(Z),1\right])}\Delta^2\\
&+ o(\Delta^2),
\end{aligned}$$

where $L(u,Z) = \mathbb{E}\left[e^{uZ}\right]$ *and* $L(u,\phi(Z)) = \mathbb{E}\left[\phi(Z)e^{uZ}\right]$.

B.2 The Bivariate Case

Theorem S.5. *For the model* (5) *with* $m = 2$*, the first and second moments are given in closed form up to order* Δ^2 *by the following expressions:*

$$
\begin{aligned}
&\mathbb{E}\left[\Delta Y_{1,t}\right] = \left(\mu_1 + \lambda_1 M\left[Z_1, 1\right]\right)\Delta + o(\Delta^2) \\
&\mathbb{E}\left[\Delta Y_{2,t}\right] = \left(\mu_2 + \lambda_2 M\left[Z_2, 1\right]\right)\Delta + o(\Delta^2) \\
&\mathbb{E}\left[(\Delta Y_{1,t} - \mathbb{E}\left[\Delta Y_{1,t}\right])^2\right] = \left(\sigma_1^2 + \lambda_1 M\left[Z_1, 2\right]\right)\Delta \\
&\qquad\qquad + \left((\nu_{1,1} - \lambda_1^2) M\left[Z_1, 1\right] + \beta_{1,1}\lambda_1 M\left[Z_1\phi_{1,1}(Z_1), 1\right]\right) M\left[Z_1, 1\right]\Delta^2 + o(\Delta^2) \\
&\mathbb{E}\left[(\Delta Y_{1,t} - \mathbb{E}\left[\Delta Y_{1,t}\right])(\Delta Y_{2,t} - \mathbb{E}\left[\Delta Y_{2,t}\right])\right] = \rho\sigma_1\sigma_2\Delta \\
&\qquad\qquad + \frac{1}{2}\left(2\left(\nu_{1,2} - \lambda_1\lambda_2\right) M\left[Z_1, 1\right] M\left[Z_2, 1\right]\right. \\
&\qquad\qquad\quad + \beta_{1,2}\lambda_2 M\left[Z_1, 1\right] M\left[Z_2\phi_{1,2}(Z_2), 1\right] \\
&\qquad\qquad\quad \left. + \beta_{2,1}\lambda_1 M\left[Z_2, 1\right] M\left[Z_1\phi_{2,1}(Z_1), 1\right]\right)\Delta^2 + o(\Delta^2) \\
&\mathbb{E}\left[(\Delta Y_{2,t} - \mathbb{E}\left[\Delta Y_{2,t}\right])^2\right] = \left(\sigma_2^2 + \lambda_2 M\left[Z_2, 2\right]\right)\Delta \\
&\qquad\qquad + \left((\nu_{2,2} - \lambda_2^2) M\left[Z_2, 1\right] + \beta_{2,2}\lambda_2 M\left[Z_2\phi_{2,2}(Z_2), 1\right]\right) M\left[Z_2, 1\right]\Delta^2 + o(\Delta^2).
\end{aligned}
$$

The expressions above depend upon expectations, variances and covariances of the stochastic jump intensities. They are in turn given by the following expressions:

$$\begin{aligned}
\lambda_1 :&= \mathbb{E}\left[\lambda_{1,t}\right] = \frac{\alpha_1\lambda_{1,\infty}\left(\alpha_2-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right]\right)+\alpha_2\lambda_{2,\infty}\beta_{1,2}M\left[\phi_{1,2}(Z_2),1\right]}{\left(\alpha_1-\beta_{1,1}M\left[\phi_{1,1}(Z_1),1\right]\right)\left(\alpha_2-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right]\right)-\beta_{2,1}\beta_{1,2}M\left[\phi_{2,1}(Z_1),1\right]M\left[\phi_{1,2}(Z_2),1\right]}\\
\lambda_2 :&= \mathbb{E}\left[\lambda_{2,t}\right] = \frac{\alpha_2\lambda_{2,\infty}\left(\alpha_1-\beta_{1,1}M\left[\phi_{1,1}(Z_1),1\right]\right)+\alpha_1\lambda_{1,\infty}\beta_{2,1}M\left[\phi_{2,1}(Z_1),1\right]}{\left(\alpha_1-\beta_{1,1}M\left[\phi_{1,1}(Z_1),1\right]\right)\left(\alpha_2-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right]\right)-\beta_{2,1}\beta_{1,2}M\left[\phi_{2,1}(Z_1),1\right]M\left[\phi_{1,2}(Z_2),1\right]}\\
\nu_{1,1} :&= \mathbb{E}\left[\lambda_{1,t}^2\right]\\
&= (2\beta_{1,2}M\left[\phi_{1,2}(Z_2),1\right](2(\alpha_2-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right])(-\beta_{1,1}\lambda_1\lambda_2M\left[\phi_{1,1}(Z_1),1\right]\\
&\qquad\qquad + \lambda_1\left(\beta_{1,1}\beta_{2,1}M\left[\phi_{1,1}(Z_1)\phi_{2,1}(Z_1),1\right]-\beta_{2,1}\lambda_1M\left[\phi_{2,1}(Z_1),1\right]+\lambda_2(\alpha_1+\alpha_2-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right])\right)\\
&\qquad\qquad + \lambda_2\left(\beta_{1,2}\beta_{2,2}M\left[\phi_{1,2}(Z_2)\phi_{2,2}(Z_2),1\right]-\beta_{1,2}\lambda_2M\left[\phi_{1,2}(Z_2),1\right]\right))\\
&\qquad\quad + \beta_{1,2}M\left[\phi_{1,2}(Z_2),1\right]\left(\lambda_1\left(\beta_{2,1}^2M\left[\phi_{2,1}(Z_1),2\right]-2\beta_{2,1}\lambda_2M\left[\phi_{2,1}(Z_1),1\right]\right)+\lambda_2\left(2\alpha_2\lambda_2+\beta_{2,2}^2M\left[\phi_{2,2}(Z_2),2\right]-2\beta_{2,2}\lambda_2M\left[\phi_{2,2}(Z_2),1\right]\right)\right))\\
&\quad - \left(\lambda_1\left(2\alpha_1\lambda_1+\beta_{1,1}^2M\left[\phi_{1,1}(Z_1),2\right]-2\beta_{1,1}\lambda_1M\left[\phi_{1,1}(Z_1),1\right]-2\beta_{1,2}\lambda_2M\left[\phi_{1,2}(Z_2),1\right]\right)+\beta_{1,2}^2\lambda_2M\left[\phi_{1,2}(Z_2),2\right]\right)\\
&\qquad \times (2\beta_{2,1}\beta_{1,2}M\left[\phi_{2,1}(Z_1),1\right]M\left[\phi_{1,2}(Z_2),1\right]-2(\alpha_2-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right])(\alpha_1+\alpha_2-\beta_{1,1}M\left[\phi_{1,1}(Z_1),1\right]-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right])))\\
&\quad\times\frac{1}{4(\alpha_1+\alpha_2-\beta_{1,1}M\left[\phi_{1,1}(Z_1),1\right]-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right])((\alpha_1-\beta_{1,1}M\left[\phi_{1,1}(Z_1),1\right])(\alpha_2-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right])-\beta_{2,1}\beta_{1,2}M\left[\phi_{2,1}(Z_1),1\right]M\left[\phi_{1,2}(Z_2),1\right])}\\
\nu_{1,2} :&= \mathbb{E}\left[\lambda_{1,t}\lambda_{2,t}\right]\\
&= \left((\alpha_2-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right])\left(2\alpha_1^2\lambda_1\lambda_2+2\alpha_1\alpha_2\lambda_1\lambda_2-2\beta_{1,1}\lambda_1M\left[\phi_{1,1}(Z_1),1\right](\alpha_1\lambda_2+\beta_{2,1}\lambda_1M\left[\phi_{2,1}(Z_1),1\right])+2\alpha_1\beta_{1,1}\beta_{2,1}\lambda_1M\left[\phi_{1,1}(Z_1)\phi_{2,1}(Z_1),1\right]\right.\right.\\
&\qquad\qquad +\beta_{1,1}^2\beta_{2,1}\lambda_1M\left[\phi_{1,1}(Z_1),2\right]M\left[\phi_{2,1}(Z_1),1\right]-2\beta_{1,2}\lambda_2M\left[\phi_{1,2}(Z_2),1\right](\alpha_1\lambda_2+\beta_{2,1}\lambda_1M\left[\phi_{2,1}(Z_1),1\right])\\
&\qquad\qquad \left.+\beta_{1,2}^2\beta_{2,1}\lambda_2M\left[\phi_{2,1}(Z_1),1\right]M\left[\phi_{1,2}(Z_2),2\right]+2\alpha_1\beta_{1,2}\beta_{2,2}\lambda_2M\left[\phi_{1,2}(Z_2)\phi_{2,2}(Z_2),1\right]-2\alpha_1\beta_{2,2}\lambda_1\lambda_2M\left[\phi_{2,2}(Z_2),1\right]\right)\\
&\quad -\beta_{1,1}M\left[\phi_{1,1}(Z_1),1\right](2(\alpha_2-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right])(\beta_{1,1}\beta_{2,1}\lambda_1M\left[\phi_{1,1}(Z_1)\phi_{2,1}(Z_1),1\right]+\lambda_2(\alpha_1\lambda_1-\beta_{1,1}\lambda_1M\left[\phi_{1,1}(Z_1),1\right]-\beta_{1,2}\lambda_2M\left[\phi_{1,2}(Z_2),1\right])\\
&\qquad\qquad + \lambda_1(\alpha_2\lambda_2-\beta_{2,1}\lambda_1M\left[\phi_{2,1}(Z_1),1\right]-\beta_{2,2}\lambda_2M\left[\phi_{2,2}(Z_2),1\right])+\beta_{1,2}\beta_{2,2}\lambda_2M\left[\phi_{1,2}(Z_2)\phi_{2,2}(Z_2),1\right])\\
&\qquad\quad + \beta_{1,2}M\left[\phi_{1,2}(Z_2),1\right]\left(\beta_{2,1}^2\lambda_1M\left[\phi_{2,1}(Z_1),2\right]-2\beta_{2,1}\lambda_1\lambda_2M\left[\phi_{2,1}(Z_1),1\right]+\lambda_2\left(2\alpha_2\lambda_2+\beta_{22}^2M\left[\phi_{2,2}(Z_2),2\right]-2\beta_{2,2}\lambda_2M\left[\phi_{2,2}(Z_2),1\right]\right)\right))\\
&\quad \left.+\alpha_1\beta_{1,2}M\left[\phi_{1,2}(Z_2),1\right]\left(\beta_{2,1}^2\lambda_1M\left[\phi_{2,1}(Z_1),2\right]-2\beta_{2,1}\lambda_1\lambda_2M\left[\phi_{2,1}(Z_1),1\right]+\lambda_2\left(2\alpha_2\lambda_2+\beta_{2,2}^2M\left[\phi_{2,2}(Z_2),2\right]-2\beta_{2,2}\lambda_2M\left[\phi_{2,2}(Z_2),1\right]\right)\right)\right)\\
&\quad\times\frac{1}{2(\alpha_1+\alpha_2-\beta_{1,1}M\left[\phi_{1,1}(Z_1),1\right]-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right])((\alpha_1-\beta_{1,1}M\left[\phi_{1,1}(Z_1),1\right])(\alpha_2-\beta_{2,2}M\left[\phi_{2,2}(Z_2),1\right])-\beta_{2,1}\beta_{1,2}M\left[\phi_{2,1}(Z_1),1\right]M\left[\phi_{1,2}(Z_2),1\right])}\\
\nu_{2,2} :&= \mathbb{E}\left[\lambda_{2,t}^2\right],
\end{aligned}$$

where $\nu_{2,2}$ is obtained from $\nu_{1,1}$ by interchanging indexes. Its full expression is suppressed to save space. Furthermore, we define $\nu_{2,1} := \nu_{1,2}$.

Theorem S.6. *For the model* (5) *with* $m = 2$*, the third and fourth moments are given in closed form up to order* Δ^2 *by the following expressions:*

