Multivariate Hawkes Processes and Their Simulations

Yuanda Chen

September, 2016

Abstract

In this article we will extend our discussion to the multivariate Hawkes processes, which are mutually exciting processes. Except for the self-excitation feature inherited from the univariate case, they also have the cross-exciting between different dimensions, which can be readily seen from the definition below. The simulation of such processes can also be performed by thinning, as given by (Ogata, 1981), similarly to the univariate case, with only one extra step that decides which dimension an accepted point belongs to.

1 Definition

Definition 1.1. *(Hawkes, 1971, p.86, (19) and (20)) Let* $\boldsymbol{N}(t) = \left(N^1(t), N^2(t), \ldots, N^M(t)\right)$ *be a simple multivariate point process satisfying*

(i) $N(0) = 0$*, and for* $m = 1, 2, \ldots, M$*,*

(ii) $\lambda^m(t)$ *is a left-continuous stochastic process given by the Stieltjes integral*

$$\begin{aligned} \lambda^m(t) &= \mu_m + \sum_{n=1}^{M} \int_0^t \alpha_{mn} e^{-\beta_{mn}(t-s)} dN^n(s) \\ &= \mu_m + \sum_{n=1}^{M} \sum_{\{k:t_k^n<t\}} \alpha_{mn} e^{-\beta_{mn}(t-t_k^n)} \end{aligned} \tag{1}$$

where $\mu_m > 0$*,* $\alpha_{mn} \geq 0$ *and* $\beta_{mn} \geq 0$ *for* $m, n = 1, 2, \ldots, M$*.*

(iii) $\lambda^m(t)$*, independently for each* m*, is the stochastic intensity of the marginal point process* $N^m(t)$

$$P\{N^m(t+h) - N^m(t) = 1 | \mathcal{F}_{t^-}^{\boldsymbol{N}}\} = \lambda^m(t)h + o(h)$$

(iv) The point process is orderly

$$P\{N(t+h) - N(t) \geq 2 | \mathcal{F}_{t^-}^{\boldsymbol{N}}\} = o(h)$$

where $\mathcal{F}_t^{\boldsymbol{N}}$ *is the natural filtration of the process, is called an* M*-variate Hawkes process with exponential decays on* $[0, \infty)$*.*

Remark 1.1. *The intensity for a multivariate Hawkes process on* $(-\infty, \infty)$ *can be written in the vector form as (Hawkes, 1971, p.86, (20))*

$$\boldsymbol{\lambda}(t) = \boldsymbol{\mu} + \int_{-\infty}^{t} \boldsymbol{\gamma}(t-u) d\boldsymbol{N}(u).$$

where $\boldsymbol{\gamma}$ *is an* M*-by-*M *matrix. Assuming stationarity, we have (Hawkes, 1971, p.87, (21))*

$$\boldsymbol{\lambda} = E[\boldsymbol{\lambda}(t)] = (\boldsymbol{I} - \boldsymbol{\Gamma})^{-1} \boldsymbol{\mu}$$

where

$$\mathbf{\Gamma} = \int_0^\infty \boldsymbol{\gamma}(u)du$$

and moreover, when the elements of $\boldsymbol{\gamma}$ *are* $\gamma_{mn}(u) = \alpha_{mn}e^{-\beta_{mn}u}$,

$$\mathbf{\Gamma} = \left\{\frac{\alpha_{mn}}{\beta_{mn}}\right\}_{m,n=1,2,\ldots,M}.$$

A sufficient condition for the process to be stationary is that $\rho(\mathbf{\Gamma}) < 1$, *where* $\rho(\mathbf{\Gamma})$ *is the spectral radius of* $\mathbf{\Gamma}$ *which is defined as*

$$\rho(\mathbf{\Gamma}) = \max_{x\in\mathcal{E}(\mathbf{\Gamma})} |x|$$

where $\mathcal{E}(\mathbf{\Gamma})$ *is the set of eigenvalues of* $\mathbf{\Gamma}$. *Moreover, when* $\rho(\mathbf{\Gamma}) < 1$ *the existence and uniqueness of such a simple multivariate point process, as defined in Definition 1.1, are guaranteed by (Abergel et al., 2016, p.104, Proposition 8.3) or (Jedidi and Abergel, 2013, p.5, Proposition 2.4).*

