SUPPLEMENTARY MATERIAL: This is the supplementary material of "Mutually exciting point processes with latency" by Yoann Potiron and Vladimir Volkov published in the Journal of the American Statistical Association. Our numerical study is carried over in Supplement A. Examples are given in the Supplement B. All proofs of the theory are shown in Supplement C. Additional empirical results belong to Supplement D.

Appendices

A Numerical study

The performance of the model is now explored via a simple multidimensional simulation experiment. Consider the 5-dimensional specification of Equation (4) with intensity given by

$$\lambda^{(i)}(t,\theta^*) = n\nu^{*,(i)} + \sum_{j=1}^{5} \int_0^{t^-} nh^{(i,j)}(n(t-s),\theta_{ker}^{*,(i,j)})dN_s^{(j)}, \tag{A1}$$

where $\nu^* = \begin{bmatrix} 0.01 & 0.01 & 0.01 & 0.01 & 0.01 \end{bmatrix}'$, $h(t,\theta)$ is a gamma kernel defined in Equation (B2) and the kernel parameters θ_{ker}^* are chosen such that

$$h(t,\theta_{ker}^*) =$$

$$\begin{bmatrix} 0.15\frac{t^8 exp(-t)}{\Gamma(9)} & 0.16\frac{t^{10} exp(-t/1.5)}{\Gamma(11)} & 0.14\frac{t^{10} exp(-t/1.2)}{\Gamma(11)} & 0.25\frac{t^5 exp(-t/2)}{\Gamma(6)} & 0.14\frac{t^9 exp(-t/1.1)}{\Gamma(10)} \\ 0.14\frac{t^5 exp(-t/2)}{\Gamma(6)} & 0.15\frac{t^{11} exp(-t/2)}{\Gamma(12)} & 0.24\frac{t^{10} exp(-t/1.5)}{\Gamma(11)} & 0.15\frac{t^9 exp(-t/2.1)}{\Gamma(10)} & 0.25\frac{t^8 exp(-t/1.7)}{\Gamma(9)} \\ 0.14\frac{t^7 exp(-t/1.8)}{\Gamma(8)} & 0.25\frac{t^8 exp(-t/1.2)}{\Gamma(9)} & 0.15\frac{t^9 exp(-t/1.6)}{\Gamma(10)} & 0.24\frac{t^{10} exp(-t/1.8)}{\Gamma(11)} & 0.15\frac{t^9 exp(-t/1.6)}{\Gamma(10)} \\ 0.25\frac{t^8 exp(-t/2)}{\Gamma(9)} & 0.14\frac{t^7 exp(-t/1.2)}{\Gamma(8)} & 0.14\frac{t^6 exp(-t/1.3)}{\Gamma(7)} & 0.24\frac{t^8 exp(-t/1.5)}{\Gamma(9)} & 0.16\frac{t^7 exp(-t/2)}{\Gamma(8)} \\ 0.24\frac{t^9 exp(-t/1.5)}{\Gamma(10)} & 0.15\frac{t^{5.5} exp(-t/2)}{\Gamma(6.5)} & 0.14\frac{t^6 exp(-t/1.5)}{\Gamma(7)} & 0.26\frac{t^{5.6} exp(-t/2.1)}{\Gamma(6.6)} & 0.15\frac{t^7 exp(-t/1.8)}{\Gamma(8)} \end{bmatrix}. \tag{A2}$$

From Equation (B2), we can deduce that the Hawkes process generated by the intensity (A1) has a baseline equal to $\widetilde{\nu}^{*,(i)} = n\nu^{*,(i)}$ and a gamma kernel $\widetilde{h}(t,\theta)$ with true value parameters

$(n\alpha^{(i,j)}, \frac{\beta^{(i,j)}}{n}, D^{(i,j)})$. Thus, there is a unique relationship between (ν^*, h) and $(\widetilde{\nu}^*, \widetilde{h})$. The choice of parameter values mimics the broad characteristics of the empirical data discussed in Section 5. From Equation (B3), latency and co-latency are obtained as a function of parameters equal to $\widetilde{L}^{(i,j)} = \frac{\beta^{(i,j)}(D^{(i,j)}-1)}{n}$. Thus, we obtain latency and co-latency values below 15 milliseconds. The simulation exercise involves 500 independent replications with the sample size order $n = 100,000$ and setting up $T = 1$ trading interval to generate the data. Since we recall that n corresponds to the order of the number of observations, we note that the sample size order $n = 100,000$ is more conservative than the average number of observations in our empirical study, see Table 1. The elements of the kernel $h(t, \theta^*_{ker})$ are estimated using the MLE approach presented in Section 3.1. The model defined in Equation (A1) meets the theoretical assumptions from Section 4.