$$\begin{aligned}
\mathbb{E}\left[(\Delta Y_{1,t} - \mathbb{E}[\Delta Y_{1,t}])^3\right] &= \lambda_1 M[Z_1, 3]\,\Delta \\
&\quad + \frac{3}{2}\left(\beta_{1,1}\lambda_1 M[Z_1,2]\, M[Z_1\phi_{1,1}(Z_1),1] + M[Z_1,1]\left(2(\nu_{1,1}-\lambda_1^2)M[Z_1,2]\right.\right. \\
&\qquad \left.\left. + \beta_{1,1}\lambda_1 M\left[Z_1^2\phi_{1,1}(Z_1),1\right]\right)\right)\Delta^2 + o(\Delta^2) \\
\mathbb{E}\left[(\Delta Y_{2,t} - \mathbb{E}[\Delta Y_{2,t}])^3\right] &= \lambda_2 M[Z_2, 3]\,\Delta \\
&\quad + \frac{3}{2}\left(\beta_{2,2}\lambda_2 M[Z_2,2]\, M[Z_2\phi_{2,2}(Z_2),1] + M[Z_2,1]\left(2(\nu_{2,2}-\lambda_2^2)M[Z_2,2]\right.\right. \\
&\qquad \left.\left. + \beta_{2,2}\lambda_2 M\left[Z_2^2\phi_{2,2}(Z_2),1\right]\right)\right)\Delta^2 + o(\Delta^2) \\
\mathbb{E}\left[(\Delta Y_{1,t} - \mathbb{E}[\Delta Y_{1,t}])^2(\Delta Y_{2,t} - \mathbb{E}[\Delta Y_{2,t}])\right] &= \frac{1}{2}\left(2(\nu_{1,2}-\lambda_1\lambda_2)M[Z_1,2]\,M[Z_2,1]\right. \\
&\qquad + \beta_{1,2}\lambda_2 M[Z_1,2]\,M[Z_2\phi_{1,2}(Z_2),1] \\
&\qquad \left. + \beta_{2,1}\lambda_1 M[Z_2,1]\,M\left[Z_1^2\phi_{2,1}(Z_1),1\right]\right)\Delta^2 + o(\Delta^2) \\
\mathbb{E}\left[(\Delta Y_{1,t} - \mathbb{E}[\Delta Y_{1,t}])(\Delta Y_{2,t} - \mathbb{E}[\Delta Y_{2,t}])^2\right] &= \frac{1}{2}\left(2(\nu_{1,2}-\lambda_1\lambda_2)M[Z_1,1]\,M[Z_2,2]\right. \\
&\qquad + \beta_{2,1}\lambda_1 M[Z_2,2]\,M[Z_1\phi_{2,1}(Z_1),1] \\
&\qquad \left. + \beta_{1,2}\lambda_2 M[Z_1,1]\,M\left[Z_2^2\phi_{1,2}(Z_2),1\right]\right)\Delta^2 + o(\Delta^2) \\
\mathbb{E}\left[(\Delta Y_{1,t} - \mathbb{E}[\Delta Y_{1,t}])^4\right] &= \lambda_1 M[Z_1, 4]\,\Delta \\
&\quad + \Big(3\sigma_1^4 + 6\lambda_1\sigma_1^2 M[Z_1,2] + 3\nu_{1,1}M[Z_1,2]^2 + 4(\nu_{1,1}-\lambda_1^2)M[Z_1,1]\,M[Z_1,3] \\
&\qquad + 2\beta_{1,1}\lambda_1\left(M[Z_1,3]\,M[Z_1\phi_{1,1}(Z_1),1] + \tfrac{3}{2}M[Z_1,2]\,M\left[Z_1^2\phi_{1,1}(Z_1),1\right]\right. \\
&\qquad \left. + M[Z_1,1]\,M\left[Z_1^3\phi_{1,1}(Z_1),1\right]\right)\Big)\Delta^2 + o(\Delta^2) \\
\mathbb{E}\left[(\Delta Y_{2,t} - \mathbb{E}[\Delta Y_{2,t}])^4\right] &= \lambda_2 M[Z_2, 4]\,\Delta \\
&\quad + \Big(3\sigma_2^4 + 6\lambda_2\sigma_2^2 M[Z_2,2] + 3\nu_{2,2}M[Z_2,2]^2 + 4(\nu_{2,2}-\lambda_2^2)M[Z_2,1]\,M[Z_2,3] \\
&\qquad + 2\beta_{2,2}\lambda_2\left(M[Z_2,3]\,M[Z_2\phi_{2,2}(Z_2),1] + \tfrac{3}{2}M[Z_2,2]\,M\left[Z_2^2\phi_{2,2}(Z_2),1\right]\right. \\
&\qquad \left. + M[Z_2,1]\,M\left[Z_2^3\phi_{2,2}(Z_2),1\right]\right)\Big)\Delta^2 + o(\Delta^2).
\end{aligned}$$

Theorem S.7. *For the model* (5) *with* $m = 2$*, the auto- and cross-covariances of the regular and squared processes are given in closed form up to order* Δ^2 *by the following expressions:*

$$\begin{aligned}
\mathbb{E}[(\Delta Y_{i,t} - \mathbb{E}[\Delta Y_{i,t}])(\Delta Y_{j,t+\Delta} - \mathbb{E}[\Delta Y_{j,t+\Delta}])] &= M[Z_j,1]\left((\nu_{i,j}-\lambda_i\lambda_j)M[Z_i,1]\right. \\
&\quad \left.+\beta_{j,i}\lambda_i M[Z_i\phi_{j,i}(Z_i),1]\right)\Delta^2 \\
&\quad + o(\Delta^2) \\
\mathbb{E}\left[(\Delta Y_{i,t} - \mathbb{E}[\Delta Y_{i,t}])^2(\Delta Y_{j,t+\Delta} - \mathbb{E}[\Delta Y_{j,t+\Delta}])^2\right] &= M[Z_j,2]\left((\nu_{i,j}-\lambda_i\lambda_j)M[Z_i,2]\right. \\
&\quad \left.+\beta_{j,i}\lambda_i M\left[Z_i^2\phi_{j,i}(Z_i),1\right]\right)\Delta^2 \\
&\quad + o(\Delta^2),
\end{aligned}$$

$i, j = 1, 2.$

Theorem S.8. *For the model (5) with $m = 2$, the marginal cgf is given in closed form up to order Δ^2 by the following expression:*

$$
\begin{aligned}
\overline{K}(\Delta, u_1) = &\left(\mu_1 u_1 + \frac{\sigma_1^2 u_1^2}{2} + \lambda_1 \left(L(u_1, Z_1) - 1\right)\right)\Delta \\
&+ \frac{1}{2}\left(\beta_{1,1}\lambda_1 (L(u_1, \phi_{1,1}(Z_1)) - M\left[\phi_{1,1}(Z_1), 1\right]) + (\nu_{1,1} - \lambda_1^2)\left(L(u_1, Z_1) - 1\right)\right)(L(u_1, Z_1) - 1)\Delta^2 \\
&+ o(\Delta^2),
\end{aligned}
$$

where $L(u_1, Z_1) = \mathbb{E}\left[e^{u_1 Z_1}\right]$ *and* $L(u_1, \phi_{1,1}(Z_1)) = \mathbb{E}\left[\phi_{1,1}(Z_1) e^{u_1 Z_1}\right]$.

Theorem S.9. *For the model (5) with $m = 2$, the bivariate cgf is given in closed form up to order Δ^2 by the following expression:*

$$
\begin{aligned}
\overline{K}(\Delta, u_1, u_2) = &\left(\mu_1 u_1 + \mu_2 u_2 + \frac{\sigma_1^2 u_1^2 + \sigma_2^2 u_2^2}{2} + u_1 u_2 \rho \sigma_1 \sigma_2 + \lambda_1 \left(L(u_1, Z_1) - 1\right) + \lambda_2 \left(L(u_2, Z_2) - 1\right)\right)\Delta \\
&+ \frac{1}{2}\Big(\Big(\beta_{1,1}\lambda_1 \left(L(u_1, \phi_{1,1}(Z_1)) - M\left[\phi_{1,1}(Z_1), 1\right]\right) + \beta_{1,2}\lambda_2 \left(L(u_2, \phi_{1,2}(Z_2)) - M\left[\phi_{1,2}(Z_2), 1\right]\right) \\
&+ \left(L(u_1, Z_1) - 1\right)\left(\nu_{1,1} - \lambda_1^2\right) + \left(L(u_2, Z_2) - 1\right)\left(\nu_{1,2} - \lambda_1\lambda_2\right)\Big)\left(L(u_1, Z_1) - 1\right) \\
&+ \Big(\beta_{2,1}\lambda_1 \left(L(u_1, \phi_{2,1}(Z_1)) - M\left[\phi_{2,1}(Z_1), 1\right]\right) + \beta_{2,2}\lambda_2 \left(L(u_2, \phi_{2,2}(Z_2)) - M\left[\phi_{2,2}(Z_2), 1\right]\right) \\
&+ \left(L(u_2, Z_2) - 1\right)\left(\nu_{2,2} - \lambda_2^2\right) + \left(L(u_1, Z_1) - 1\right)\left(\nu_{1,2} - \lambda_1\lambda_2\right)\Big)\left(L(u_2, Z_2) - 1\right)\Big)\Delta^2 \\
&+ o(\Delta^2),
\end{aligned}
$$

where $L(u_i, Z_i) = \mathbb{E}\left[e^{u_i Z_i}\right]$ *and* $L(u_j, \phi_{i,j}(Z_j)) = \mathbb{E}\left[\phi_{i,j}(Z_j) e^{u_j Z_j}\right]$.

C Saddlepoint Approximations: General Expressions

Throughout this and the next sections, we let $X_{i,t,\Delta} := -(Y_{i,t+\Delta} - Y_{i,t})$ denote bank i's sign-changed P&L over a time interval of length Δ. Furthermore, in the univariate case ($m = 1$, omitting the index i), we denote its cumulative distribution function (cdf) by $P(\Delta, x) = \mathbb{P}[X_{t,\Delta} \leq x]$ with associated probability density function (pdf) $p(\Delta, x)$ and decumulative distribution function (ddf) $\overline{P}(\Delta, x)$; and similarly in the bivariate case ($m = 2$, with $i = 1, 2$).

C.1 Univariate Saddlepoint Approximations

Saddlepoint approximations are of great use if the cgf is known analytically but the corresponding pdf is not. We start by considering the cgf of $X_{t,\Delta}$, defined as[3]

$$K(\Delta, u) = \log \mathbb{E}\left[e^{u X_{t,\Delta}}\right], \qquad u \in \mathbb{R}.$$

As is well-known, the cgf is connected to the pdf $p(\Delta, x)$ of $X_{t,\Delta}$ by the Fourier inversion formula

$$p(\Delta, x) = (2\pi \mathrm{i})^{-1} \int_{\hat{u}-\mathrm{i}\infty}^{\hat{u}+\mathrm{i}\infty} \exp\left(K(\Delta, u) - ux\right) \mathrm{d}u. \tag{C.1}$$

The saddlepoint method, dating back to Daniels (1954), suitably chooses the path of integration $\hat{u}$ in (C.1).

Let $\hat{u} = \hat{u}(\Delta, x)$ be the solution of

$$K_1(\Delta, u) = x, \qquad \text{with} \qquad K_1(\Delta, u) = \frac{\partial K(\Delta, u)}{\partial u}. \tag{C.2}$$

Then an approximation to (C.1) can be taken to be (see Section C.3.1)

$$p^{(0)}(\Delta, x) = (2\pi)^{-1/2} \exp\left(K(\Delta, \hat{u}) - \hat{u}x\right) \left(\frac{\partial^2 K(\Delta, \hat{u})}{\partial u^2}\right)^{-1/2}. \tag{C.3}$$

The approximation (C.3) can be seen to feature a Gaussian leading term (see Section C.3.2). To obtain a non-Gaussian leading term — particularly important in our setting with jumps — we let $\hat{w} = \hat{w}(x)$ be the solution of

$$K_{0,1}(\Delta, w) = x, \qquad \text{with} \qquad K_{0,1}(\Delta, w) = \frac{\partial K_0(\Delta, w)}{\partial w}, \tag{C.4}$$

[3] In this and the next sections, we use K to denote the cgf of $\mathbf{X}_{t,\Delta}$, just like in the main text. In Sections A and B, we use $\overline{K}$ to denote the cgf of $-\mathbf{X}_{t,\Delta}$.

for some (non-Gaussian) cgf K_0, with corresponding pdf f_0 and cdf F_0. Then an approximation to (C.1) can be taken to be (see Section C.3.2)[4]

$$p^{(0)}(\Delta, x) = f_0(\Delta, x) \exp\left((K(\Delta, \hat{u}) - \hat{u}x) - (K_0(\Delta, \hat{w}) - \hat{w}x)\right) \times \left(\frac{\partial^2 K_0(\Delta, \hat{w})}{\partial w^2}\right)^{1/2} \left(\frac{\partial^2 K(\Delta, \hat{u})}{\partial u^2}\right)^{-1/2}. \quad \text{(C.5)}$$

The accuracy of this approximation may be optimized further with an alternative, judicious choice of x in (C.4) so that the exponential term in the approximation vanishes; see e.g., Butler (2007), Section 16.1. We do not pursue this here and throughout.

To obtain an approximation to the cdf, following Lugannani & Rice (1980) and Wood et al. (1993), we start by considering the Fourier inversion formula

$$\mathbb{P}[X_{t,\Delta} > x] \equiv \overline{P}(\Delta, x) = (2\pi \mathrm{i})^{-1} \int_{\hat{u}-\mathrm{i}\infty}^{\hat{u}+\mathrm{i}\infty} \exp\left(K(\Delta, u) - ux\right) \frac{\mathrm{d}u}{u}. \quad \text{(C.6)}$$

Then, after a change of variables and an isolation of the singular part (see Section C.3.3), an approximation to (C.6) can be taken to be

$$\begin{aligned} &\overline{P}^{(0)}(\Delta, x) \\ &= \exp\left((K(\Delta, \hat{u}) - \hat{u}x) - (K_0(\Delta, \hat{w}) - \hat{w}x)\right) \\ &\times \left(1 - F_0(\Delta, x) + f_0(\Delta, x)\left\{\frac{1}{\hat{u}}\left(\frac{\partial^2 K_0(\Delta, \hat{w})}{\partial w^2}\right)^{1/2}\left(\frac{\partial^2 K(\Delta, \hat{u})}{\partial u^2}\right)^{-1/2} - \frac{1}{\hat{w}}\right\}\right). \quad \text{(C.7)} \end{aligned}$$

We next consider the problem of deriving saddlepoint approximations to the Expected Shortfall. We start by considering the Fourier inversion formula (Martin 2006)

$$\mathcal{L}(\Delta, x) \equiv \mathbb{E}\left[X_{t,\Delta} \mathbb{1}_{\{X_{t,\Delta} > x\}}\right] = (2\pi \mathrm{i})^{-1} \int_{\hat{u}-\mathrm{i}\infty}^{\hat{u}+\mathrm{i}\infty} \exp\left(K(\Delta, u) - ux\right) K_1(\Delta, u) \frac{\mathrm{d}u}{u}. \quad \text{(C.8)}$$

[4]Saddlepoint approximations may alternatively be derived by first exponentially tilting the pdf to obtain a family of densities constituting a regular exponential family, and next expanding it by computing an Edgeworth approximation. This alternative approach was initiated by Esscher (1932) who argued that indirect Edgeworth approximation, by first exponentially tilting the density and next approximating it for a well-chosen point of approximation, achieves much higher accuracy especially in the tails of the distribution. In particular, the saddlepoint approximation for the pdf agrees with Esscher's approach given by the indirect Edgeworth expansion; see Section 5.3 of Butler (2007) for details.

To isolate the singular part, we write

$$\mathcal{L}(\Delta,x) = (2\pi\mathrm{i})^{-1}\int_{\hat{u}-\mathrm{i}\infty}^{\hat{u}+\mathrm{i}\infty} \exp\left(K(\Delta,u)-ux\right)K_1(\Delta,0)\frac{\mathrm{d}u}{u} + (2\pi\mathrm{i})^{-1}\int_{\hat{u}-\mathrm{i}\infty}^{\hat{u}+\mathrm{i}\infty} \exp\left(K(\Delta,u)-ux\right)\left(K_1(\Delta,u)-K_1(\Delta,0)\right)\frac{\mathrm{d}u}{u},$$

which gives rise to the following approximation:

$$\mathcal{L}^{(0)}(\Delta,x) = K_1(\Delta,0)\overline{P}^{(0)}(\Delta,x) + \frac{x-K_1(\Delta,0)}{\hat{u}}p^{(0)}(\Delta,x). \tag{C.9}$$

Notice that $K_1(\Delta,0)=\mathbb{E}[X_{t,\Delta}]$.

