Hawkes Process Models for High-Frequency Trade Arrivals

T. Zamrik • 12-DEC-2025

1. Abstract

We study multivariate Hawkes processes as models for high-frequency trade arrivals, capturing the self-exciting and cross-exciting structure of order flow across asset classes. The intensity of each arrival stream depends on the full history of all streams via a matrix of exponential kernels, whose parameters are estimated by maximum likelihood using an expectation-maximisation algorithm. Calibrated on simulated tick data, the model reproduces the clustering of trades, the elevated autocorrelation of inter-arrival times, and the asymmetric cross-excitation patterns characteristic of correlated liquid assets.

2. Introduction

High-frequency financial data exhibit pronounced clustering: trades arrive in bursts, with short inter-arrival times concentrated around market events, news releases, and large order executions. Standard Poisson models, which assume independent and identically distributed inter-arrivals, fail to capture this empirical regularity.

The Hawkes process, introduced by Hawkes (1971), provides a tractable self-exciting point process framework in which each event temporarily elevates the arrival intensity of future events. Its multivariate extension allows cross-excitation between streams, making it naturally suited to modelling the joint arrival dynamics of trades across correlated assets.

This note develops the multivariate Hawkes model for trade arrivals, derives the expectation-maximisation (EM) estimator, and analyses the implied clustering and cross-excitation structure. Section 2 defines the model. Section 3 derives the log-likelihood and EM algorithm. Section 4 presents numerical experiments. Section 5 concludes.

3. Model

3.1 Univariate Hawkes Process

Let $N(t)$ be a counting process with stochastic intensity $\lambda(t)$. The univariate Hawkes process is defined by:

$$\lambda(t) = \mu + \sum_{t_i < t} \phi(t - t_i), \tag{3.1}$$

where $\mu > 0$ is the baseline intensity and $\phi : [0, \infty) \to [0, \infty)$ is the excitation kernel. We use the exponential kernel:

$$\phi(t) = \alpha\, e^{-\beta t}, \qquad \alpha, \beta > 0, \tag{3.2}$$

which gives the intensity the Markovian representation $d\lambda(t) = -\beta(\lambda(t) - \mu)\, dt + \alpha\, dN(t)$. Stationarity requires the branching ratio $n = \alpha/\beta < 1$. The stationary mean intensity is $\bar{\lambda} = \mu/(1 - n)$.

3.2 Multivariate Extension

Let $\mathbf{N}(t) = (N_1(t), \dots, N_D(t))^\top$ be a D-dimensional counting process. The intensity of stream d is:

$$\lambda_d(t) = \mu_d + \sum_{d'=1}^{D} \sum_{t_i^{d'} < t} \alpha_{dd'}\, e^{-\beta_{dd'}(t - t_i^{d'})}, \tag{3.3}$$

where $\alpha_{dd'} \geq 0$ measures the excitation from stream d' to stream d and $\beta_{dd'} > 0$ is the corresponding decay rate. The matrix $A = (\alpha_{dd'}/\beta_{dd'})$ is the branching ratio matrix; stationarity requires that the spectral radius $\rho(A) < 1$.

3.3 Cluster Representation

The Hawkes process admits a Poisson cluster (immigration-birth) representation. Each event is either an immigrant, arriving at baseline rate μ_d, or an offspring triggered by a previous event. The parameter $\alpha_{dd'}/\beta_{dd'}$ is the expected number of type-d offspring generated by each type-d' event.

4. Estimation

4.1 Log-likelihood

Given observations $\{t_i^d\}$ for $d = 1, \dots, D$ on $[0, T]$, the log-likelihood is:

$$\ell(\theta) = \sum_{d=1}^{D} \left[\sum_i \log \lambda_d(t_i^d) - \int_0^T \lambda_d(s)\, ds \right]. \tag{4.1}$$

The integrated intensity has the closed form:

$$\int_0^T \lambda_d(s)\,ds = \mu_d T + \sum_{d'} \frac{\alpha_{dd'}}{\beta_{dd'}} \sum_{t_i^{d'} \le T} \left(1 - e^{-\beta_{dd'}(T - t_i^{d'})}\right). \tag{4.2}$$

4.2 EM Algorithm

Direct maximisation of ℓ via gradient methods is feasible but requires careful initialisation. The EM algorithm exploits the cluster representation: treating the parentage of each event (immigrant or offspring of a specific ancestor) as latent data yields closed-form M-step updates.