Figure 1 shows an example sample path of the intensity (upper panels) and associated counting (lower panels) processes, for a bivariate Hawkes process with exponential decays, with the following parameters

$$\boldsymbol{\mu} = \begin{pmatrix}0.1\\0.5\end{pmatrix}, \qquad \boldsymbol{\alpha} = \begin{pmatrix}0.1 & 0.7\\0.5 & 0.2\end{pmatrix}, \qquad \text{and} \quad \boldsymbol{\beta} = \begin{pmatrix}1.2 & 1.0\\0.8 & 0.6\end{pmatrix}.$$

The first four points in each dimension are plotted, however $\lambda^1(t)$ (top-left panel) and $\lambda^2(t)$ (top-right panel) each has eight jumps, because of the built-in cross-excitation in (1). For example, the first jump of $\lambda^2(t)$ occurring at t_1^1 is caused by the first point in $N^1(t)$, so the jump size is determined by $\alpha_{21} = 0.5$; while the self-excitation effect from this point, causes $\lambda^1(t)$ to jump, also at t_1^1, with a size of $\alpha_{11} = 0.1$.

2 Simulation

In (Ogata, 1981) the author gave a proposition that states the simulation of multivariate point processes, by distributing the accepted points to each dimension with probabilities proportional to their intensities.

Proposition 2.1. *(Ogata, 1981, p.24, Proposition 1) Let* $\boldsymbol{N}(t) = \left(N^1(t), N^2(t), \ldots, N^M(t)\right)$ *be an* M*-variate point process on an interval* $[0,T]$ *with stochastic intensities* $\lambda^{m*}(t) = \lambda^m(t|\mathcal{F}_{t^-}^{\boldsymbol{N}})$ *for* $m = 1,2,\ldots,M$. *Suppose there is a one-dimensional* $\mathcal{F}_{t^-}^{\boldsymbol{N}}$*-predictable process* $\bar{\lambda}^*(t)$ *which is defined path-wisely satisfying*

$$\sum_{m=1}^{M} \lambda^{m*}(t) \le \bar{\lambda}^*(t), \qquad 0 < t \le T$$

and set

$$\lambda^{0*}(t) = \bar{\lambda}^*(t) - \sum_{m=1}^{M} \lambda^{m*}(t).$$

Let $\bar{t}_1, \ldots, \bar{t}_{\bar{N}(T)} \in (0,T]$ *be the points of the process* $\bar{N}(t)$ *with stochastic intensity* $\bar{\lambda}^*(t)$. *For each of the points,* $\bar{t}_k$ *where* $k = 1,2,\ldots,\bar{N}(T)$ *attach a mark* $m = 0,1,\ldots,M$ *with probability* $\lambda^{m*}(\bar{t}_k)/\bar{\lambda}^*(\bar{t}_k)$,*respectively. Then the points with marks* $m = 1,2,\ldots,M$, *provide an* M*-variate point process with stochastic intensities* $\lambda^{m*}(t)$.

Remark 2.1. *The sketch of the proof for Proposition 2.1 is given below. For the technical details please refer to (Jacod, 1975) and (Kabanov et al., 1976). Define a random measure for the* M*-variate point process by*

$$M(dt, m) = \bar{N}(dt) \cdot I(t, m)$$

for $m = 1,2,\ldots,M$, *where* $I(t,m)$ *is the random measure of the marks under the condition that there is a point at* t. *By the conditions of the proposition,* $I(t,m)$ *has the following properties.*