An appropriate (non-Gaussian) base in the setting of our Hawkes pure jump model with $X_{t,\Delta}=\int_\Delta Z_t \mathrm{d}N_t$, taking for ease of exposition Z to be the sign-changed counterpart of the jump magnitude Z in the main text and in Sections A–B, is

$$F_0(\Delta,x) = (1-\lambda\Delta)+\lambda\Delta F_Z(x), \qquad x\geq 0,$$

which can be viewed as the cdf of the terminal value of a process that has at most one jump, with cdf F_Z, over an interval of length Δ, with jump probability $\lambda\Delta$. It is henceforth referred to as the *Bernoulli base*. In this case,

$$K_0(\Delta,w) = \log(1-\lambda\Delta+\lambda\Delta\, L(w,Z)),$$

where $L(\cdot,Z)$ is the moment generating function of Z. With exponential jump size distribution $F_Z(x)=1-\exp(-\gamma x)$, $\gamma>0$, hence $L(u,Z)=\mathbb{E}\left[e^{uZ}\right]=\frac{\gamma}{\gamma-u}$, $u<\gamma$, and $M[Z,n]=\mathbb{E}[Z^n]=\frac{n!}{\gamma^n}$, $n=1,2,\ldots$, this specializes to

$$\begin{cases} F_0(\Delta,x) &= 1-\lambda\Delta\exp(-\gamma x);\\ f_0(\Delta,x) &= \lambda\Delta\gamma\exp(-\gamma x);\\ K_0(\Delta,w) &= \log\left(1-\lambda\Delta+\lambda\Delta\frac{\gamma}{\gamma-w}\right);\\ \hat{w}(\Delta,x) &= \frac{2\gamma x-\lambda\Delta\gamma x-\sqrt{\lambda\Delta\gamma x}\sqrt{4-4\lambda\Delta+\lambda\Delta\gamma x}}{2(x-\lambda\Delta x)}. \end{cases}$$

Notice the explicitness of these formulae.

The final step in obtaining the saddlepoint approximations, is to specify $K(\Delta,u)$. In the case of a Poissonian jump model,

$$K_{\mathrm{Poi}}(\Delta,u) = \lambda\Delta\left(L(u,Z)-1\right).$$

Hence, with exponential jump size distribution, we obtain

$$K_{\mathrm{Poi}}(\Delta, u) = \lambda\Delta\left(\frac{\gamma}{\gamma - u} - 1\right).$$

Notice that $K_{\mathrm{Poi}}(\Delta, u)$ and $K_0(\Delta, u)$ are equal at the leading order in Δ. While the cgf $K(\Delta, u)$ for the (affine) Hawkes jump model considered here may be obtained by numerically solving two ODE's, see Duffie et al. (2000) and Errais et al. (2010), to maintain explicitness, we rather use the expansions in Δ of the cgf $K(\Delta, u)$ derived explicitly in Section B of this Supplement.

C.2 Bivariate Saddlepoint Approximations

The cgf of the bivariate vector $\mathbf{X}_{t,\Delta} = (X_{1,t,\Delta}, X_{2,t,\Delta})$ is given by

$$K(\Delta, u_1, u_2) = \log \mathbb{E}\left[\exp\left(u_1 X_{1,t,\Delta} + u_2 X_{2,t,\Delta}\right)\right], \qquad \mathbf{u} = (u_1, u_2) \in \mathbb{R}^2.$$

We suppose that the set $\Lambda := \{(u_1, u_2) | K(\Delta, u_1, u_2) < \infty\}$ takes the product form $\Lambda = \Lambda_1\Lambda_2$ where $\Lambda_i = [-c_i, d_i]$, $c_i, d_i \geq 0$, $c_i + d_i > 0$, $i = 1, 2$, so the origin $(0, 0)$ is included in Λ and the intervals are non-empty.

We start with the Fourier inversion formula for the pdf $p(\Delta, x_1, x_2)$ of $(X_{1,t,\Delta}, X_{2,t,\Delta})$:

$$p(\Delta, x_1, x_2) = (2\pi \mathrm{i})^{-2} \int_{\hat{u}_1 - \mathrm{i}\infty}^{\hat{u}_1 + \mathrm{i}\infty} \int_{\hat{u}_2 - \mathrm{i}\infty}^{\hat{u}_2 + \mathrm{i}\infty} \exp\left(K(\Delta, u_1, u_2) - (u_1 x_1 + u_2 x_2)\right) \mathrm{d}u_1 \mathrm{d}u_2. \quad \text{(C.10)}$$

The vector $(\hat{u}_1, \hat{u}_2) \in \mathbb{R}^2$ in the limits of integration is chosen so as to ensure that on the path of integration $-c_1 < \hat{u}_1 < d_1$ and $-c_2 < \hat{u}_2 < d_2$ so $-c_1 < \mathrm{Re}(u_1) < d_1$ and $-c_2 < \mathrm{Re}(u_2) < d_2$ is satisfied. An appropriate choice of $(\hat{u}_1, \hat{u}_2)$ dictating the path of integration is the key to the saddlepoint approximation.

We determine $(\hat{u}_1, \hat{u}_2) = (\hat{u}_1(x_1, x_2), \hat{u}_2(x_1, x_2))$ as the solution to the equations

$$\frac{\partial K(\Delta, u_1, u_2)}{\partial u_1} = x_1, \qquad \frac{\partial K(\Delta, u_1, u_2)}{\partial u_2} = x_2. \quad \text{(C.11)}$$

Note that $K(\Delta, u_1, u_2)$ is convex in both variables ensuring the existence of a solution to this set of two equations.

We also determine $\hat{u}_0 = \hat{u}_0(x_1)$ as the solution to the equation

$$\frac{\partial K(\Delta, u_1, 0)}{\partial u_1} = x_1, \quad \text{(C.12)}$$

and $\hat{u}_{u_2} = \hat{u}_{u_2}(x_1)$ as the solution to the equation

$$\frac{\partial K(\Delta, u_1, u_2)}{\partial u_1} = x_1. \tag{C.13}$$

Similarly, for the base cgf $K_0(\Delta, w)$, we consider the associated saddlepoint $\hat{w} = \hat{w}(x_2)$ as the solution to the equation

$$K_{0,1}(\Delta, w) = x_2. \tag{C.14}$$

Following Section 16.3 in Butler (2007), this leads to the following saddlepoint approximation for the conditional pdf:

$$\begin{aligned} p^{(0)}(\Delta, x_2|x_1) = f_0(\Delta, x_2) e^{\left(K(\Delta, \hat{u}_{\hat{u}_2}, \hat{u}_2) - \hat{u}_{\hat{u}_2} x_1\right) - (K(\Delta, \hat{u}_0, 0) - \hat{u}_0 x_1) - \hat{u}_2 x_2 - (K_0(\Delta, \hat{w}) - \hat{w} x_2)} \\ \times \left(\frac{\partial^2 K_0(\Delta, \hat{w})}{\partial w^2}\right)^{1/2} \left(\frac{\partial^2 K(\Delta, \hat{u}_0, 0)}{\partial u_1^2}\right)^{1/2} \left(|K''(\Delta, \hat{u}_1, \hat{u}_2)|\right)^{-1/2}, \end{aligned} \tag{C.15}$$

where K'' denotes the Hessian of K and $|\cdot|$ denotes the determinant. The accuracy of this approximation may again be optimized further with an alternative, judicious choice of x_2 in (C.14) so that the exponential term in the approximation vanishes. We do not pursue this here and throughout.

Furthermore, we have the following saddlepoint approximation for the conditional ddf:

$$\begin{aligned} &\overline{P}^{(0)}(\Delta, x_2|x_1) \\ &= \exp\left((K(\Delta, \hat{u}_{\hat{u}_2}, \hat{u}_2) - \hat{u}_{\hat{u}_2} x_1) - (K(\Delta, \hat{u}_0, 0) - \hat{u}_0 x_1) - \hat{u}_2 x_2 - (K_0(\Delta, \hat{w}) - \hat{w} x_2)\right) \\ &\quad \times \Bigg(1 - F_0(\Delta, x_2) + f_0(\Delta, x_2) \\ &\qquad \times \left\{\frac{1}{\hat{u}_2}\left(\frac{\partial^2 K_0(\Delta, \hat{w})}{\partial w^2}\right)^{1/2} \left(\frac{\partial^2 K(\Delta, \hat{u}_0, 0)}{\partial u_1^2}\right)^{1/2} \left(|K''(\Delta, \hat{u}_1, \hat{u}_2)|\right)^{-1/2} - \frac{1}{\hat{w}}\right\}\Bigg). \end{aligned} \tag{C.16}$$

In view of the generic identity

$$f(x_1, x_2) = f(x_2|x_1) f(x_1),$$

which expresses the bivariate probability density $f(x_1, x_2)$ as the product of the conditional density of X_2 evaluated at x_2 given $X_1 = x_1$ and the probability density of X_1 at x_1, we have the following saddlepoint approximation for our bivariate pdf:

$$
\begin{aligned}
&p^{(0)}(\Delta, x_1, x_2) \\
&= f_0(\Delta, x_1) \exp\left((K(\Delta, \hat{u}_1(x_1)) - \hat{u}_1(x_1)x_1) - (K_0(\Delta, \hat{w}(x_1)) - \hat{w}(x_1)x_1)\right) \\
&\quad \times \left(\frac{\partial^2 K_0(\Delta, \hat{w}(x_1))}{\partial w^2}\right)^{1/2} \left(\frac{\partial^2 K(\Delta, \hat{u}_1(x_1))}{\partial u_1^2}\right)^{-1/2} \\
&\quad \times f_0(\Delta, x_2) e^{\left(K(\Delta, \hat{u}_{\hat{u}_2}, \hat{u}_2) - \hat{u}_{\hat{u}_2} x_1\right) - (K(\Delta, \hat{u}_0, 0) - \hat{u}_0 x_1) - \hat{u}_2 x_2 - (K_0(\Delta, \hat{w}) - \hat{w} x_2)} \\
&\quad \times \left(\frac{\partial^2 K_0(\Delta, \hat{w})}{\partial w^2}\right)^{1/2} \left(\frac{\partial^2 K(\Delta, \hat{u}_0, 0)}{\partial u_1^2}\right)^{1/2} \left(|K''(\Delta, \hat{u}_1, \hat{u}_2)|\right)^{-1/2}. \qquad \text{(C.17)}
\end{aligned}
$$

Here, $\hat{u}_1(x_1)$ and $\hat{w}(x_1)$ are the saddlepoints of the first marginal cgf's K and K_0 as a function of x_1 (not x_2).

C.3 Derivations

C.3.1 Derivation of (C.3)

Note that the function

$$u \mapsto K(\Delta, u) - ux,$$

is minimal at $u = \hat{u}$ (observe that K is convex in u, which guarantees existence and uniqueness of this minimum), and a second-order Taylor series expansion around $u = \hat{u}$ given by

$$K(\Delta, u) - ux = K(\Delta, \hat{u}) - \hat{u}x + \frac{1}{2}(u - \hat{u})^2 \left(\frac{\partial^2 K(\Delta, \hat{u})}{\partial u^2}\right) + O(|u - \hat{u}|^3),$$

leads to

$$K(\Delta, u) - ux = K(\Delta, \hat{u}) - \hat{u}x + \frac{1}{2}v^2 \left(\frac{\partial^2 K(\Delta, \hat{u})}{\partial u^2}\right) + O(|v|^3),$$

using $u = \hat{u} + \mathrm{i}v$, $v \in \mathbb{R}$, on the path of integration in (C.1). Therefore, upon suitably substituting this Taylor series expansion up to the second order in (C.1) one obtains the following expression, which one may take as the leading term in an approximation to (C.1):

$$p^{(0)}(\Delta, x) = (2\pi)^{-1} \exp\left(K(\Delta, \hat{u}) - \hat{u}x\right) \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}v^2 \frac{\partial^2 K(\Delta, \hat{u})}{\partial u^2}\right) \mathrm{d}v.$$

The right-hand side in the expression above readily simplifies to the right-hand side of (C.3). If desired, one may normalize to unity the integral over x of the right-hand side of (C.3). This is simply achieved by dividing the right-hand side of (C.3) by the integral over

x.

C.3.2 Derivation of (C.5)

Rather than Tayloring $u \mapsto K(\Delta, u) - ux$ around $\hat{u}$, one may consider another ("base") cgf K_0, analytic at 0, and approximate the quadratic behavior of $u \mapsto K(\Delta, u) - ux$ in the neighborhood of $\hat{u}$ by invoking the function

$$w \mapsto K_0(\Delta, w) - wx_0,$$

with x_0 a fixed point inducing $\hat{w}$ through $K_{0,1}(\Delta, \hat{w}) = x_0$, such that

$$(K(\Delta, u) - ux) - (K(\Delta, \hat{u}) - \hat{u}x) = (K_0(\Delta, w) - wx_0) - (K_0(\Delta, \hat{w}) - \hat{w}x_0), \qquad \text{(C.18)}$$

which implicitly dictates the change of variable $u(w)$. Now take $x_0 = x$ for a given x. Then, with the change of variable $u(w)$,

$$\begin{aligned} p(\Delta, x) = (2\pi \mathrm{i})^{-1} &\exp\left((K(\Delta, \hat{u}) - \hat{u}x) - (K_0(\Delta, \hat{w}) - \hat{w}x)\right) \\ &\times \int_{\hat{w}-\mathrm{i}\infty}^{\hat{w}+\mathrm{i}\infty} \exp\left(K_0(\Delta, w) - wx\right) \frac{\mathrm{d}u(w)}{\mathrm{d}w} \mathrm{d}w, \end{aligned} \qquad \text{(C.19)}$$

and an approximation at the leading order is given by

$$\begin{aligned} p^{(0)}(\Delta, x) = (2\pi \mathrm{i})^{-1} &\exp\left((K(\Delta, \hat{u}) - \hat{u}x) - (K_0(\Delta, \hat{w}) - \hat{w}x)\right) \\ &\times \int_{\hat{w}-\mathrm{i}\infty}^{\hat{w}+\mathrm{i}\infty} \exp\left(K_0(\Delta, w) - wx\right) \mathrm{d}w \frac{\mathrm{d}u(\hat{w})}{\mathrm{d}w}, \end{aligned} \qquad \text{(C.20)}$$

which reduces to (C.5), upon realizing from (C.18) with $x_0 = x$ that

$$\frac{\mathrm{d}u(\hat{w})}{\mathrm{d}w} = \left(\frac{\partial^2 K_0(\Delta, \hat{w})}{\partial w^2}\right)^{1/2} \left(\frac{\partial^2 K(\Delta, \hat{u})}{\partial u^2}\right)^{-1/2}.$$

When $K_0(\Delta, w) = \frac{w^2}{2}$, which corresponds to a standard Gaussian distribution, (C.3) occurs.