Figure A1 illustrates histograms of the kernel estimates $\hat{h}$. The trapezoidal rule is used to numerically compute the integral in Equation (A1). The estimates of all kernel functions are close to their theoretical values and the confidence intervals behave as expected. One case where the confidence intervals are especially narrow is represented by $\hat{h}^{(3,2)}$. This is expected as the kernel specification has the highest mean rate. All kernels approach zero within 50 intervals.

Now we verify the CLT of the latency estimator. Histograms of latency and co-latency estimates $\widehat{L_T^{(i,j)}}$ are presented in Figure A2. The obtained variance lies within the range of 2-3 intervals. Table A1 verifies the finite sample properties. The bias ranges from -0.10 to -0.30, which only affects around 1% of latency values. The estimated latency has relatively similar in sample standard error, which is calculated from $(\widehat{L_T^{(i,j)}} - L^{(i,j)})$, and estimated standard error $\sqrt{\text{Cov}\left[\widehat{\eta^{(i,j)}, \eta^{(k,l)}}\right]}$. This indicates that our variance estimator performs reasonably well with increasing n. To confirm the behavior of variance estimators and asymptotic Gaussianity, we provide histograms of the standardized errors related to the latency estimator $\widehat{L_T^{(i,j)}}$ in Figure A3.

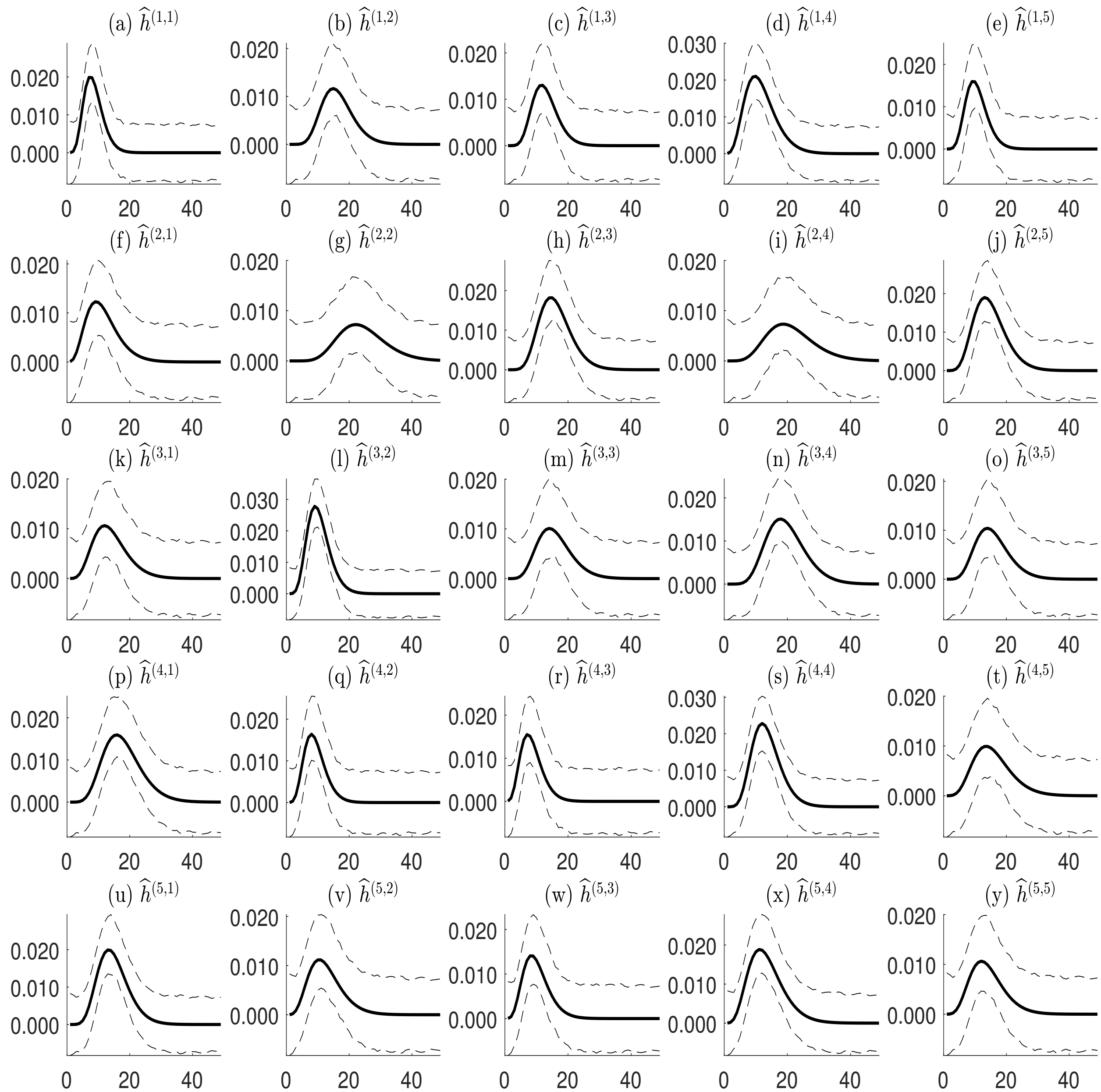

Figure A1: Histograms of the kernel estimates $\hat{h}$ over 50 time intervals. The solid line represents the estimated kernels. The confidence intervals (dashed line) are represented by 2.5% and 97.5% percentiles. The histograms are generated from 500 independent replications with the sample size order $n = 100,000$.