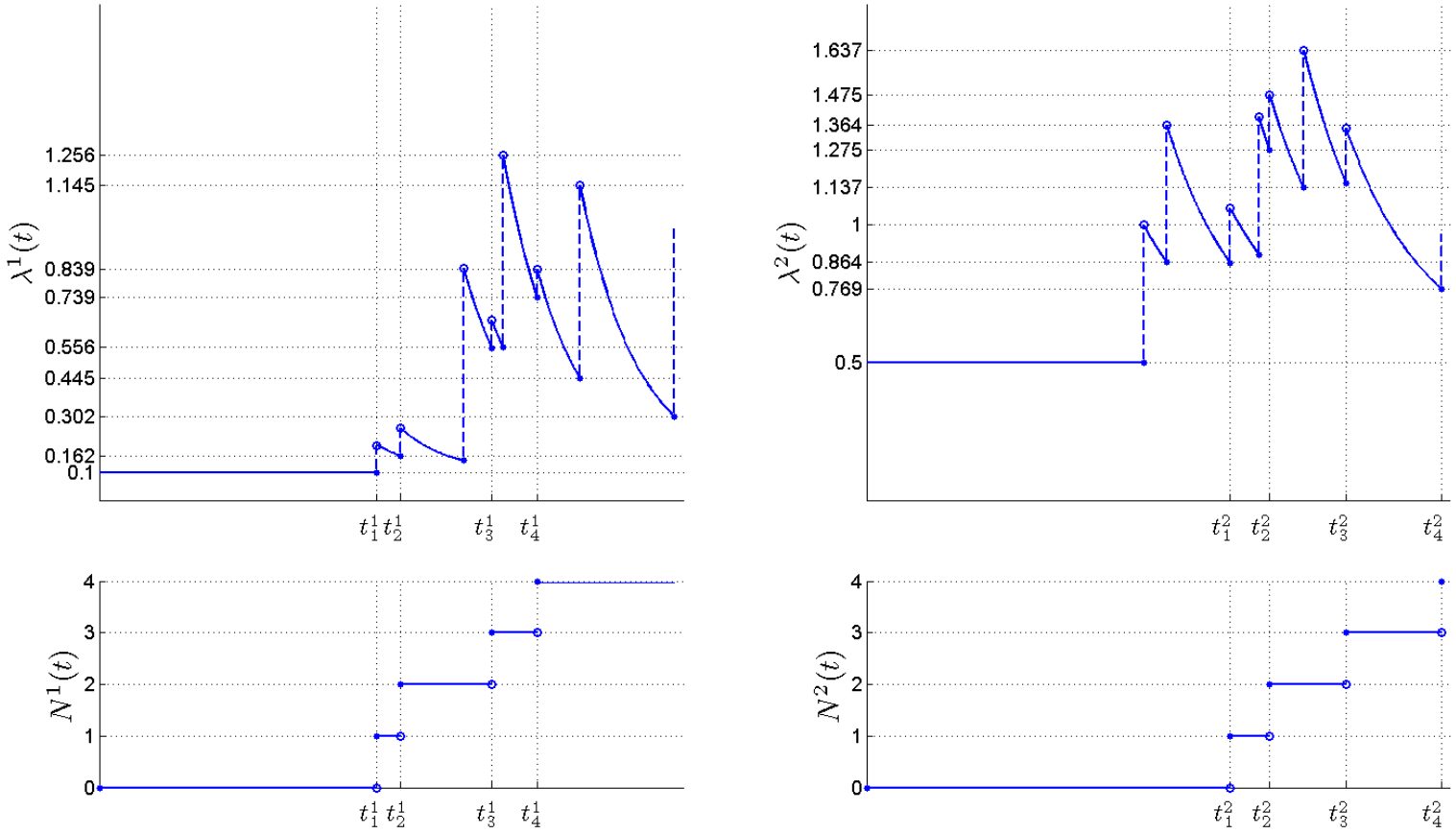

Figure 1: An illustrative example of the conditional intensity and the associated counting processes for a bivariate Hawkes process with exponential decays, with parameters $\boldsymbol{\mu} = \begin{pmatrix} 0.1 \\ 0.5 \end{pmatrix}$, $\boldsymbol{\alpha} = \begin{pmatrix} 0.1 & 0.7 \\ 0.5 & 0.2 \end{pmatrix}$ and $\boldsymbol{\beta} = \begin{pmatrix} 1.2 & 1.0 \\ 0.8 & 0.6 \end{pmatrix}$. Notice both the self- and cross-excitation effects in the intensity processes $\lambda^1(t)$ and $\lambda^2(t)$. For example, the first jump of $\lambda^2(t)$ occurring at t_1^1 is caused by the first point in $N^1(t)$, so the jump size is determined by $\alpha_{21} = 0.5$; while the self-excitation effect from this point, causes $\lambda^1(t)$ to jump, also at t_1^1, with a size of $\alpha_{11} = 0.1$.