E-step. Compute the probability that event j of type d was triggered by event i of type d':

$$p_{ij}^{dd'} = \frac{\alpha_{dd'}\, e^{-\beta_{dd'}(t_j^d - t_i^{d'})}}{\lambda_d(t_j^d)}, \qquad t_i^{d'} < t_j^d. \tag{4.3}$$

M-step. Update parameters:

$$\hat{\mu}_d = \frac{1}{T} \sum_j p_{j,\mathrm{imm}}^d, \qquad \hat{\alpha}_{dd'} = \frac{\sum_{j>i} p_{ij}^{dd'}}{\sum_i (1 - e^{-\beta_{dd'}(T - t_i^{d'})})/\beta_{dd'}}. \tag{4.4}$$

The decay rates $\beta_{dd'}$ are updated by a one-dimensional Newton step on the concentrated log-likelihood.

4.3 Algorithm

```
Input : event streams {t_i^d}, d=1..D; initial theta_0; tolerance eps
Output: MLE estimate theta*

Initialise: mu, alpha, beta <- theta_0
Repeat:
  -- E-step --
  For each event j of type d:
    Compute lambda_d(t_j) using current params
    For each ancestor i of type d' with t_i < t_j:
      p_ij <- alpha[d,d'] * exp(-beta[d,d'] * (t_j - t_i)) / lambda_d(t_j)
    p_imm[j] <- mu_d / lambda_d(t_j)
  -- M-step --
  mu_d        <- (1/T) * sum_j p_imm[j]
  alpha[d,d'] <- sum_{j>i} p_ij / sum_i (1 - exp(-beta*(T-t_i))) / beta
  beta[d,d']  <- Newton step on concentrated likelihood
  delta <- || theta_new - theta_old || / || theta_old ||
Until delta < eps
Return theta
```

5. Numerical Experiments

5.1 Simulation

We simulate a bivariate Hawkes process ($D = 2$) with parameters:

$$\mu = (0.5,\ 0.3), \qquad A = \begin{pmatrix} 0.4 & 0.2 \\ 0.15 & 0.35 \end{pmatrix}, \qquad \beta = \begin{pmatrix} 2.0 & 1.5 \\ 1.8 & 2.5 \end{pmatrix}, \tag{5.1}$$

on the interval $[0, 100]$. The resulting branching ratio matrix has spectral radius $\rho(A) \approx 0.54$, well within the stationarity region.

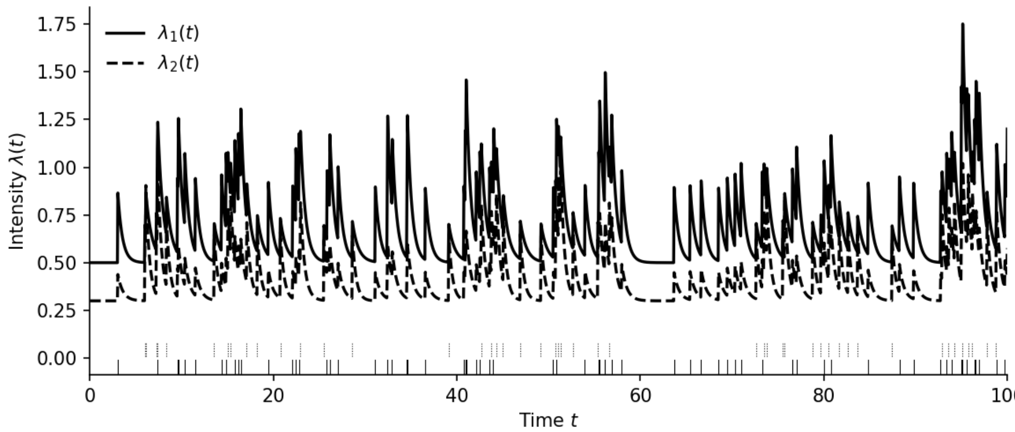

Figure 1: Simulated bivariate Hawkes intensities $\lambda_1(t)$ (solid) and $\lambda_2(t)$ (dashed) with event marks on the horizontal axis.