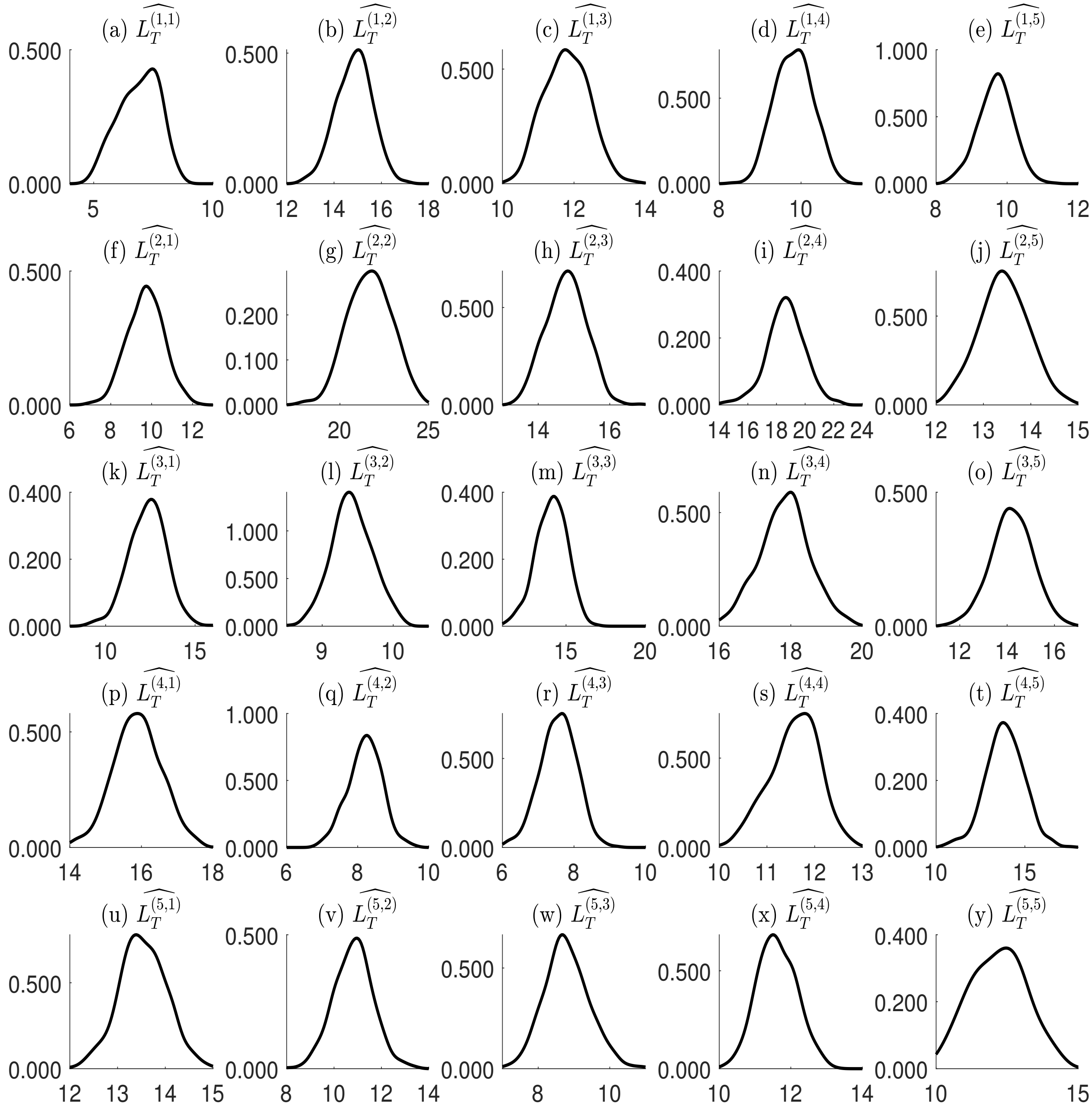

Figure A2: Histograms of latency and co-latency estimates $\widehat{L_T^{(i,j)}}$. The histograms are generated from 500 independent replications with the sample size order $n = 100,000$.