C.3.3 Derivation of (C.7)

Invoking the change of variable $u(w)$ defined in (C.18) with $x_0 = x$ yields

$$\overline{P}(\Delta, x) = (2\pi \mathrm{i})^{-1} \exp\left((K(\Delta, \hat{u}) - \hat{u}x) - (K_0(\Delta, \hat{w}) - \hat{w}x)\right) \times \int_{\hat{w}-\mathrm{i}\infty}^{\hat{w}+\mathrm{i}\infty} \exp\left(K_0(\Delta, w) - wx\right) \left(\frac{1}{u(w)} \frac{\mathrm{d}u(w)}{\mathrm{d}w}\right) \mathrm{d}w.$$

We isolate the singular part, as follows:

$$\begin{aligned}\overline{P}(\Delta, x) &= (2\pi \mathrm{i})^{-1} \exp\left((K(\Delta, \hat{u}) - \hat{u}x) - (K_0(\Delta, \hat{w}) - \hat{w}x)\right) \\ &\quad \times \left(\int_{\hat{w}-\mathrm{i}\infty}^{\hat{w}+\mathrm{i}\infty} \exp\left(K_0(\Delta, w) - wx\right) \frac{\mathrm{d}w}{w} \right. \\ &\quad \left. + \int_{\hat{w}-\mathrm{i}\infty}^{\hat{w}+\mathrm{i}\infty} \exp\left(K_0(\Delta, w) - wx\right) \left(\frac{1}{u(w)} \frac{\mathrm{d}u(w)}{\mathrm{d}w} - \frac{1}{w}\right) \mathrm{d}w \right).\end{aligned}$$

Hence,

$$\begin{aligned}&\overline{P}(\Delta, x) \\ &= \exp\left((K(\Delta, \hat{u}) - \hat{u}x) - (K_0(\Delta, \hat{w}) - \hat{w}x)\right) \\ &\quad \times \left(1 - F_0(\Delta, x) + (2\pi \mathrm{i})^{-1} \int_{\hat{w}-\mathrm{i}\infty}^{\hat{w}+\mathrm{i}\infty} \exp\left(K_0(\Delta, w) - wx\right) \left(\frac{1}{u(w)} \frac{\mathrm{d}u(w)}{\mathrm{d}w} - \frac{1}{w}\right) \mathrm{d}w \right).\end{aligned}$$

Finally, invoking a Taylor series expansion of $\frac{1}{u(w)} \frac{\mathrm{d}u(w)}{\mathrm{d}w} - \frac{1}{w}$ around $\hat{w}$,

$$\frac{1}{\hat{u}} \left(\frac{\partial^2 K_0(\Delta, \hat{w})}{\partial w^2}\right)^{1/2} \left(\frac{\partial^2 K(\Delta, \hat{u})}{\partial u^2}\right)^{-1/2} - \frac{1}{\hat{w}} + O(|w - \hat{w}|),$$

yields

$$\begin{aligned}&\overline{P}^{(0)}(\Delta, x) \\ &= \exp\left((K(\Delta, \hat{u}) - \hat{u}x) - (K_0(\Delta, \hat{w}) - \hat{w}x)\right) \\ &\quad \times \left(1 - F_0(\Delta, x) + f_0(\Delta, x) \left\{\frac{1}{\hat{u}} \left(\frac{\partial^2 K_0(\Delta, \hat{w})}{\partial w^2}\right)^{1/2} \left(\frac{\partial^2 K(\Delta, \hat{u})}{\partial u^2}\right)^{-1/2} - \frac{1}{\hat{w}}\right\}\right). \quad \text{(C.21)}\end{aligned}$$

D Saddlepoint Approximations: Explicit Expressions and Proofs

In this section, we first provide detailed technical derivations of our saddlepoint approximations for the univariate case. Next, we provide detailed technical derivations of our saddlepoint approximations for the marginal tail in the bivariate model and for the bivariate case.

D.1 The Univariate Model

Let us first assume that $\mu = \sigma = 0$ and $\phi(Z) \equiv 1$. Then, we state the following results:

Lemma S.1. *The saddlepoint approximation to the pdf of the Hawkes jump model* (3), (5)–(7) *with $m = 1$ and $p = 1$ is given at order Δ^2 by the following expression:*

$$
\begin{aligned}
&p^{(0)}(\Delta, x) \\
&= \lambda\Delta\gamma \exp(-\gamma x) \\
&\quad \times \exp\left(\frac{\gamma x(2\alpha(\beta+\lambda) - \beta(\beta+2\lambda))}{4(\alpha-\beta)}\Delta\right) \\
&\quad \times \left(1 + \frac{1}{64}\left(-\frac{8\beta(2\alpha-\beta)}{\alpha-\beta} + \frac{3\beta^2\gamma(2\alpha-\beta)^2 x}{\lambda(\alpha-\beta)^2} + 4\lambda(\gamma x - 4)\right)\Delta\right) \\
&\quad + o(\Delta^2).
\end{aligned}
$$

Lemma S.2. *The saddlepoint approximation to the pdf of the Hawkes jump model given in Lemma S.1 satisfies the relative error property:*

$$
\frac{p(\Delta, x)}{p^{(0)}(\Delta, x)} = 1 + O(\Delta^{1/2}).
$$

Theorem S.10. *The saddlepoint approximation to the ddf of the Hawkes jump model* (3), (5)–(7) *with* $m = 1$ *and* $p = 1$ *is given at order* Δ^2 *by the following expression:*

$$\begin{aligned}
&\overline{P}^{(0)}(\Delta, x) \\
&= \exp\left(\frac{\gamma x(2\alpha(\beta+\lambda) - \beta(\beta+2\lambda))}{4(\alpha-\beta)}\Delta\right) \\
&\quad \times \Bigg(\lambda\Delta\exp(-\gamma x) + \lambda\Delta\gamma\exp(-\gamma x) \\
&\quad \times \frac{8\beta\lambda\left(2\alpha^2 - 3\alpha\beta + \beta^2\right) + 3\beta^2\gamma(2\alpha-\beta)^2 x + 4\lambda^2(\alpha-\beta)^2(\gamma x + 4)}{64\gamma\lambda(\alpha-\beta)^2}\Delta\Bigg) \\
&\quad + o(\Delta^2).
\end{aligned}$$

Corollary S.1. *The saddlepoint approximation to* $\mathcal{L}(\Delta, x)$ *defined in* (10) *in the univariate Hawkes jump model of Theorem S.10 takes the following form:*

$$\mathcal{L}^{(0)}(\Delta, x) = \frac{\lambda}{\gamma}\Delta\overline{P}^{(0)}(\Delta, x) + \frac{x - (\lambda/\gamma)\Delta}{\hat{u}}p^{(0)}(\Delta, x),$$

with $\overline{P}^{(0)}(\Delta, x)$, $\hat{u}$ *and* $p^{(0)}(\Delta, x)$ *given by Theorem S.10, Equation* (D.1) *and Lemma S.1, respectively.*

D.1.1 Proof of Lemma S.1

We start with an expansion of the cgf of the univariate Hawkes jump-diffusion model, up to the second order in Δ (see Theorem S.3 in this Supplement):

$$\begin{aligned}
&\overline{K}(\Delta, u) \\
&= \left(\mu u + \frac{\sigma^2 u^2}{2} + \lambda\left(L(u, Z) - 1\right)\right)\Delta \\
&+ \frac{\beta\lambda(L(u,Z)-1)\left(2L(u,\phi(Z))(\alpha - \beta M\left[\phi(Z),1\right]) + \beta M\left[\phi(Z),2\right](L(u,Z)-1) - 2\alpha M\left[\phi(Z),1\right] + 2\beta M\left[\phi(Z),1\right]^2\right)}{4(\alpha - \beta M\left[\phi(Z),1\right])}\Delta^2 \\
&+ o(\Delta^2),
\end{aligned}$$

where $L(u, Z) = \mathbb{E}\left[e^{uZ}\right]$ and $L(u, \phi(Z)) = \mathbb{E}\left[\phi(Z)e^{uZ}\right]$. When restricting attention to the special case in which $\mu = \sigma = 0$ and $\phi(Z) \equiv 1$, the expression for the cgf reduces to

$$\overline{K}(\Delta, u) = \lambda(L(u, Z) - 1)\Delta + \frac{\beta\lambda(2\alpha - \beta)(L(u, Z) - 1)^2}{4(\alpha - \beta)}\Delta^2 + o(\Delta^2).$$

With (one-sided, positive) exponential jump size distribution $F_Z(x) = 1 - \exp(-\gamma x)$, $\gamma > 0$, hence $L(u, Z) = \frac{\gamma}{\gamma - u}$, $u < \gamma$, taking for ease of exposition Z to be the sign-changed counterpart of the jump magnitude Z in the main text and in Sections A–B, this specializes to

$$K(\Delta, u) = \frac{\lambda u}{\gamma - u}\Delta + \frac{\beta\lambda u^2(2\alpha - \beta)}{4(\alpha - \beta)(u - \gamma)^2}\Delta^2 + o(\Delta^2).$$

Henceforth, we omit terms of $o(\Delta^2)$ in $K(\Delta, u)$.

We next compute a Taylor series expansion of the saddlepoint $\hat{u} = \hat{u}(\Delta, x)$ solving

$$K_1(\Delta, u) = x, \qquad \text{with} \qquad K_1(\Delta, u) = \frac{\partial K(\Delta, u)}{\partial u}.$$

It is given by (the coefficient of $\sqrt{\Delta}$ in the expansion below is chosen such that it induces $\hat{u} < \gamma$)

$$\begin{aligned} \hat{u} = \gamma &- \sqrt{\frac{\gamma\lambda}{x}}\sqrt{\Delta} - \frac{\beta\gamma(2\alpha - \beta)}{4(\alpha - \beta)}\Delta \\ &- \frac{\beta\sqrt{\gamma}(2\alpha - \beta)\left(-3\beta\gamma x(2\alpha - \beta) - 8\lambda(\alpha - \beta)\right)}{32\sqrt{\lambda}(\alpha - \beta)^2\sqrt{x}}\Delta^{3/2} + O(\Delta^2). \end{aligned} \tag{D.1}$$

We consider a Bernoulli base. Indeed, an appropriate (non-Gaussian) base in the setting of our Hawkes pure jump model is given by

$$F_0(\Delta, x) = (1 - \lambda\Delta) + \lambda\Delta F_Z(x), \qquad x \geq 0,$$

which can be viewed as the cdf of the terminal value of a process that has at most one jump, with cdf F_Z, over an interval of length Δ, with jump probability $\lambda\Delta$. The Bernoulli base strikes a balance between accuracy and simplicity of the f_0 and F_0 functions that appear in the saddlepoint approximations. Other bases (such as a triangular jump model or a Poissonian jump model) may be considered at the expense of more complex f_0 and F_0 functions. In this case,

$$K_0(\Delta, w) = \log\left(1 - \lambda\Delta + \lambda\Delta\, L(w, Z)\right).$$

With (one-sided, positive) exponential jump size distribution this specializes to

$$\begin{aligned} F_0(\Delta, x) &= 1 - \lambda\Delta\exp(-\gamma x), \quad f_0(\Delta, x) = \lambda\Delta\gamma\exp(-\gamma x), \\ K_0(\Delta, w) &= \log\left(1 - \lambda\Delta + \lambda\Delta\frac{\gamma}{\gamma - w}\right). \end{aligned}$$

Note that K_0 and K are equal up to order Δ.

We next compute the exact as well as a Taylor series expansion of the saddlepoint $\hat{w} = \hat{w}(\Delta, x)$, solving

$$K_{0,1}(\Delta, w) = x, \qquad \text{with} \qquad K_{0,1}(\Delta, w) = \frac{\partial K_0(\Delta, w)}{\partial w}.$$

They are given by the following expressions, respectively:

$$\hat{w} = \frac{\gamma(2 - \lambda\Delta)x - \sqrt{\gamma\lambda\Delta}\sqrt{x(\lambda\Delta(\gamma x - 4) + 4)}}{2(1 - \lambda\Delta)x},$$

and

$$\hat{w} = \gamma - \sqrt{\frac{\gamma\lambda}{x}}\sqrt{\Delta} + \frac{\gamma\lambda}{2}\Delta + \frac{\left(-4\sqrt{\gamma}\lambda^{3/2} - \gamma^{3/2}\lambda^{3/2}x\right)}{8\sqrt{x}}\Delta^{3/2} + O(\Delta^2). \tag{D.2}$$

For consistency with our treatment of $\hat{u}$, we will henceforth use in our computations the latter Taylor expansion of $\hat{w}$ instead of the former exact expression for $\hat{w}$. Note that $\hat{u}$ and $\hat{w}$ agree up to order $\Delta^{1/2}$.

We next compute an expansion up to the first order in Δ of $(K(\Delta, \hat{u}) - \hat{u}x) - (K_0(\Delta, \hat{w}) - \hat{w}x)$. It is given by

$$\begin{aligned}(K(\Delta, \hat{u}) - \hat{u}x) - (K_0(\Delta, \hat{w}) - \hat{w}x) = {} & \frac{\gamma x(2\alpha(\beta + \lambda) - \beta(\beta + 2\lambda))}{4(\alpha - \beta)}\Delta \\ & + O(\Delta^{3/2}).\end{aligned}$$

We finally compute $\left(\frac{\partial^2 K_0(\Delta,\hat{w})}{\partial w^2}\right)^{1/2}\left(\frac{\partial^2 K(\Delta,\hat{u})}{\partial u^2}\right)^{-1/2}$ up to the first order in Δ. It is given by

$$\begin{aligned}&\left(\frac{\partial^2 K_0(\Delta, \hat{w})}{\partial w^2}\right)^{1/2}\left(\frac{\partial^2 K(\Delta, \hat{u})}{\partial u^2}\right)^{-1/2} \\ &= 1 + \frac{1}{64}\left(-\frac{8\beta(2\alpha - \beta)}{\alpha - \beta} + \frac{3\beta^2\gamma(2\alpha - \beta)^2 x}{\lambda(\alpha - \beta)^2} + 4\lambda(\gamma x - 4)\right)\Delta + O(\Delta^{3/2}).\end{aligned}$$

Then the stated result follows from (C.5) upon collecting the terms we derived above.