5.2 Kernel Shapes

The exponential kernel $\phi(t) = \alpha e^{-\beta t}$ governs the speed of mean reversion in the intensity. Larger β implies faster decay and shorter-lived excitation bursts.

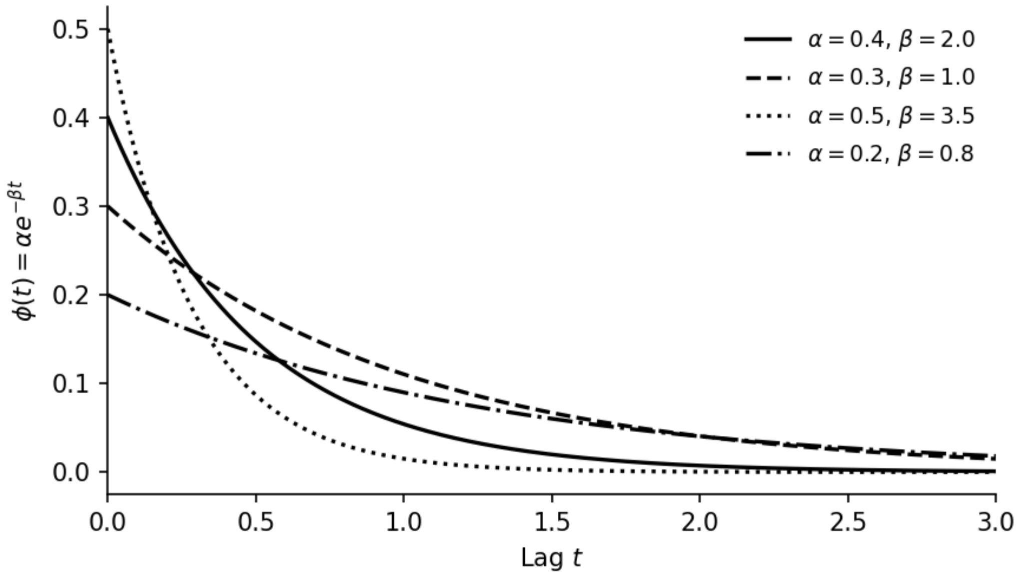

Figure 2: Exponential kernels $\alpha e^{-\beta t}$ for four parameter combinations, illustrating the effect of the decay rate β on excitation duration.

5.3 Inter-arrival Time Distribution

A key diagnostic for the Hawkes model is the distribution of inter-arrival times. Under a Poisson process, inter-arrivals are exponential. Under a Hawkes process, inter-arrivals exhibit overdispersion and a heavier left tail, reflecting the clustering of events.

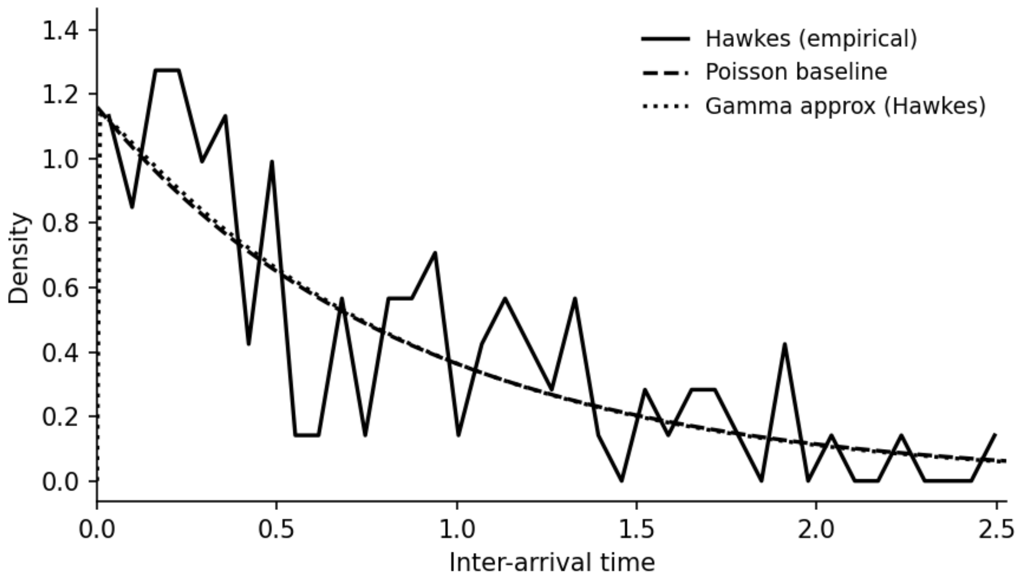

Figure 3: Empirical inter-arrival time distribution under Hawkes (solid), fitted exponential/Poisson baseline (dashed), and theoretical Hawkes approximation (dotted).