(i) $I(t,m) = \delta_p(m)$ *with probability* $\lambda^{m*}(t)/\bar{\lambda}^*(t)$ *for* $m, p = 0, 1, \ldots, M$, *where* $\delta_p(m)$ *is a Dirac measure.*

(ii) For fixed t, $\bar{N}(dt)$ *and* $\{I(t,m)\}_{m=0,1,\ldots,M}$ *are conditionally independent given* $\mathcal{F}_{t^-}^{\boldsymbol{N}}$.

Then for each mark $m = 1, 2, \ldots, M$, *the intensity measure of the marked point process is given by*

$$\begin{aligned}\nu(dt,m) &= E[M(dt,m)|\mathcal{F}_{t^-}^{\boldsymbol{N}}] \\ &= E[\bar{N}(dt)\cdot I(t,m)|\mathcal{F}_{t^-}^{\boldsymbol{N}}] \\ &= E[\bar{N}(dt)|\mathcal{F}_{t^-}^{N}]\cdot E[I(t,m)|\mathcal{F}_{t^-}^{\boldsymbol{N}}]\end{aligned}$$

By discussion in Remark **??**,

$$E[\bar{N}(dt)|\mathcal{F}_{t^-}^{\boldsymbol{N}}] = \bar{\lambda}^*(t)dt$$

and also

$$\begin{aligned}E[I(t,m)|\mathcal{F}_{t^-}^{\boldsymbol{N}}] &= 1\cdot P\{I(t,m)=1|\mathcal{F}_{t^-}^{\boldsymbol{N}}\} + 0\cdot P\{I(t,m)=0|\mathcal{F}_{t^-}^{\boldsymbol{N}}\} \\ &= \lambda^{m*}(t)/\bar{\lambda}^*(t)\end{aligned}$$

Therefore for each $m = 1, 2, \ldots, M$, *we get* $\nu(dt,m) = \lambda^{m*}(t)dt$. *Since the predictable random measure corresponds uniquely to the multivariate process (see (Jacod, 1975) and (Kabanov et al., 1976)), this completes the proof.*

The simulation of a multivariate Hawkes process with exponential decays on a fixed interval is similar to the univariate case, with only one extra step that decides which dimension an accepted point belongs to. By Proposition 2.1, given that a point is accepted at time s, it should be distributed to dimension m with probability $\lambda^m(s)/\sum_{m=1}^M \lambda^m(s)$, for $m = 1, 2, \ldots, M$. The procedure for simulating an M-variate Hawkes process, defined in Definition 1.1, on the interval $[0, T]$ is summarized below, where $\mathcal{T}^m$ represents the ordered set of accepted points in dimension m and n^m counts the number of points in $\mathcal{T}^m$, for $m = 1, 2, \ldots, M$. As before, s is always the newest candidate point generated.