Table A1: Finite sample properties of the latency estimator $\widehat{L_T^{(i,j)}}$. The bias is computed as an average of $(\widehat{L_T^{(i,j)}} - L^{(i,j)})$, in sample standard error is calculated from $(\widehat{L_T^{(i,j)}} - L^{(i,j)})$, and the estimated standard error is $\sqrt{\text{Cov}\left[\widehat{\eta^{(i,j)}, \eta^{(k,l)}}\right]}$. The estimates are generated from 500 independent replications with the sample size order $n = 100,000$.

	Bias	In st. error	Est. st. error		Bias	In st. error	Est. st. error
$\widehat{L_T^{(1,1)}}$	-0.11	2.50	2.62	$\widehat{L_T^{(1,2)}}$	-0.15	2.25	3.85
$\widehat{L_T^{(1,3)}}$	-0.20	1.84	3.43	$\widehat{L_T^{(1,4)}}$	-0.18	1.35	3.13
$\widehat{L_T^{(1,5)}}$	-0.21	1.45	3.11	$\widehat{L_T^{(2,1)}}$	-0.30	2.63	3.11
$\widehat{L_T^{(2,2)}}$	-0.13	2.35	3.18	$\widehat{L_T^{(2,3)}}$	-0.22	1.65	3.84
$\widehat{L_T^{(2,4)}}$	-0.25	3.95	4.31	$\widehat{L_T^{(2,5)}}$	-0.16	1.54	3.66
$\widehat{L_T^{(3,1)}}$	-0.26	3.08	3.50	$\widehat{L_T^{(3,2)}}$	-0.18	1.84	3.06
$\widehat{L_T^{(3,3)}}$	-0.19	2.28	3.31	$\widehat{L_T^{(3,4)}}$	-0.17	2.08	4.22
$\widehat{L_T^{(3,5)}}$	-0.20	2.62	3.76	$\widehat{L_T^{(4,1)}}$	-0.11	1.98	3.98
$\widehat{L_T^{(4,2)}}$	-0.20	1.42	2.86	$\widehat{L_T^{(4,3)}}$	-0.24	1.51	2.75
$\widehat{L_T^{(4,4)}}$	-0.41	1.53	3.39	$\widehat{L_T^{(4,5)}}$	-0.13	3.20	3.72
$\widehat{L_T^{(5,1)}}$	-0.12	1.47	3.68	$\widehat{L_T^{(5,2)}}$	-0.18	2.46	3.28
$\widehat{L_T^{(5,3)}}$	-0.20	1.86	2.96	$\widehat{L_T^{(5,4)}}$	-0.15	1.63	3.40
$\widehat{L_T^{(5,5)}}$	-0.19	2.04	3.03				