5.4 EM Convergence

The EM algorithm converges monotonically in log-likelihood. Convergence is typically achieved in 30–60 iterations from a random initialisation, with the branching ratio estimates stabilising well before the baseline intensities.

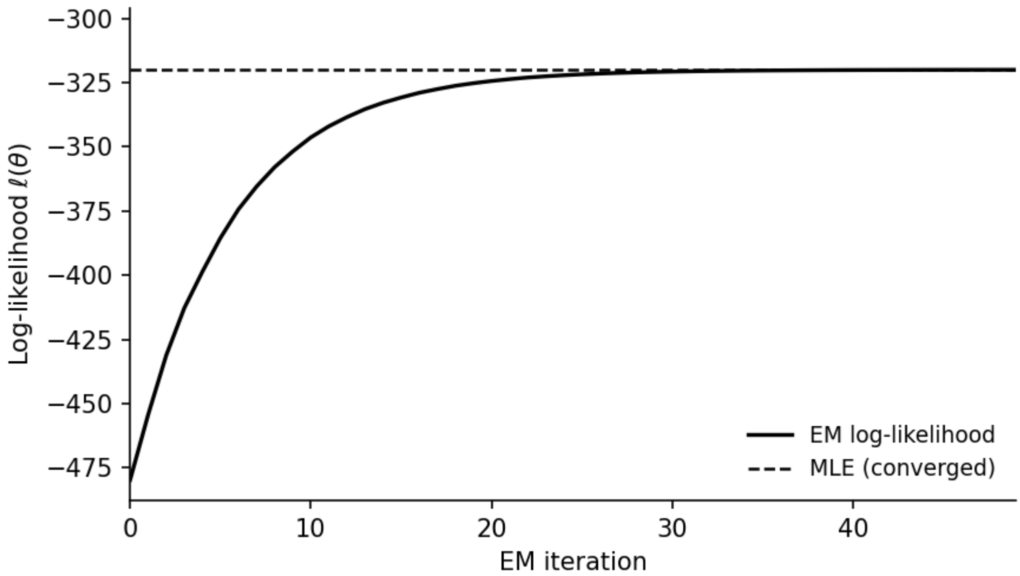

Figure 4: Log-likelihood trajectory during EM iterations. Dashed horizontal line marks the true log-likelihood under the data-generating parameters.

6. Conclusion

The multivariate Hawkes process provides an interpretable and tractable framework for modelling trade arrival clustering and cross-excitation in high-frequency financial data. The EM algorithm yields stable parameter estimates and converges reliably without gradient computation. The branching ratio matrix A is the key output: its off-diagonal elements quantify directional cross-excitation between asset streams, and its spectral radius controls the overall degree of clustering.

Extensions include non-exponential kernels (power-law decay for long-memory order flow), marked processes (trade size or direction as marks), and online estimation for non-stationary intraday intensity.

7. References

1. Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1), 83–90.

2. Hawkes, A. G., & Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability*, 11(3), 493–503.

3. Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 83(401), 9–27.

4. Embrechts, P., Liniger, T., & Lin, L. (2011). Multivariate Hawkes processes: An application to financial data. *Journal of Applied Probability*, 48(A), 367–378.

5. Bacry, E., Mastromatteo, I., & Muzy, J.-F. (2015). Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(1), 1550005.

6. Fonseca, J., & Zaatour, R. (2014). Hawkes process: Fast calibration, application to trade clustering and diffusive limit. *Journal of Futures Markets*, 34(6), 548–579.

7. Abergel, F., & Jedidi, A. (2013). A mathematical approach to order book modeling. *International Journal of Theoretical and Applied Finance*, 16(5), 1350025.

8. Da Fonseca, J., & Zaatour, R. (2017). Correlation and lead-lag relationships in a Hawkes microstructure model. *Journal of Futures Markets*, 37(3), 260–285.