1. Set $\mathcal{T}^1 = \cdots = \mathcal{T}^M = \emptyset$, $s = 0$ and $n^1 = \cdots = n^M = 0$.
2. Repeat the following until $s > T$.
   (1) Set $\bar{\lambda} = \sum_{m=1}^M \lambda^m(s^+) = \sum_{m=1}^M \left(\mu_m + \sum_{n=1}^M \sum_{\tau\in\mathcal{T}^n} \alpha_{mn} e^{-\beta_{mn}(s-\tau)}\right)$
   (2) Generate u from a uniform distribution on $[0, 1]$.
   (3) Generate $w = -\ln u/\bar{\lambda}$ as the interarrival to the next candidate point.
   (4) Set the new candidate point $s = s + w$.
   (5) Generate u from a uniform distribution on $[0, 1]$.
      (i) If $D \le \sum_{m=1}^M \lambda^m(s)/\bar{\lambda}$ (where $\lambda^m(s) = \mu_m + \sum_{n=1}^M \sum_{\tau\in\mathcal{T}^n} \alpha_{mn} e^{-\beta_{mn}(s-\tau)}$), then do the following:
         (a) Find $k \in \{1, 2, \ldots, M\}$ such that $\sum_{m=1}^{k-1} \lambda^m(s) < D\bar{\lambda} \le \sum_{m=1}^k \lambda^m(s)$.
         (b) Assign candidate point s to dimension k by setting $n^k = n^k + 1$, $t_{n^k}^k = s$ and $\mathcal{T}^k = \mathcal{T}^k \bigcup \{t_{n^k}^k\}$.
      (ii) else, do nothing.
3. If $t_{n^k}^k > T$, then $\mathcal{T}^1, \ldots, \mathcal{T}^k \backslash \{t_{n^k}^k\}, \ldots, \mathcal{T}^M$ contain the simulated points for each dimension;
4. else $\mathcal{T}^m$ for $m = 1, 2, \ldots, M$ contain the simulated points.

Algorithm 1 is the pseudo-code for the simulation procedure described above, where the input $\boldsymbol{\mu}$ is a vector with elements μ_m for $m = 1, 2, \ldots, M$ and both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are M-by-M matrices. Comparing to the univariate case, the main change is the newly added "while" loop at line 10, where k loops through 1 to M and stops at the first k such that $D\bar{\lambda} \le \sum_{m=1}^k \lambda^m(s)$. Since this k is the first such k, we actually

have $\sum_{m=1}^{k-1} \lambda^m(s) < D\bar{\lambda} \leq \sum_{m=1}^{k} \lambda^m(s)$ and dimension k is where this point belongs to. Notice that this "while" loop eventually stops because we only reach line 10 if the condition in line 8 is satisfied, that is $D\bar{\lambda} \leq \sum_{m=1}^{M} \lambda^m(s)$. Also notice that the use of s^+ in summation of intensities at line 3 is meant to cover both the cases whether s was accepted as a point in some dimension k or not; while the summation of intensities is taken at s in line 8, because that is where we try to decide whether to accept s. The "if" statement at line 18 is used to check whether the last accepted point is out of $[0, T]$. We only need to check $t^k_{n^k}$ because whenever the stopping criterion at line 2 is met, k still stores the information of the last accepted point.

Algorithm 1: Simulation of an M-variate Hawkes Process with Exponential Kernels $\gamma_{mn}(u) = \alpha_{mn} e^{-\beta_{mn} u}$ for $m, n = 1, 2, \ldots, M$, on $[0, T]$.

Input: $\boldsymbol{\mu}_{M\times 1}$, $\boldsymbol{\alpha}_{M\times M}$, $\boldsymbol{\beta}_{M\times M}$, T

1 Initialize $\mathcal{T}^1 = \cdots = \mathcal{T}^M = \emptyset$, $n^1 = \cdots = n^M = 0$, $s = 0$;
2 **while** $s < T$ **do**
3 Set $\bar{\lambda} = \sum_{m=1}^{M} \lambda^m(s^+) = \sum_{m=1}^{M} \left(\mu_m + \sum_{n=1}^{M} \sum_{\tau \in \mathcal{T}^n} \alpha_{mn} e^{-\beta_{mn}(s-\tau)}\right)$;
4 Generate $u \sim$ `uniform(0,1)`;
5 Let $w = -\ln u / \bar{\lambda}$; `// so that` $w \sim$ `exponential(`$\bar{\lambda}$`)`
6 Set $s = s + w$; `// so that` s `is the next candidate point`
7 Generate $D \sim$ `uniform(0,1)`;
`// accepting with probability` $\sum_{m=1}^{M} \lambda^m(s)/\bar{\lambda}$
8 **if** $D\bar{\lambda} \leq \sum_{m=1}^{M} \lambda^m(s) = \sum_{m=1}^{M} \left(\mu_m + \sum_{n=1}^{M} \sum_{\tau \in \mathcal{T}^n} \alpha_{mn} e^{-\beta_{mn}(s-\tau)}\right)$ **then**
9 $k = 1$;
`// searching for the first` k `such that` $D\bar{\lambda} \leq \sum_{m=1}^{k} \lambda^m(s)$
10 **while** $D\bar{\lambda} > \sum_{m=1}^{k} \lambda^m(s)$ **do**
11 $k = k + 1$
12 **end**
13 $n^k = n^k + 1$; `// updating the number of points in dimension` k
14 $t^k_{n^k} = s$; `// naming it` $t^k_{n^k}$
15 $\mathcal{T}^k = \mathcal{T}^k \bigcup \{t^k_{n^k}\}$; `// adding` $t^k_{n^k}$ `to the ordered set` $\mathcal{T}^k$
16 **end**
17 **end**
18 **if** $t^k_{n^k} \leq T$ **then**
19 **return** $\mathcal{T}^m$ *for* $m = 1, 2, \ldots, n$;
20 **else**
21 **return** $\mathcal{T}^1, \ldots, \mathcal{T}^k \backslash \{t^k_{n^k}\}, \ldots, \mathcal{T}^M$;
22 **end**