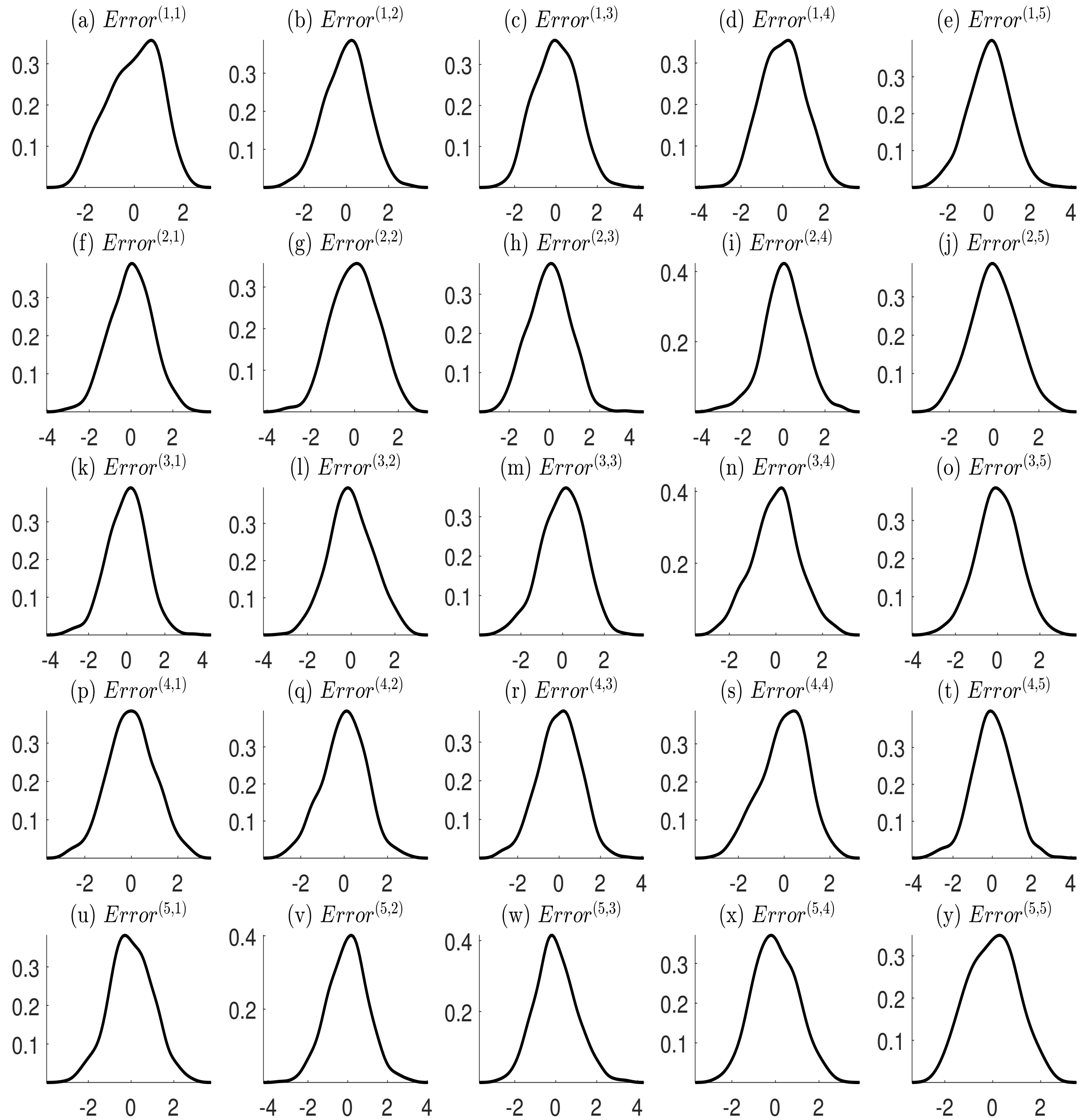

Figure A3: Histograms of the standardized errors for the latency estimator $\widehat{L_T^{(i,j)}}$. The histograms are generated from 500 independent replications with the sample size order $n = 100,000$.

Now we conduct hypothesis testing to confirm the size and power of the tests provided in Section 4.3 for different sample sizes. While the sample size order in our empirical application is larger than $n = 100,000$ providing a better size and power, we aim to demonstrate that the proposed latency test can be used in other areas of statistics. We compare the sample size orders of $n = 10,000$, $n = 50,000$ and $n = 100,000$. The sample size order of $n = 50,000$ is similar to a size of dataset from Ogata (1988) who considered earthquakes of magnitude 6 or more that occurred in Japan and its vicinity over almost 100 years. The significance level is set up at 5% level. Table A2 reports power and size at the 5% level of the tests: $H_A : L^{(1,1)} = 0$, $H_B : L^{(1,2)} = 0$, $H_C : L^{(1,3)} = 0$, $H_D : L^{(1,4)} = 0$, $H_E : L^{(1,5)} = 0$ against one-sided alternatives. For smaller samples the tests are slightly undersized in most cases but approach 5% for the sample size order of $n = 100,000$. The power is bigger than 0.93 in all cases and approaches 1. This highlights the potential of applying our method not only in finance but also in other areas of statistics such as seismology.

Table A2: Power and size at the 5% level of the tests: $H_A : L^{(1,1)} = 0$, $H_B : L^{(1,2)} = 0$, $H_C : L^{(1,3)} = 0$, $H_D : L^{(1,4)} = 0$, $H_E : L^{(1,5)} = 0$ against one-sided alternatives. 500 independent replications are used for simulation.

Null	H_A	H_B	H_C	H_D	H_E	H_A	H_B	H_C	H_D	H_E
Sample order			Size					Power		
10,000	0.040	0.016	0.010	0.044	0.011	0.932	0.944	0.964	0.956	0.948
50,000	0.032	0.036	0.060	0.036	0.044	0.938	0.951	0.952	0.958	0.952
100,000	0.039	0.044	0.046	0.048	0.046	0.974	0.996	1.000	0.959	1.000

B Examples

In this supplement, we provide five examples of kernels, i.e. exponential, gamma, Weibull, generalized gamma, and mixture of several kernels, which meet the assumptions of our framework. We insist on the fact that latency is not well-defined when the kernel is exponential as the mode is always equal to 0 in that case. For the remaining examples, the latency is defined as the mode of the kernel.

B.1 Exponential kernel

The conventional exponential kernel is defined as

$$h^{(i,j)}(t, \theta_{ker}^{(i,j)}) = \alpha^{(i,j)} \frac{exp(-t/\beta^{(i,j)})}{\beta^{(i,j)}}, \quad \alpha^{(i,j)} \in (0, h_+], \beta^{(i,j)} \in \mathbb{R}_+^*. \tag{B1}$$

This is a particular case of the generalized gamma kernel (5) when $p^{(i,j)} = D^{(i,j)} = 1$. On the one hand, the exponential kernel satisfies the assumptions of this paper, so it is a valid kernel form. On the other hand, we insist on the fact that latency is not well-defined when the kernel is exponential as the mode is always equal to 0 in that case. Thus, the exponential kernel is not suitable for estimating latency.