D.1.2 Proof of Lemma S.2

To prove the relative error property, we compute the higher-order terms in the saddlepoint expansion. Before proving the result, it is instructive to first consider a Gaussian base and show that the relative error property does not hold if one uses a Gaussian base. In particular, the Gaussian base is useful to demonstrate the computation of the higher order terms.

So consider the saddlepoint expansion with a Gaussian base, using the change of variable from u to w in

$$(K(\Delta,u)-ux)-(K(\Delta,\hat{u})-\hat{u}x)=\frac{1}{2}\left(w-\hat{w}\right)^2,$$

where $\hat{w}$ is chosen such that

$$K(\Delta,\hat{u})-\hat{u}x=-\frac{1}{2}\hat{w}^2,$$

so

$$K(\Delta,u)-ux=\frac{1}{2}w^2-w\hat{w}, \tag{D.3}$$

and the minima on the LHS and RHS of (D.3) agree.

By repeatedly differentiating both sides of (D.3), with $K^{(n)}(\Delta,\cdot)$ and $u^{(n)}(\cdot)$ shorthand notation for the n-th order derivative of $K(\Delta,\cdot)$ and $u(\cdot)$, respectively, we obtain:

$$\begin{aligned}
&K'\big(\Delta,u(w)\big)u'(w)-u'(w)x=w-\hat{w},\\
\Rightarrow\ &K^{(2)}\big(\Delta,u(w)\big)\Big(u'(w)\Big)^2+K'\big(\Delta,u(w)\big)u^{(2)}(w)-u^{(2)}(w)x=1,\\
\Rightarrow\ &K^{(3)}\big(\Delta,u(w)\big)\Big(u'(w)\Big)^3+3K^{(2)}\big(\Delta,u(w)\big)u'(w)u^{(2)}(w)\\
&\qquad+K'\big(\Delta,u(w)\big)u^{(3)}(w)-u^{(3)}(w)x=0,\\
\Rightarrow\ &K^{(4)}(\Delta,u(w))\Big(u'(w)\Big)^4+6K^{(3)}(\Delta,u(w))\Big(u'(w)\Big)^2u^{(2)}(w)+3K^{(2)}(\Delta,u(w))\Big(u^{(2)}(w)\Big)^2\\
&\qquad+4K^{(2)}(\Delta,u(w))u'(w)u^{(3)}(w)+K'(\Delta,u(w))u^{(4)}(w)-u^{(4)}(w)x=0.
\end{aligned}$$

By evaluating the last two equations at $w=\hat{w}$ and recalling that $K'(\Delta,\hat{u})=x$, $u(\hat{w})=\hat{u}$ and $u'(\hat{w})=(K^{(2)}(\hat{u}))^{-1/2}$, we obtain:

$$u^{(2)}(\hat{w})=\frac{-K^{(3)}(\Delta,\hat{u})\big(u'(\hat{w})\big)^2}{3K^{(2)}(\Delta,\hat{u})}=\frac{-K^{(3)}(\Delta,\hat{u})}{3\big(K^{(2)}(\Delta,\hat{u})\big)^2}, \tag{D.4}$$

$$u^{(3)}(\hat{w})=-\frac{K^{(4)}(\Delta,\hat{u})\big(u'(\hat{w})\big)^4+6K^{(3)}(\Delta,\hat{u})\big(u'(\hat{w})\big)^2u^{(2)}(\hat{w})+3K^{(2)}(\Delta,\hat{u})\big(u^{(2)}(\hat{w})\big)^2}{4K^{(2)}(\Delta,\hat{u})u'(\hat{w})}. \tag{D.5}$$

Note that, with the change of variable $u(w)$,

$$p(\Delta, x) = (2\pi\mathrm{i})^{-1} \int_{\hat{u}-\mathrm{i}\infty}^{\hat{u}+\mathrm{i}\infty} \exp\left(K(\Delta, u) - ux\right) \mathrm{d}u$$
$$= (2\pi\mathrm{i})^{-1} \int_{\hat{w}-\mathrm{i}\infty}^{\hat{w}+\mathrm{i}\infty} \exp\left(\frac{1}{2}w^2 - w\hat{w}\right) \frac{\mathrm{d}u(w)}{\mathrm{d}w} \mathrm{d}w.$$

Next, we approximate $\frac{\mathrm{d}u(w)}{\mathrm{d}w}$ near the saddlepoint $w = \hat{w}$ by

$$u'(w) \approx u'(\hat{w}) + u^{(2)}(\hat{w})(w - \hat{w}) + \frac{1}{2}u^{(3)}(\hat{w})(w - \hat{w})^2.$$

Specifically, as $w = \hat{w} + \mathrm{i}v$, $v \in \mathbb{R}$, over the integral path, we have $u'(w) \approx u'(\hat{w}) + \mathrm{i} \cdot u^{(2)}(\hat{w})v - \frac{1}{2}u^{(3)}(\hat{w})v^2$. This yields

$$p^{(1)}(\Delta, x) = \underbrace{(2\pi)^{-1} \int_{-\infty}^{\infty} \exp\left(\frac{w^2}{2} - w\hat{w}\right) u'(\hat{w})\mathrm{d}v}_{\text{(I)}} + \underbrace{\mathrm{i} \cdot (2\pi)^{-1} \int_{-\infty}^{\infty} \exp\left(\frac{w^2}{2} - w\hat{w}\right) u^{(2)}(\hat{w})v \, \mathrm{d}v}_{\text{(II)}}$$
$$- \underbrace{(2\pi)^{-1} \cdot \frac{1}{2}u^{(3)}(\hat{w}) \int_{-\infty}^{\infty} \exp\left(\frac{w^2}{2} - w\hat{w}\right) v^2 \, \mathrm{d}v}_{\text{(III)}}.$$

Term (I) is equal to $p^{(0)}(\Delta, x)$, the leading term of the saddlepoint approximation. As for term (II),

$$\text{(II)} = \mathrm{i} \cdot (2\pi)^{-1} u^{(2)}(\hat{w}) \cdot \int_{-\infty}^{\infty} \exp\left(\frac{1}{2}\left[\hat{w}^2 + 2\mathrm{i}v\hat{w} - v^2\right] - \hat{w}^2 - \mathrm{i}v\hat{w}\right) v \, \mathrm{d}v$$
$$= \mathrm{i} \cdot (2\pi)^{-1} u^{(2)}(\hat{w}) \exp\left(-\frac{\hat{w}^2}{2}\right) \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{v^2}{2}\right) v \, \mathrm{d}v = 0.$$

Similarly,

$$\text{(III)} = (2\pi)^{-1} \cdot \frac{1}{2}u^{(3)}(\hat{w}) \exp\left(-\frac{\hat{w}^2}{2}\right) \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{v^2}{2}\right) v^2 \, \mathrm{d}v$$
$$= (2\pi)^{-1/2} \cdot \frac{1}{2}u^{(3)}(\hat{w})\phi(\hat{w}) \cdot \int_{-\infty}^{\infty} \exp\left(-\frac{v^2}{2}\right) v^2 \, \mathrm{d}v \quad \text{where } \phi \text{ is the standard normal pdf}$$
$$= (2\pi)^{-1/2} \cdot \frac{1}{2}u^{(3)}(\hat{w})\phi(\hat{w}) \cdot \sqrt{2\pi} = \frac{1}{2}u^{(3)}(\hat{w})\phi(\hat{w}).$$

In summary,

$$p^{(1)}(\Delta,x) = (\mathrm{I}) - (\mathrm{III}) = \phi(\hat{w})\left[\Big(K^{(2)}(\Delta,\hat{u})\Big)^{-1/2} - \frac{1}{2}u^{(3)}(\hat{w})\right]$$
$$= \phi(\hat{w})\Big(K^{(2)}(\Delta,\hat{u})\Big)^{-1/2}\cdot\left[1 - \frac{1}{2}\cdot\frac{u^{(3)}(\hat{w})}{(K^{(2)}(\Delta,\hat{u}))^{-1/2}}\right]$$
$$= p^{(0)}(\Delta,x)\cdot\left[1 - \frac{1}{2}\cdot\frac{u^{(3)}(\Delta,\hat{w})}{(K^{(2)}(\hat{u}))^{-1/2}}\right].$$

Finally, by substituting (D.4)–(D.5), and $u'(\hat{w}) = (K^{(2)}(\Delta,\hat{u}))^{-1/2}$, we arrive at:

$$-\frac{1}{2}\cdot\frac{u^{(3)}(\hat{w})}{(K^{(2)}(\Delta,\hat{u}))^{-1/2}}$$
$$= \frac{K^{(4)}(\Delta,\hat{u})\big(u'(\hat{w})\big)^4 + 6K^{(3)}(\Delta,\hat{u})\big(u'(\hat{w})\big)^2 u^{(2)}(\hat{w}) + 3K^{(2)}(\Delta,\hat{u})\big(u^{(2)}(\hat{w})\big)^2}{8\sqrt{K^{(2)}(\Delta,\hat{u})}\cdot u'(\hat{w})}$$
$$= \frac{K^{(4)}(\Delta,\hat{u})\big(K^{(2)}(\Delta,\hat{u})\big)^2 + 6K^{(3)}(\Delta,\hat{u})\big(K^{(2)}(\Delta,\hat{u})\big)^{-1}u^{(2)}(\hat{w}) + 3K^{(2)}(\Delta,\hat{u})\big(u^{(2)}(\hat{w})\big)^2}{8}$$
$$= \frac{1}{8}\cdot\frac{K^{(4)}(\Delta,\hat{u})}{\big(K^{(2)}(\Delta,\hat{u})\big)^2} - \frac{1}{4}\cdot\frac{\big(K^{(3)}(\Delta,\hat{u})\big)^2}{\big(K^{(2)}(\Delta,\hat{u})\big)^3} + \frac{1}{24}\cdot\frac{\big(K^{(3)}(\Delta,\hat{u})\big)^2}{\big(K^{(2)}(\Delta,\hat{u})\big)^3}$$
$$= \frac{1}{8}\cdot\frac{K^{(4)}(\Delta,\hat{u})}{\big(K^{(2)}(\Delta,\hat{u})\big)^2} - \frac{5}{24}\cdot\frac{\big(K^{(3)}(\Delta,\hat{u})\big)^2}{\big(K^{(2)}(\Delta,\hat{u})\big)^3}.$$

In the case of the Hawkes jump model, however, the Gaussian base leads to an inaccurate saddlepoint approximation. This can be seen from the fact that, when using a Gaussian base, $p^{(0)}(\Delta,x) = O(\Delta^{1/4})$ in the tails, whereas the jump model is of order $O(\Delta)$ in the tails. Thus, the relative error property is not satisfied for the Hawkes jump model when using a Gaussian base.

Now let us consider a non-Gaussian base and proceed similarly. By repeatedly differentiating both sides of (C.18) with $x_0 = x$ we obtain:

$$
\begin{aligned}
& K'\big(\Delta, u(w)\big)u'(w) - u'(w)x = K_0'(\Delta, w) - x, \\
\Rightarrow\ & K^{(2)}\big(\Delta, u(w)\big)\Big(u'(w)\Big)^2 + K'\big(\Delta, u(w)\big)u^{(2)}(w) - u^{(2)}(w)x = K_0^{(2)}(\Delta, w), \\
\Rightarrow\ & K^{(3)}\big(\Delta, u(w)\big)\Big(u'(w)\Big)^3 + 3K^{(2)}\big(\Delta, u(w)\big)u'(w)u^{(2)}(w) + K'\big(\Delta, u(w)\big)u^{(3)}(w) - u^{(3)}(w)x \\
& = K_0^{(3)}(\Delta, w), \\
\Rightarrow\ & K^{(4)}(\Delta, u(w))\Big(u'(w)\Big)^4 + 6K^{(3)}(\Delta, u(w))\Big(u'(w)\Big)^2 u^{(2)}(w) + 3K^{(2)}(\Delta, u(w))\Big(u^{(2)}(w)\Big)^2 \\
& \quad + 4K^{(2)}(\Delta, u(w))u'(w)u^{(3)}(w) + K'(\Delta, u(w))u^{(4)}(w) - u^{(4)}(w)x = K_0^{(4)}(\Delta, w).
\end{aligned}
$$

Next, by evaluating at $w = \hat{w}$ and using $K'(\Delta, \hat{u}) = x$, $u(\hat{w}) = \hat{u}$, we obtain:

$$u'(\hat{w}) = \big(K_0^{(2)}(\Delta, \hat{w})\big)^{1/2} \cdot \big(K^{(2)}(\Delta, \hat{u})\big)^{-1/2}, \tag{D.6}$$

$$u^{(2)}(\hat{w}) = \frac{K_0^{(3)}(\Delta, \hat{w}) - K^{(3)}(\Delta, \hat{u})\big(u'(\hat{w})\big)^3}{3K^{(2)}(\Delta, \hat{u})u'(\hat{w})}, \tag{D.7}$$

$$u^{(3)}(\hat{w}) = -\frac{-K_0^{(4)}(\Delta, \hat{w}) + K^{(4)}(\Delta, \hat{u})\big(u'(\hat{w})\big)^4 + 6K^{(3)}(\Delta, \hat{u})\big(u'(\hat{w})\big)^2 u^{(2)}(\hat{w}) + 3K^{(2)}(\Delta, \hat{u})\big(u^{(2)}(\hat{w})\big)^2}{4K^{(2)}(\Delta, \hat{u})u'(\hat{w})}. \tag{D.8}$$

With the change of variable $u(w)$, we have

$$
\begin{aligned}
p(\Delta, x) = (2\pi\mathrm{i})^{-1} \exp\left((K(\Delta, \hat{u}) - \hat{u}x) - (K_0(\Delta, \hat{w}) - \hat{w}x)\right) \\
\times \int_{\hat{w}-\mathrm{i}\infty}^{\hat{w}+\mathrm{i}\infty} \exp\left(K_0(\Delta, w) - wx\right) \frac{\mathrm{d}u(w)}{\mathrm{d}w}\mathrm{d}w.
\end{aligned}
$$