Figure 2 shows an example of the simulation of a bivariate Hawkes process with exponential decays, with the same parameters as in Figure 1. The lowest panel plots the values of D (line 7 of Algorithm 1) generated for each candidate point. The corresponding values $D\bar{\lambda}$ used for the "if" statement at line 8 are plotted in the third panel, marked either with circles, when $D\bar{\lambda} > \lambda^1(t) + \lambda^2(t)$ (pink), indicating rejections; crosses, when $\lambda^1(t) < D\bar{\lambda} \leq \lambda^1(t) + \lambda^2(t)$ (yellow), indicating acceptances to $\mathcal{T}^2$; or asterisks, when $D\bar{\lambda} \leq \lambda^1(t)$ (red), indicating acceptances to $\mathcal{T}^1$). The top two panels plot the accepted points in N^1 (1 point) and N^2 (3 points), together with the resulting intensities $\lambda^1(t)$ and $\lambda^2(t)$.

References

Abergel, F., M. Anane, A. Chakraborti, A. Jedidi, and I. M. Toke
2016. *Limit Order Books (Physics of Society: Econophysics and Sociophysics)*. Cambridge University Press.

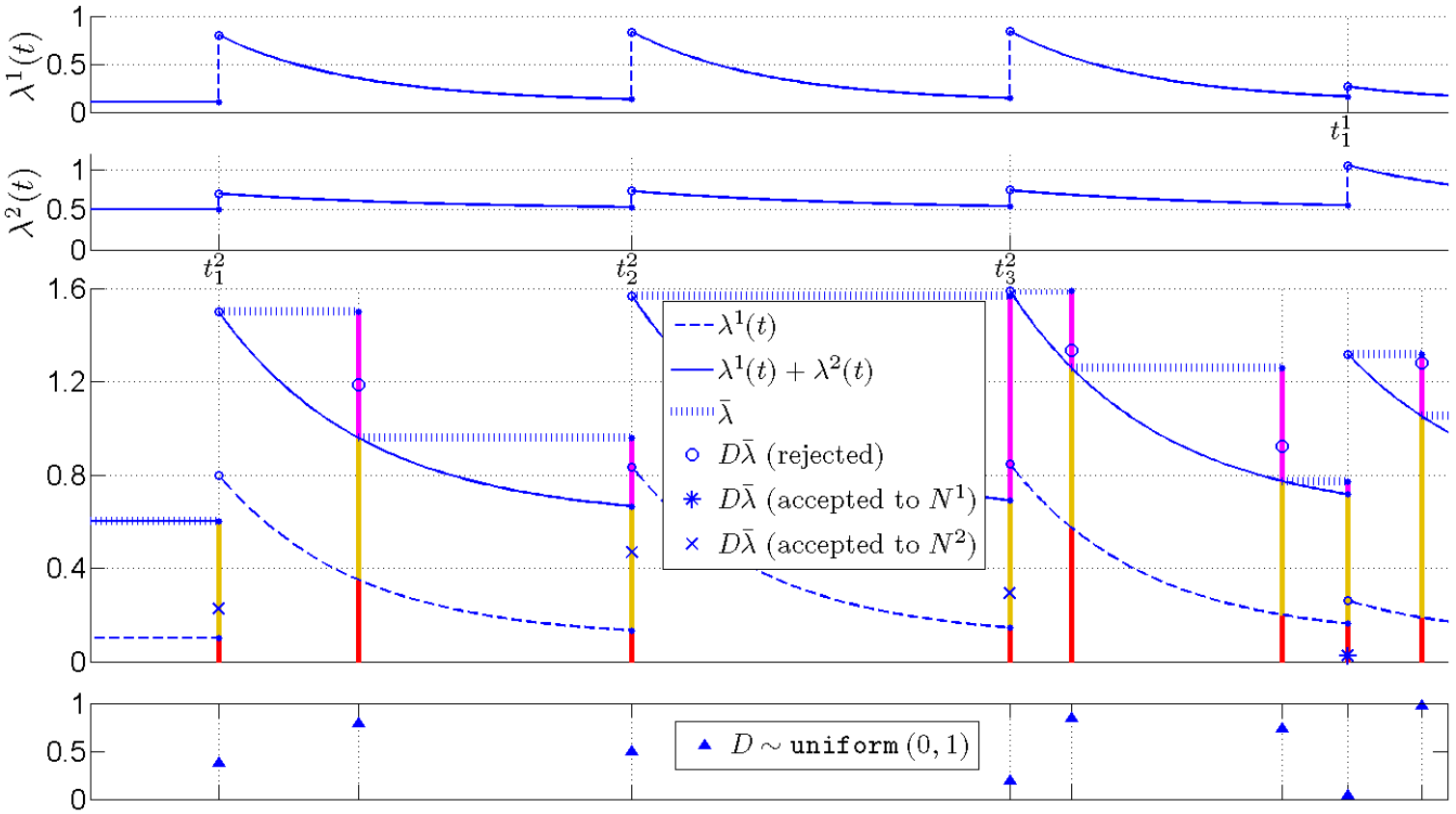

Figure 2: An illustrative example of simulating a bivariate Hawkes process with exponential decays, with parameters $\boldsymbol{\mu} = \begin{pmatrix} 0.1 \\ 0.5 \end{pmatrix}$, $\boldsymbol{\alpha} = \begin{pmatrix} 0.1 & 0.7 \\ 0.5 & 0.2 \end{pmatrix}$ and $\boldsymbol{\beta} = \begin{pmatrix} 1.2 & 1.0 \\ 0.8 & 0.6 \end{pmatrix}$. The lowest panel