B.2 Gamma kernel

The gamma kernel is defined as

$$h^{(i,j)}(t, \theta_{ker}^{(i,j)}) = \alpha^{(i,j)} \frac{t^{(D^{(i,j)}-1)} exp(-t/\beta^{(i,j)})}{(\beta^{(i,j)})^{D^{(i,j)}} \gamma(D^{(i,j)})}, \quad \alpha^{(i,j)} \in (0, h_+], \beta^{(i,j)} \in \mathbb{R}_+^*, D^{(i,j)} \in \mathbb{R}_+^*. \tag{B2}$$

This is a particular case of the generalized gamma kernel (5) when $p^{(i,j)} = 1$. We define latency as the mode, which can be expressed as

$$L^{(i,j)} = \beta^{(i,j)}(D^{(i,j)} - 1). \tag{B3}$$

When $L^{(i,j)} > 0$, or equivalently $D^{(i,j)} > 1$, a latency between an event in process j and its impact on process i is introduced. When $L^{(i,j)} \leq 0$, or equivalently $D^{(i,j)} \leq 1$, there is no latency between an event in process j and its impact on process i. Figure B1 illustrates an example of gamma kernel defined in Equation (B2) and exponential kernel defined in Equation (B1) for 20 intervals.

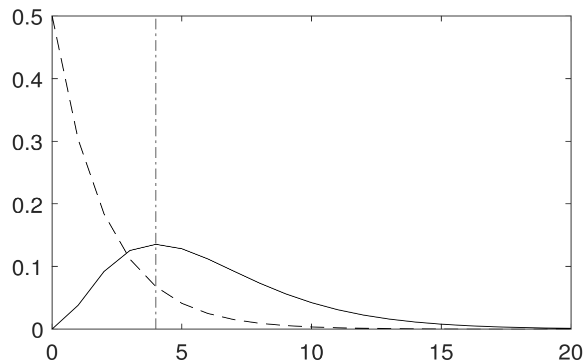

Figure B1: An example of gamma kernel defined in Equation (B2) and exponential kernel defined in Equation (B1) for 20 intervals. The solid line represents the gamma kernel with parameters $\alpha = 1$, $\beta = 2$, and $D = 3$ and the vertical line shows the latency $L = 4$. The dashed line represents the exponential kernel with $\alpha = 1$, $\beta = 2$, and $D = 1$.

B.3 Weibull kernel

The Weibull kernel is defined as

$$h^{(i,j)}(t, \theta_{ker}^{(i,j)}) = \alpha^{(i,j)} \frac{D^{(i,j)} t^{(D^{(i,j)}-1)} \exp(-(t/\beta^{(i,j)})^{D^{(i,j)}})}{(\beta^{(i,j)})^{D^{(i,j)}}}, \qquad \text{(B4)}$$
$$\alpha^{(i,j)} \in (0, h_+], \beta^{(i,j)} \in \mathbb{R}_+^*, D^{(i,j)} \in \mathbb{R}_+^*.$$

This is a particular case of the generalized gamma kernel (5) when $p^{(i,j)} = D^{(i,j)}$. We define latency as the mode, which can be expressed as

$$L^{(i,j)} = \beta^{(i,j)} \Big(\frac{D^{(i,j)} - 1}{D^{(i,j)}} \Big)^{1/D^{(i,j)}}. \qquad \text{(B5)}$$

B.4 Generalized gamma kernel

The generalized gamma kernel is defined as

$$h^{(i,j)}(t,\theta_{ker}^{(i,j)}) = \alpha^{(i,j)}\frac{p^{(i,j)}t^{(D^{(i,j)}-1)}\exp(-(t/\beta^{(i,j)})^{p^{(i,j)}})}{(\beta^{(i,j)})^{D^{(i,j)}}\gamma(D^{(i,j)}/p^{(i,j)})}, \alpha^{(i,j)} \in (0,h_+], \beta^{(i,j)}, D^{(i,j)}, p^{(i,j)} \in \mathbb{R}_+^* \quad \text{(B6)}$$

We define latency as the mode, which can be expressed as

$$L^{(i,j)} = \beta^{(i,j)}\Big(\frac{D^{(i,j)}-1}{p^{(i,j)}}\Big)^{1/p^{(i,j)}}. \quad \text{(B7)}$$