Under the approximation of $\frac{\mathrm{d}u(w)}{\mathrm{d}w}$ near the saddlepoint $w = \hat{w}$ given by

$$u'(w) \approx u'(\hat{w}) + u^{(2)}(\hat{w})(w - \hat{w}) + \frac{1}{2}u^{(3)}(\hat{w})(w - \hat{w})^2,$$

we obtain $p^{(1)}(\Delta, x) = \text{(I)} + \text{(II)}$, where

$$
\begin{aligned}
\text{(I)} &= (2\pi\mathrm{i})^{-1} \exp\left(\Big[K(\Delta, \hat{u}) - \hat{u}x\Big] - \Big[K_0(\Delta, \hat{w}) - \hat{w}x\Big]\right) \\
&\quad \times \int_{\hat{w}-i\infty}^{\hat{w}+i\infty} \exp\Big(K_0(\Delta, w) - wx\Big)u'(\hat{w})\,\mathrm{d}w, \\
\text{(II)} &= (2\pi\mathrm{i})^{-1} \exp\left(\Big[K(\Delta, \hat{u}) - \hat{u}x\Big] - \Big[K_0(\Delta, \hat{w}) - \hat{w}x\Big]\right) \\
&\quad \times \int_{\hat{w}-i\infty}^{\hat{w}+i\infty} \exp\Big(K_0(\Delta, w) - wx\Big)\left[u^{(2)}(\hat{w})(w - \hat{w}) + \frac{1}{2}u^{(3)}(\hat{w})(w - \hat{w})^2\right] \mathrm{d}w.
\end{aligned}
$$

Term (I) yields $p^{(0)}(\Delta, x)$, the leading term of the saddlepoint approximation, which for the Hawkes jump model under the non-Gaussian, Bernoulli base is $O(\Delta)$ in the tails; see Lemma S.1. As for term (II), as $w = \hat{w} + \mathrm{i}v$, $v \in \mathbb{R}$, over the integral path, we have $u'(w) \approx u'(\hat{w}) + \mathrm{i} \cdot u^{(2)}(\hat{w})v - \frac{1}{2}u^{(3)}(\hat{w})v^2$, hence,

$$
\begin{aligned}
\text{(II)} = (2\pi)^{-1} \exp&\left(\left[K(\Delta, \hat{u}) - \hat{u}x\right] - \left[K_0(\Delta, \hat{w}) - \hat{w}x\right]\right) \\
&\times \int_{-\infty}^{\infty} \exp\left(K_0(\Delta, \hat{w} + \mathrm{i}v) - (\hat{w} + \mathrm{i}v)x\right) \left[\mathrm{i} \cdot u^{(2)}(\hat{w})v - \frac{1}{2}u^{(3)}(\hat{w})v^2\right] \mathrm{d}v.
\end{aligned}
$$

For the Hawkes jump model under the non-Gaussian base, we have from (D.7)–(D.8) that

$$
\begin{aligned}
u^{(2)}(\hat{w}) &= O(\Delta^0), \\
u^{(3)}(\hat{w}) &= O(\Delta^{-1/2}).
\end{aligned}
$$

Thus, in a neighborhood in w of order $O(\Delta^{1/2})$ of the saddlepoint $\hat{w}$, we have that (II) $= O(\Delta^{3/2})$, which is dominated by the leading term (I). This, along with an argument as in Aït-Sahalia & Yu (2006) that shows that the approximation error induced by using the expansion of K at order Δ^2 is of higher order, proves the relative error property for the Hawkes jump model using the non-Gaussian base.

D.1.3 Proof of Theorem S.10

To obtain the saddlepoint approximation to the ddf, we also need to compute Taylor series expansions of the reciprocals of the saddlepoints $\hat{u}$ and $\hat{w}$. They are given by

$$
\begin{aligned}
\hat{u}^{-1} &= \frac{1}{\gamma} + \frac{\sqrt{\Delta}}{\sqrt{\frac{\gamma^3 x}{\lambda}}} + \frac{\left(\frac{\beta\gamma(2\alpha-\beta)}{\alpha-\beta} + \frac{4\lambda}{x}\right)}{4\gamma^2}\Delta + O(\Delta^{3/2}), \\
\hat{w}^{-1} &= \frac{1}{\gamma} + \frac{\sqrt{\Delta}}{\sqrt{\frac{\gamma^3 x}{\lambda}}} + \frac{(-\gamma x + 2)\lambda}{2\gamma^2 x}\Delta + O(\Delta^{3/2}).
\end{aligned}
$$

Clearly, $\hat{u}^{-1}$ and $\hat{w}^{-1}$ agree up to order $\Delta^{1/2}$.

We finally compute a Taylor series expansion of the following term,

$$
\left\{\frac{1}{\hat{u}}\left(\frac{\partial^2 K_0(\Delta, \hat{w})}{\partial w^2}\right)^{1/2}\left(\frac{\partial^2 K(\Delta, \hat{u})}{\partial u^2}\right)^{-1/2} - \frac{1}{\hat{w}}\right\},
$$

that appears in the ddf saddlepoint approximation. It is given by

$$\begin{aligned}&\left\{\frac{1}{\hat{u}}\left(\frac{\partial^2 K_0(\Delta,\hat{w})}{\partial w^2}\right)^{1/2}\left(\frac{\partial^2 K(\Delta,\hat{u})}{\partial u^2}\right)^{-1/2}-\frac{1}{\hat{w}}\right\}\\&=\frac{\left(8\beta\lambda\left(2\alpha^2-3\alpha\beta+\beta^2\right)+3\beta^2\gamma(2\alpha-\beta)^2x+4\lambda^2(\alpha-\beta)^2(\gamma x+4)\right)}{64\gamma\lambda(\alpha-\beta)^2}\Delta\\&\quad+O(\Delta^{3/2}).\end{aligned}$$

Then the stated result follows from (C.7) upon collecting the terms we derived above.

D.1.4 Proof of Corollary S.1

The proof of Corollary S.1 now follows from (C.9).

D.2 Jump Amplification Function

So far, we have set the jump amplification to unity in this Supplement. With the jump amplification function of Section 2 added, the saddlepoint approximation to the pdf becomes:

Lemma S.3. *The saddlepoint approximation to the pdf of the Hawkes jump model* (3), (5)–(7) *with $m=1$ and $p=1$ is given at order Δ^2 by the following expression:*

$$\begin{aligned}&p^{(0)}(\Delta,x)\\&=\lambda\Delta\gamma\exp(-\gamma x)\\&\quad\times\exp\left(\frac{\gamma x}{2}\left(\lambda+\frac{\beta q_1(\alpha q_2-\beta q_1)}{\xi q_2(\alpha-\beta)}\right)\Delta\right)\\&\quad\times\Bigg(1+\left(2q_2\lambda(\alpha-\beta)\left(\beta\left(\gamma^3\beta+4\gamma^2\xi\beta+2\gamma\xi^2(2\beta+\lambda)+2\xi^3(\beta+2\lambda)\right)-\alpha q_2\left(\gamma^2\beta+2\gamma\xi\beta+2(\beta+\lambda)\xi^2\right)\right)\right.\\&\qquad\left.+\left(3q_1^2\beta^2(q_2\alpha-q_1\beta)^2+\lambda^2\xi^2q_2^2(\alpha-\beta)^2\right)\gamma x\right)\\&\quad\times\frac{1}{16\lambda\xi^2q_2^2(\alpha-\beta)^2}\Delta\Bigg)+o(\Delta^2),\end{aligned}$$

with $q_1=\gamma+\xi$ and $q_2=\gamma+2\xi$.

Furthermore, with the jump amplification function added, the saddlepoint approximation to the ddf is given in Theorem 1, and the approximated saddlepoint becomes:

$$\begin{aligned}\hat{u}&=\gamma-\sqrt{\frac{\gamma\lambda}{x}}\sqrt{\Delta}-\frac{q_1\beta\gamma(q_2\alpha-q_1\beta)}{2q_2\xi(\alpha-\beta)}\Delta\\&\quad-\frac{\beta\sqrt{\gamma}\left(-3q_1^2\beta\gamma(q_2\alpha-q_1\beta)^2x-2q_2\lambda(\alpha-\beta)\left(q_2\alpha(\gamma^2+2\gamma\xi+2\xi^2)-\beta(\gamma^3+4\gamma^2\xi+4\gamma\xi^2+2\xi^3)\right)\right)}{8\sqrt{\lambda}q_2^2\xi^2(\alpha-\beta)^2\sqrt{x}}\Delta^{3/2}\\&\quad+O(\Delta^2).\end{aligned}\tag{D.9}$$

D.3 The Marginal Tail in the Bivariate Model

For ease of exposition, we restrict attention to the special case in which $\mu_i = \sigma_i = 0$ and $\phi_{i,j}(Z_j) \equiv 1$. We state the following results:

Lemma S.4. *The saddlepoint approximation to the first marginal pdf of the Hawkes jump model* (3), (5)–(7) *with $m = 2$ and $p_j = 1$, $j = 1, 2$, is given at order Δ^2 by the following expression:*

$$
\begin{aligned}
&p^{(0)}(\Delta, x_1) \\
&= \lambda_1 \Delta \gamma_1 \exp(-\gamma_1 x_1) \\
&\quad \times \exp\left(\frac{\gamma_1 x_1(\lambda_1(2\lambda_1 r + s) + \lambda_2 t)}{4\lambda_1 r}\Delta\right) \\
&\quad \times \left(1 + \vphantom{\frac{1}{1}}\right. \\
&\quad \left.\frac{\lambda_1^2\left(4\lambda_1^2 r^2(\gamma_1 x_1 - 4) - 8\lambda_1 r s + 3\gamma_1 s^2 x_1\right) + 2\lambda_1\lambda_2 t(3\gamma_1 s x_1 - 4\lambda_1 r) + 3\gamma_1\lambda_2^2 t^2 x_1}{64\lambda_1^3 r^2}\Delta\right) \\
&\quad + o(\Delta^2),
\end{aligned}
$$

with r, s and t defined in (D.11)–(D.13).

Theorem S.11. *The saddlepoint approximation to the first marginal ddf of the Hawkes jump model* (3), (5)–(7) *with $m = 2$ and $p_j = 1$, $j = 1, 2$, is given at order Δ^2 by the following expression:*

$$
\begin{aligned}
&\overline{P}^{(0)}(\Delta, x_1) \\
&= \exp\left(\frac{\gamma_1 x_1(\lambda_1(2\lambda_1 r + s) + \lambda_2 t)}{4\lambda_1 r}\Delta\right) \\
&\quad \times \left(\lambda_1 \Delta \exp(-\gamma_1 x_1) + \lambda_1 \Delta \gamma_1 \exp(-\gamma_1 x_1)\right. \\
&\quad \left. \times \frac{\lambda_1^2\left(4\lambda_1^2 r^2(\gamma_1 x_1 + 4) + 8\lambda_1 r s + 3\gamma_1 s^2 x_1\right) + 2\lambda_1\lambda_2 t(4\lambda_1 r + 3\gamma_1 s x_1) + 3\gamma_1\lambda_2^2 t^2 x_1}{64\gamma_1\lambda_1^3 r^2}\Delta\right) \\
&\quad + o(\Delta^2),
\end{aligned}
\tag{D.10}
$$

with r, s and t defined in (D.11)–(D.13).

Corollary S.2. *We have the following saddlepoint approximation for $\mathcal{L}(\Delta, x_1)$ defined in* (10) *in the bivariate Hawkes jump model of Theorem S.11:*

$$
\begin{aligned}
\mathcal{L}^{(0)}(\Delta, x_1) &= K_1(\Delta, 0)\overline{P}^{(0)}(\Delta, x_1) + \frac{x_1 - K_1(\Delta, 0)}{\hat{u}_1} p^{(0)}(\Delta, x_1) \\
&= \frac{\lambda_1}{\gamma_1}\Delta\overline{P}^{(0)}(\Delta, x_1) + \frac{x_1 - (\lambda_1/\gamma_1)\Delta}{\hat{u}_1} p^{(0)}(\Delta, x_1),
\end{aligned}
$$

with $\overline{P}^{(0)}(\Delta, x_1)$, $\hat{u}_1$ *and* $p^{(0)}(\Delta, x_1)$ *given by Theorem S.11, Equation* (D.14) *and Lemma S.4, respectively.*

D.3.1 Proof of Lemma S.4

We start with an expansion of the cgf of the bivariate Hawkes jump-diffusion model, up to the second order in Δ (see Theorem S.8 in this Supplement). We restrict attention to the special case in which $\mu_i = \sigma_i = 0$ and $\phi_{i,j}(Z_j) \equiv 1$, $i, j = 1, 2$:

$$\begin{aligned}\overline{K}(\Delta, u_1) = &\lambda_1 \left(L(u_1, Z_1) - 1\right) \Delta \\ &+ \frac{(\lambda_1 s + \lambda_2 t)\left(L(u_1, Z_1) - 1\right)^2}{4r}\Delta^2 \\ &+ o(\Delta^2),\end{aligned}$$

where $L(u_1, Z_1) = \mathbb{E}\left[e^{u_1 Z_1}\right]$ and

$$r := (\alpha_1 + \alpha_2 - \beta_{1,1} - \beta_{2,2})((\alpha_1 - \beta_{1,1})(\alpha_2 - \beta_{2,2}) - \beta_{1,2}\beta_{2,1}); \tag{D.11}$$

$$\begin{aligned}s := &2\alpha_1^2\beta_{1,1}(\alpha_2 - \beta_{2,2}) + \alpha_1\beta_{1,1}((\alpha_2 - \beta_{2,2})(2\alpha_2 - 3\beta_{1,1} - 2\beta_{2,2}) - 2\beta_{1,2}\beta_{2,1}) \\ &- (\alpha_2\beta_{1,1} + \beta_{1,2}\beta_{2,1} - \beta_{1,1}\beta_{2,2})(\alpha_2\beta_{1,1} - \beta_{1,2}\beta_{2,1} - \beta_{1,1}(\beta_{1,1} + \beta_{2,2}));\end{aligned} \tag{D.12}$$

$$t := \beta_{1,2}^2(\alpha_2(\alpha_1 + \alpha_2 - \beta_{1,1}) - \beta_{1,2}\beta_{2,1} + \beta_{2,2}(\beta_{1,1} - \alpha_1)). \tag{D.13}$$

With (one-sided, positive) exponential jump size distribution $F_{Z_1}(x) = 1 - \exp(-\gamma_1 x)$, $\gamma_1 > 0$, hence $L(u_1, Z_1) = \frac{\gamma_1}{\gamma_1 - u_1}$, $u_1 < \gamma_1$, taking for ease of exposition Z_1 to be the sign-changed counterpart of the jump magnitude Z_1 in the main text and in Sections A–B, this specializes to

$$\begin{aligned}K(\Delta, u_1) = &\frac{\lambda_1 u_1}{\gamma_1 - u_1}\Delta \\ &+ \frac{(\lambda_1 s + \lambda_2 t)u_1^2}{4r(u_1 - \gamma_1)^2}\Delta^2 \\ &+ o(\Delta^2).\end{aligned}$$

Henceforth, we omit terms of $o(\Delta^2)$ in $K(\Delta, u_1)$.