plots the values of D (line 7 of Algorithm 1) generated for each candidate points. The corresponding values $D\bar{\lambda}$ used for the "if" statement at line 8 are plotted in the third panel, marked either with circles, when $D\bar{\lambda} > \lambda^1(t) + \lambda^2(t)$ (pink), indicating rejections; crosses, when $\lambda^1(t) < D\bar{\lambda} \leq \lambda^1(t) + \lambda^2(t)$ (yellow), indicating acceptances to $\mathcal{T}^2$; or asterisks, when $D\bar{\lambda} \leq \lambda^1(t)$ (red), indicating acceptances to $\mathcal{T}^1$). The top two panels plot the accepted points in N^1 (1 point) and N^2 (3 points), together with the resulting intensities $\lambda^1(t)$ and $\lambda^2(t)$.

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

Jacod, J.
1975. Multivariate point processes: predictable projection, radon-nikodym derivatives, representation of martingales. *Z. Wahrscheinlichkeitstheorie verw Gebiete*, 31(3):235–253.

Jedidi, A. and F. Abergel
2013. Stability and price scaling limit of a hawkes-process based order book model. Working papers, HAL.

Kabanov, Y. M., R. S. Lipcer, and A. N. Širyaev
1976. Criteria of absolute continuity of measures corresponding to multivariate point processes. In *Proceedings of the Third Japan — USSR Symposium on Probability Theory*, Pp. 232–252. Springer Nature.

Ogata, Y.
1981. On lewis' simulation method for point processes. *IEEE Trans. Inform. Theory*, 27(1):23–31.