B.5 Mixture of several kernels

The mixture of exponential, gamma, Weibull and generalized gamma kernel is defined as

$$h^{(i,j)}(t,\theta_{ker}^{(i,j)}) = h_{exp}^{(i,j)}(t,\theta_{ker}^{(i,j)}) + h_{gam}^{(i,j)}(t,\theta_{ker}^{(i,j)}) + h_{Wei}^{(i,j)}(t,\theta_{ker}^{(i,j)}) + h_{gengam}^{(i,j)}(t,\theta_{ker}^{(i,j)}), \quad \text{(B8)}$$

where

$$\begin{aligned}
h_{exp}^{(i,j)}(t,\theta_{ker}^{(i,j)}) &= \sum_{k=1}^{K_1^{(i,j)}} \alpha_{k,1}^{(i,j)}\frac{\exp(-t/\beta_{k,1}^{(i,j)})}{\beta_{k,1}^{(i,j)}},\\
h_{gam}^{(i,j)}(t,\theta_{ker}^{(i,j)}) &= \sum_{k=1}^{K_2^{(i,j)}} \alpha_{k,2}^{(i,j)}\frac{t^{(D_{k,2}^{(i,j)}-1)}\exp(-t/\beta_{k,2}^{(i,j)})}{(\beta_{k,2}^{(i,j)})^{D_{k,2}^{(i,j)}}\gamma(D_{k,2}^{(i,j)})},\\
h_{Wei}^{(i,j)}(t,\theta_{ker}^{(i,j)}) &= \sum_{k=1}^{K_3^{(i,j)}} \alpha_{k,3}^{(i,j)}\frac{D_{k,3}^{(i,j)}t^{(D_{k,3}^{(i,j)}-1)}\exp(-(t/\beta_{k,3}^{(i,j)})^{D_{k,3}^{(i,j)}})}{(\beta_{k,3}^{(i,j)})^{D_{k,3}^{(i,j)}}},\\
h_{gengam}^{(i,j)}(t,\theta_{ker}^{(i,j)}) &= \sum_{k=1}^{K_4^{(i,j)}} \alpha_{k,4}^{(i,j)}\frac{p_{k,4}^{(i,j)}t^{(D_{k,4}^{(i,j)}-1)}\exp(-(t/\beta_{k,4}^{(i,j)})^{p_{k,4}^{(i,j)}})}{(\beta_{k,4}^{(i,j)})^{D_{k,4}^{(i,j)}}\gamma(D_{k,4}^{(i,j)}/p_{k,4}^{(i,j)})},
\end{aligned}$$

and we have that

$$\sum_{l=1}^{4}\sum_{k=1}^{K^{(i,j)}} \alpha_{k,l}^{(i,j)} \leq h_+.$$

A general formula for the mode in the mixture of several kernels case is beyond the scope of this paper.

C Proofs

In this supplement, we give all the proofs of the theoretical results from Section 4, namely Theorem 1, Proposition 2, Corollary 3, Corollary 4, Corollary 5 and Corollary 6.

C.1 Notations

Before we start the proofs, we need some more formal definitions. If z is a real number, a vector or a matrix, we define its norm as $\mid z \mid = \sum_k \mid z_k \mid$. When Z is a random variable, we define its L^p-norm as $\mid\mid Z \mid\mid = \mathbb{E}[\mid Z \mid^p]^{1/p}$. When Y_n and Z_n are two sequences of random variables, we define the notation small tau as $Y_n = o_{\mathbb{P}}(Z_n)$, i.e. that $\frac{Y_n}{Z_n}\mathbf{1}_{\{Z_n \neq 0\}} \to^{\mathbb{P}} 0$, and the notation big tau $Y_n = O_{\mathbb{P}}(Z_n)$, i.e. that $\frac{Y_n}{Z_n}\mathbf{1}_{\{Z_n \neq 0\}}$ is stochastically bounded. Moreover, given a Borel space $(E, \mathbf{B}(E))$, $C_b(E, \mathbb{R})$ is defined as the set of continuous and bounded functions from the space E to $\mathbb{R}$. For a measure μ, let $\mathbb{L}^1(\mu)$ be the space of functions that are integrable with respect to μ. Finally, we define for any $i = 1, \dots, d$ the event times of the ith process as

$$(T_1^{(i)}, \dots, T_{N^{(i)}}^{(i)}).$$

Since N_t is a point process, its $\mathbf{F}$-intensity (4) can be re-expressed partly as the sum at jump times, i.e.

$$\lambda^{(i)}(t, \theta^*) = n\nu^{*,(i)} + \sum_{j=1}^{d} \sum_{k \in \mathbb{N}_* \text{ s.t. } T_k^{(j)} < t} nh^{(i,j)}(n(t - T_k^{(j)}), \theta_{ker}^{*,(i,j)}). \tag{C1}$$

C.2 Time transformation and some lemmas

Our proof strategy follows the general machinery of Clinet and Yoshida (2017), which consider large-T asymptotics. To rewrite our problem with in-fill asymptotics as a problem with large-T asymptotics, we consider a time transformation as in Clinet and Potiron (2018) and Kwan et al.