We next compute a Taylor series expansion of the saddlepoint $\hat{u}_1 = \hat{u}_1(\Delta, x_1)$ solving

$$K_1(\Delta, u_1) = x_1, \qquad \text{with} \qquad K_1(\Delta, u_1) = \frac{\partial K(\Delta, u_1)}{\partial u_1}.$$

It is given by (the coefficient of $\sqrt{\Delta}$ in the expansion below is chosen such that it induces

$\hat{u}_1 < \gamma_1$)

$$\begin{aligned}\hat{u}_1 = \gamma_1 & \\ & - \sqrt{\frac{\gamma_1\lambda_1}{x_1}}\sqrt{\Delta} \\ & - \frac{\gamma_1(\lambda_1 s + \lambda_2 t)}{4\lambda_1 r}\Delta \\ & - \frac{\sqrt{\gamma_1}(\lambda_1 s + \lambda_2 t)\left(-3\gamma_1 x_1(\lambda_1 s + \lambda_2 t) - 8\lambda_1^2 r\right)}{32\sqrt{\lambda_1}\lambda_1^2 r^2\sqrt{x_1}}\Delta^{3/2} \\ & + O(\Delta^2). \qquad \text{(D.14)}\end{aligned}$$

As before, we consider a Bernoulli base:

$$F_0(\Delta, x) = (1 - \lambda_1\Delta) + \lambda_1\Delta F_{Z_1}(x), \qquad x \geq 0,$$

and

$$K_0(\Delta, w) = \log\left(1 - \lambda_1\Delta + \lambda_1\Delta L(w, Z_1)\right).$$

With (one-sided, positive) exponential jump size distribution this specializes to

$$\begin{aligned}F_0(\Delta, x) &= 1 - \lambda_1\Delta\exp(-\gamma_1 x); \\ f_0(\Delta, x) &= \lambda_1\Delta\gamma_1\exp(-\gamma_1 x); \\ K_0(\Delta, w) &= \log\left(1 - \lambda_1\Delta + \lambda_1\Delta\frac{\gamma_1}{\gamma_1 - w}\right).\end{aligned}$$

Note that K_0 and K are again equal up to order Δ.

The exact as well as a Taylor series expansion of the saddlepoint $\hat{w} = \hat{w}(\Delta, x_1)$ are given by the following expressions, respectively:

$$\hat{w} = \frac{\gamma_1(2 - \lambda_1\Delta)x_1 - \sqrt{\gamma_1\lambda_1\Delta}\sqrt{x_1(\lambda_1\Delta(\gamma_1 x_1 - 4) + 4)}}{2(1 - \lambda_1\Delta)x_1},$$

and

$$
\begin{aligned}
\hat{w} = \gamma_1 & \\
& - \sqrt{\frac{\gamma_1 \lambda_1}{x_1}} \sqrt{\Delta} \\
& + \frac{\gamma_1 \lambda_1}{2} \Delta \\
& + \frac{\left(-4\sqrt{\gamma_1}\lambda_1^{3/2} - \gamma_1^{3/2}\lambda_1^{3/2} x_1\right)}{8\sqrt{x_1}} \Delta^{3/2} \\
& + O(\Delta^2).
\end{aligned}
$$

For consistency with our treatment of $\hat{u}_1$, we will henceforth use in our computations the latter Taylor expansion of $\hat{w}$ instead of the former exact expression for $\hat{w}$. Note that $\hat{u}_1$ and $\hat{w}$ agree up to order $\Delta^{1/2}$.

We next compute an expansion up to the first order in Δ of $(K(\Delta, \hat{u}_1) - \hat{u}_1 x_1) - (K_0(\Delta, \hat{w}) - \hat{w} x_1)$. It is given by

$$
\begin{aligned}
(K(\Delta, \hat{u}_1) - \hat{u}_1 x_1) - (K_0(\Delta, \hat{w}) - \hat{w} x_1) = & \frac{\gamma_1 x_1 (\lambda_1 (2\lambda_1 r + s) + \lambda_2 t)}{4\lambda_1 r} \Delta \\
& + O(\Delta^{3/2}).
\end{aligned}
$$

We finally compute $\left(\frac{\partial^2 K_0(\Delta, \hat{w})}{\partial w^2}\right)^{1/2} \left(\frac{\partial^2 K(\Delta, \hat{u}_1)}{\partial u_1^2}\right)^{-1/2}$ up to the first order in Δ. It is given by

$$
\begin{aligned}
& \left(\frac{\partial^2 K_0(\Delta, \hat{w})}{\partial w^2}\right)^{1/2} \left(\frac{\partial^2 K(\Delta, \hat{u}_1)}{\partial u_1^2}\right)^{-1/2} \\
& = 1 \\
& + \frac{\lambda_1^2 \left(4\lambda_1^2 r^2 (\gamma_1 x_1 - 4) - 8\lambda_1 r s + 3\gamma_1 s^2 x_1\right) + 2\lambda_1 \lambda_2 t (3\gamma_1 s x_1 - 4\lambda_1 r) + 3\gamma_1 \lambda_2^2 t^2 x_1}{64 \lambda_1^3 r^2} \Delta \\
& + O(\Delta^{3/2}).
\end{aligned}
$$

Then the stated result follows from (C.5) upon collecting the terms we derived above.

D.3.2 Proof of Theorem S.11

Taking the components derived in the proof of Lemma S.4 as input, we next compute Taylor series expansions of the reciprocals of the saddlepoints $\hat{u}_1$ and $\hat{w}$. They are given by

$$\hat{u}_1^{-1} = \frac{1}{\gamma_1} + \frac{\sqrt{\Delta}}{\sqrt{\frac{\gamma_1^3 x_1}{\lambda_1}}} + \frac{\frac{\gamma_1\left(s+\frac{\lambda_2 t}{\lambda_1}\right)}{r} + \frac{4\lambda_1}{x_1}}{4\gamma_1^2}\Delta + O(\Delta^{3/2}),$$

and

$$\hat{w}^{-1} = \frac{1}{\gamma_1} + \frac{\sqrt{\Delta}}{\sqrt{\frac{\gamma_1^3 x_1}{\lambda_1}}} + \frac{(-\gamma_1 x_1 + 2)\lambda_1}{2\gamma_1^2 x_1}\Delta + O(\Delta^{3/2}).$$

Clearly, $\hat{u}_1^{-1}$ and $\hat{w}^{-1}$ agree up to order $\Delta^{1/2}$.

We finally compute the following Taylor series expansion of the following term that appears in the ddf saddlepoint approximation:

$$\begin{aligned}&\left\{\frac{1}{\hat{u}_1}\left(\frac{\partial^2 K_0(\Delta,\hat{w})}{\partial w^2}\right)^{1/2}\left(\frac{\partial^2 K(\Delta,\hat{u}_1)}{\partial u_1^2}\right)^{-1/2} - \frac{1}{\hat{w}}\right\}\\ &= \frac{\lambda_1^2\left(4\lambda_1^2 r^2(\gamma_1 x_1 + 4) + 8\lambda_1 r s + 3\gamma_1 s^2 x_1\right) + 2\lambda_1\lambda_2 t(4\lambda_1 r + 3\gamma_1 s x_1) + 3\gamma_1\lambda_2^2 t^2 x_1}{64\gamma_1\lambda_1^3 r^2}\Delta\\ &+ O(\Delta^{3/2}).\end{aligned}$$

Then the stated result follows from (C.7) upon collecting the terms we derived above.

D.3.3 Proof of Corollary S.2

The proof of Corollary S.2 now follows from (C.9).

D.4 The Bivariate Model

For ease of exposition, we restrict attention to the special case in which $\mu_i = \sigma_i = 0$ and $\phi_{i,j}(Z_j) \equiv 1$. We state the following lemma:

Lemma S.5. *The saddlepoint approximation to the conditional pdf of the Hawkes jump model* (3), (5)–(7) *with* $m = 2$ *and* $p_j = 1$, $j = 1, 2$, *is given at order* Δ^2 *by the following* *expression:*

$$
\begin{aligned}
&p^{(0)}(\Delta, x_2|x_1) \\
&= \lambda_2 \Delta \gamma_2 \exp(-\gamma_2 x_2) \\
&\quad \times \exp\left(\frac{f\sqrt{\gamma_1\lambda_1\gamma_2\lambda_2 x_1 x_2} + \gamma_1\lambda_1 x_2 (c + \gamma_2^2\lambda_2^2)}{2\gamma_1\lambda_1\gamma_2\lambda_2} \Delta \right) \\
&\quad \times \left(1 + \frac{f\left(\sqrt{\gamma_1\lambda_1 x_1} + \sqrt{\gamma_2\lambda_2 x_2}\right)}{8\gamma_1\lambda_1\gamma_2\lambda_2} \Delta^{1/2} \right. \\
&\quad \left. + \frac{-3b^2\gamma_2 x_1 + \gamma_1^2\lambda_1^2\left(-2(f + d\gamma_2) + \gamma_1\lambda_1\gamma_2\lambda_2(\gamma_2 x_2 - 4)\right)}{16\gamma_1^3\lambda_1^3\gamma_2} \Delta \right) \\
&\quad + o(\Delta^2).
\end{aligned}
$$

Here,

$$
\begin{aligned}
b &= \gamma_1^2\left(\nu_{1,1} + (\beta_{1,1} - \lambda_1)\lambda_1\right), \\
c &= \gamma_2^2\left(\nu_{2,2} + (\beta_{2,2} - \lambda_2)\lambda_2\right), \\
d &= -\gamma_1\left(2\nu_{1,1} + 2\nu_{1,2} + \beta_{1,2}\lambda_2 + \lambda_1\left(2\beta_{1,1} + \beta_{2,1} - 2(\lambda_1 + \lambda_2)\right)\right), \\
f &= \gamma_1\gamma_2\left(2\nu_{1,2} + \beta_{2,1}\lambda_1 + (\beta_{1,2} - 2\lambda_1)\lambda_2\right).
\end{aligned}
$$

D.5 Monte Carlo Simulations

In this subsection, we analyze the accuracy of our saddlepoint approximations by comparing them to estimates obtained using Monte Carlo simulation. In particular, we compare our saddlepoint approximation to the pdf $p(\Delta, x)$, given in Lemma S.1, to its Monte Carlo simulated counterpart. We first specialize the P&L model to the Poissonian case, i.e., $\beta \equiv 0$, which admits a closed-form expression for the pdf. In this case, we separately compare the tail probabilities obtained from our saddlepoint approximations and from Monte Carlo simulations to the tail probabilities obtained from the exact, closed-form expression. This allows us to benchmark the accuracy of both our saddlepoint approximation and the Monte Carlo simulated values against the known exact expression.

The results in the Poissonian case, displayed in Figure S.1, show that the saddlepoint approximations provide more accurate estimates of the tail probabilities than the Monte Carlo simulations, even with $> 10^7$ simulated increments. Only when we increase the number of replications to more than 10^8, the Monte Carlo simulated values become competitive to the saddlepoint approximations in terms of accuracy. Accurately sampling tails is known to be challenging.

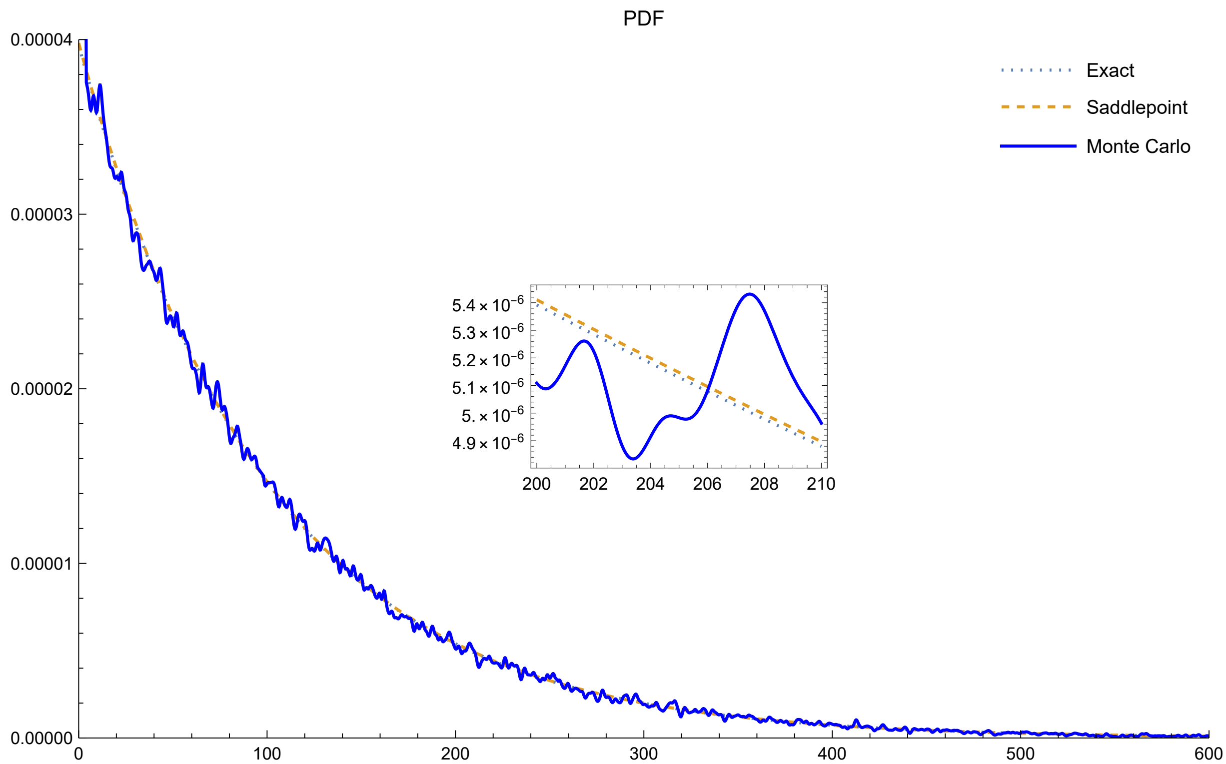

Figure S.1: Univariate pdf: Poisson model

Note: This figure plots the exact and saddlepoint approximation to the Poissonian pdf against the Monte Carlo simulated pdf, including a zoomed-in version. We simulate $2.51 * 10^7$ increments. The parameter values are $\lambda = 1.0$, $\gamma = 1/100$ and $\Delta = 1/251$. The exact and saddlepoint approximation to the pdf are nearly indistinguishable.

Next, we consider the Hawkes case, i.e., $\beta > 0$, which does not admit an exact, closed-form expression for the pdf. The results are displayed in Figure S.2. Under moderate degrees of excitation, there is a reasonable match between the saddlepoint approximations and the (noisy) Monte Carlo simulated estimates of the tail probabilities. The plot also includes the exact Poissonian pdf for comparison. As expected, the Poissonian tail is thinner than that of the Hawkes self-exciting model, which is approximated by the saddlepoint.

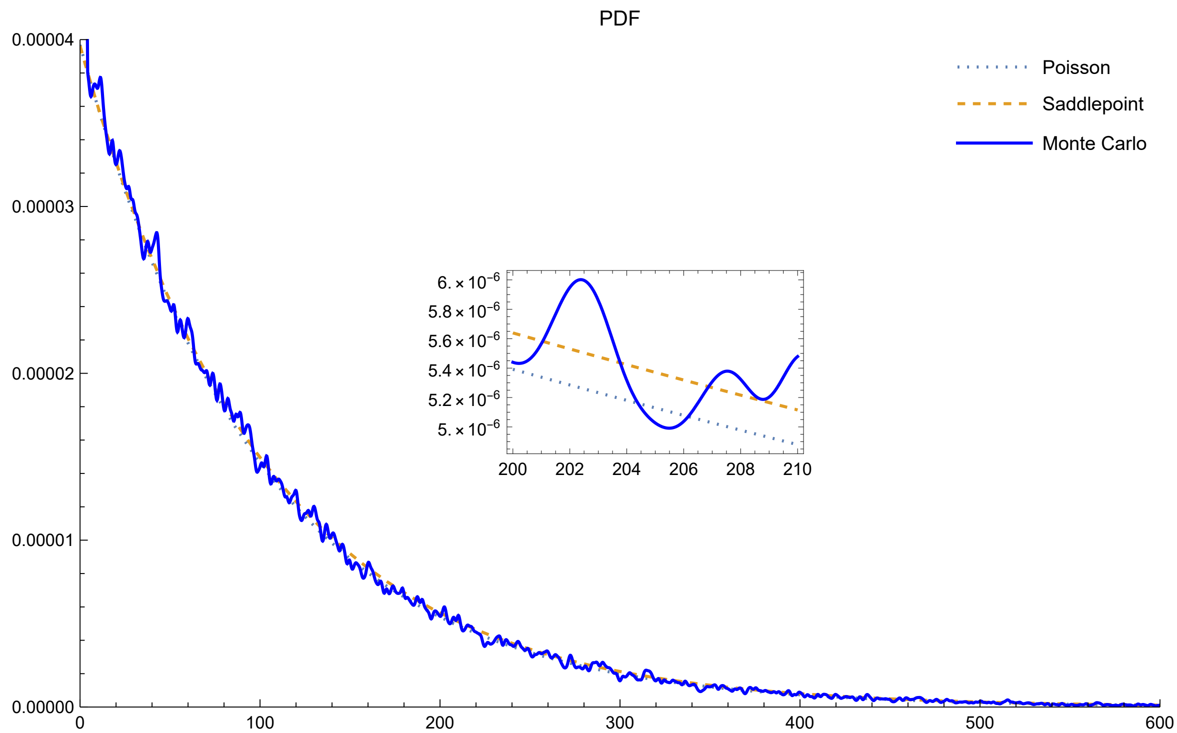

Figure S.2: Univariate pdf: Hawkes self-exciting model

Note: This figure plots the saddlepoint approximation to the Hawkes pdf against the Monte Carlo simulated pdf, including a zoomed-in version. We simulate $2.51 * 10^7$ increments. The parameter values are $\alpha = 1.5$, $\beta = 1.25$, $\lambda = 1.0$, $\gamma = 1/100$ and $\Delta = 1/251$. The plot also includes the exact Poissonian pdf (which is misspecified for this model) with $\lambda = 1$ for comparison.

E Preliminaries on Eigenvector Centrality

In the nomenclature of graphs and networks, the mutual excitation matrix $\boldsymbol{\beta}$ is a "weighted adjacency matrix of a directed graph with self-loops". To understand this, we first recall the definition of an adjacency matrix associated with a directed graph. The adjacency matrix $\mathbf{A}$ of the directed graph (V, E), with V a set of vertices (or nodes) and $E \subset V \times V$ a set of directed edges, where the directed edge (v_i, v_j) goes from node v_i to node v_j, is the square matrix that satisfies $A_{ij} = 1$ if $(v_j, v_i) \in E$ and 0 otherwise. Note the indexing convention: A_{ij} refers to (v_j, v_i). If multiple edges between nodes are allowed, one speaks of multigraphs; in this case, A_{ij} equals the number of directed edges between node v_j and v_i. Undirected graphs occur if $A_{ij} = A_{ji}$, $i = 1, \ldots, m$, with m the number of nodes, that is, if $\mathbf{A}$ is symmetric. A node v_i has a self-loop if there is a directed edge (v_i, v_i). When edge weights are introduced, associating weights, assumed to be non-negative throughout, to directed edges, one speaks of weighted graphs. If the weights are positive integers, then they can be mapped onto multigraphs (Newman 2004). The representing adjacency matrix has entries equal to the edge weights. See e.g., Brouwer & Haemers (2011) for further details. Hence, in our Hawkes jump model, the matrix consisting of the cross-excitation parameters $\beta_{i,j}$ can be seen as a weighted adjacency matrix, for a directed graph with self-loops, so that $\beta_{i,j} = A_{ij}$.

An important measure of "centrality" is that of eigenvector centrality (Bonacich 1972 and Bonacich 1987). It identifies the "most influential" node in a graph by summing a node's connections to other nodes, weighing by their centralities. For an unweighted and undirected graph, the eigenvector centrality is given by the eigenvector $\mathbf{x}$, associated with the dominant (leading or principal) eigenvalue $\ell_{\max}$ of the adjacency matrix, solving

$$\ell_{\max}\mathbf{x} = \mathbf{A}\mathbf{x}.$$

Existence and uniqueness is guaranteed by the Perron-Frobenius Theorem discussed later. Of course, this notion of centrality is related to the spectral decomposition of $\mathbf{A}$. The appropriate generalization of this centrality notion to weighted graphs is still the dominant eigenvector of the adjacency matrix with edge weights as entries (Newman 2004). The change from integer to non-integer entries, i.e., from multigraphs to weighted graphs, can be achieved by a change of the unit of flow. Furthermore, the natural generalization of this centrality notion to directed graphs is the leading eigenvector of the "forward operator" $\mathbf{A}'$ (Canright & Engø-Monsen 2007), i.e., the eigenvector $\mathbf{x}$, associated with the leading eigenvalue $\ell_{\max}$ of $\mathbf{A}'$, solving

$$\ell_{\max}\mathbf{x} = \mathbf{A}'\mathbf{x}.$$

It is a measure of "out-centrality". The leading eigenvector of the "backward operator" $\mathbf{A}$ provides a measure of "in-centrality".

Finally, one may divide each non-zero column of $\mathbf{A}$ by the column sum of its elements yielding the standardized adjacency matrix $\mathbf{A}_0$ with principal eigenvalue equal to unity. The associated principal eigenvector is related to Google's "Pagerank" centrality measure. Specifically, Pagerank is given by the principal eigenvector $\mathbf{x}$ solving

$$\mathbf{x} = \frac{1-d}{m}\mathbf{1} + d\mathbf{A}_0\mathbf{x},$$

with $0 \leq d < 1$ a damping factor and $\mathbf{1}$ a column vector of length m containing ones. That is, apart from dampening, Pagerank is the principal (right) eigenvector of the standardized adjacency matrix (Canright & Engø-Monsen 2007). Pagerank centralities automatically sum to unity.

Formally, by the Perron-Frobenius Theorem, for a square matrix $\mathbf{A}$ with non-negative entries, if for any i, j there exists a $k(i, j)$ such that $(\mathbf{A}^k)_{ij} > 0$ so that $\mathbf{A}$ is irreducible, then $\mathbf{A}$ has a unique principal positive real eigenvalue $\ell_{\max}$ such that for all other eigenvalues ℓ of $\mathbf{A}$ we have $|\ell| \leq \ell_{\max}$. Furthermore, the eigenvector $\mathbf{x}$ associated with $\ell_{\max}$ can be taken to satisfy $\mathbf{x} > 0$ (i.e., all entries of $\mathbf{x}$ are positive). See e.g., Brouwer & Haemers (2011) for details.

Against the previous backdrop, we finally point out four differences between the square matrix $\boldsymbol{\beta}$ and the more familiar matrix of linear (Pearson's) correlations coefficients i.e., a square symmetric matrix that is positive semi-definite with entries taking absolute values in the unit interval, fully characterizing the dependence structure in a multivariate Gaussian model. (i) Asymmetry vs. Symmetry: First, the matrix $\boldsymbol{\beta}$ is allowed to be asymmetric, contrary to a linear correlation matrix. In fact, the asymmetry in mutual excitation is a key feature we seek to capture. (ii) Non-standardized and non-negative vs. Standardized and real: Second, the entries of the matrix $\boldsymbol{\beta}$ are arbitrary non-negative numbers (though we will assume that stationarity of the process is maintained), not standardized to the unit interval. In this respect the matrix $\boldsymbol{\beta}$ is perhaps somewhat more similar to a variance-covariance matrix than to a matrix of linear correlation coefficients: both the variance-covariance matrix and the linear correlation matrix are symmetric, but the former is non-standardized. We note, however, that covariances and correlations can be negative, while the $\beta_{i,j}$'s are restricted to be non-negative (i.e., we restrict to mutual excitation, and do not allow for mutual dampening). (iii) Marginals and Dependence Structure vs. Dependence Structure: Third, note that the $\beta_{i,j}$-parameters do not just dictate the dependence structure but that also the marginal distributions in our model already depend upon the $\beta_{i,j}$-parameters, thus impacting the tails of the marginal loss distributions, unlike in the Gaussian model in which the dependence structure, as characterized by the matrix of linear correlation coefficients, is separated from the marginal parameters given by the means and variances. We note, however, that outside the Gaussian model, this separation does no longer hold in the sense that outside the Gaussian model linear correlation coefficients generally depend upon the marginal distributions. (iv) Cross-Sectional and Serial vs. Cross-Sectional Correlation: Finally, the mutually exciting model induces contemporaneous as well as serial cross-correlations. The latter are absent in the multivariate Brownian model, where all correlation is instantaneous.

References

Aït-Sahalia, Y., Cacho-Diaz, J. & Laeven, R. J. (2015), 'Modeling financial contagion using mutually exciting jump processes', *Journal of Financial Economics* **117**, 585–606.

Aït-Sahalia, Y., Fan, J., Laeven, R. J., Wang, C. D. & Yang, X. (2017), 'Estimation of the continuous and discontinuous leverage effects', *Joural of the American Statistical Association* **112**, 1744–1758.

Aït-Sahalia, Y., Hansen, L. P. & Scheinkman, J. A. (2010), Operator methods for continuous-time Markov processes, *in* Y. Aït-Sahalia & L. P. Hansen, eds, 'Handbook of Financial Econometrics', North-Holland, Amsterdam, The Netherlands, pp. 1–66.

Aït-Sahalia, Y. & Yu, J. (2006), 'Saddlepoint approximations for continuous-time Markov processes', *Journal of Econometrics* **134**, 507–551.

Bonacich, P. (1972), 'Factoring and weighting approaches to status scores and clique identification', *Journal of Mathematical Sociology* **2**, 113–120.

Bonacich, P. (1987), 'Power and centrality: A family of measures', *American Journal of Sociology* **92**(5), 1170–1182.

Brouwer, A. & Haemers, W. H. (2011), *Spectra of Graphs*, Springer, New York.

Butler, R. W. (2007), *Saddlepoint Approximations with Applications*, Cambridge University Press, Cambridge, UK.

Canright, G. S. & Engø-Monsen, K. (2007), Some relevant aspects of network analysis and graph theory, *in* J. Bergstra & M. Burgess, eds, 'Handbook of Network and System Administration', Elsevier, Amsterdam.

Daniels, H. (1954), 'Saddlepoint approximations in statistics', *Annals of Mathematical Statistics* **25**, 631–650.

Duffie, D., Pan, J. & Singleton, K. J. (2000), 'Transform analysis and asset pricing for affine jump-diffusions', *Econometrica* **68**, 1343–1376.

Errais, E., Giesecke, K. & Goldberg, L. R. (2010), 'Affine point processes and portfolio credit risk', *SIAM Journal on Financial Mathematics* **1**, 642–665.

Esscher, F. (1932), 'On the probability function in the collective theory of risk', *Scandinavian Actuarial Journal* **15**, 78–86.

Lugannani, R. & Rice, S. (1980), 'Saddlepoint approximation for the distribution of the sum of independent random variables', *Advances in Applied Probability* **12**, 475–490.

Martin, R. (2006), 'The saddlepoint method and portfolio optionalities', *Risk Magazine* **19**, 93–95.

Newman, M. E. (2004), 'Analysis of weighted networks', *Phys. Rev. E* **70**(5), 056131.

Wood, A. T. A., Booth, J. G. & Butler, R. W. (1993), 'Saddlepoint approximations to the CDF of some statistics with nonnormal limit distributions', *Journal of the American Statistical Association* **88**, 680–686.