(2023). More specifically, we define the time-transformed filtration as

$$\overline{\mathbf{F}}_n = (\overline{\mathcal{F}}_{t,n})_{t\in[0,nT]},$$
$$\overline{\mathcal{F}}_{t,n} = \mathcal{F}_{\frac{t}{n},n}.$$

From now on, we implicitly assume that all the defined quantities are $\overline{\mathbf{F}}_n$-adapted. For any $i = 1, \ldots, d$ we define the ith process of the time-transformed point process as

$$\begin{aligned} \overline{N}_n^{(i)} : [0, nT] &\to \mathbb{N} \\ t &\mapsto \overline{N}_{t,n}^{(i)} = N_{\frac{t}{n},n}^{(i)}, \end{aligned} \tag{C2}$$

with corresponding jump times

$$(\overline{T}_{1,n}^{(i)}, \ldots, \overline{T}_{N_n^{(i)},n}^{(i)})$$

defined such that $\overline{T}_{k,n}^{(i)} = nT_{k,n}^{(i)}$ and the rescaled time-transformed stochastic $\overline{\mathbf{F}}_n$-intensity process as

$$\begin{aligned} \overline{\lambda}_n^{(i)} : [0, nT] \times \Theta &\to \mathbb{R}^+ \\ (t, \theta) &\mapsto \overline{\lambda}_n^{(i)}(t, \theta) = \frac{\lambda_n^{(i)}(\frac{t}{n}, \theta)}{n}. \end{aligned} \tag{C3}$$

In this first lemma, we rewrite the rescaled time-transformed stochastic $\overline{\mathbf{F}}_n$-intensity in terms of the time-transformed point process.

Lemma C1. *For any $(t, \theta) \in [0, nT] \times \Theta$ and any $i = 1, \ldots, d$ we have that*

$$\overline{\lambda}_n^{(i)}(t, \theta) = \nu^{(i)} + \sum_{j=1}^{d} \int_0^{t^-} h^{(i,j)}((t - s), \theta_{ker}^{(i,j)}) d\overline{N}_{s,n}^{(j)}. \tag{C4}$$

Proof. If we substitute Equation (4) into Definition (C3), we obtain

$$\overline{\lambda}_n^{(i)}(t, \theta) = \nu^{(i)} + \sum_{j=1}^{d} \int_0^{t^-} h^{(i,j)}((t - s), \theta_{ker}^{(i,j)}) dN_{\frac{s}{n},n}^{(j)}. \tag{C5}$$

Finally, we can conclude by substituting Definition (C2) into Equation (C5). □

The next lemma shows that $\overline{N}_n$ is a multidimensional Hawkes process with the same kernel.

Lemma C2. *We have that*

$$\overline{M}_{t,n} = \overline{N}_{t,n} - \int_0^t \overline{\lambda}_n(s, \theta^*)ds \tag{C6}$$

is a $\overline{\mathbf{F}}_n$-local martingale. In particular, this implies that $\overline{N}_n$ is a multidimensional Hawkes process with the same kernel and related $\overline{\mathbf{F}}_n$-intensity $\overline{\lambda}_n(., \theta^*)$.

Proof. By definition of a compensator, we have that

$$M_{t,n} = N_{t,n} - \int_0^t \lambda_n(s, \theta^*)ds \tag{C7}$$

is a $\mathbf{F}_n$-local martingale. First, we will show that Equation (C6) is a $\overline{\mathbf{F}}_n$-local martingale. In fact we have

$$\begin{aligned} \overline{M}_{t,n} &= \overline{N}_{t,n} - \int_0^t \overline{\lambda}_n(s, \theta^*)ds \\ &= N_{\frac{t}{n},n} - \int_0^t \frac{\lambda_n(\frac{s}{n}, \theta^*)}{n}ds \\ &= N_{\frac{t}{n},n} - \int_0^{\frac{t}{n}} \lambda_n(y, \theta^*)dy \\ &= M_{\frac{t}{n},n}, \end{aligned}$$

where we used Equation (C6) in the first equality, Equation (C2) and Equation (C3) in the second equality, integral change of variable in the third equality and Equation (C7) in the fourth equality. Now, as $M_{t,n}$ is a $\mathbf{F}_n$-local martingale, it is clear that the time-transformed local martingale $M_{\frac{t}{n},n}$ is a $\overline{\mathbf{F}}_n$-local martingale. Then, it means that $\overline{M}_{\frac{t}{n},n}$ is a $\overline{\mathbf{F}}_n$-local martingale, thus we have shown the lemma. Second, we can deduce that $\overline{N}_n$ is a multidimensional Hawkes process with a mixture of generalized gamma kernels and related $\overline{\mathbf{F}}_n$-intensity $\overline{\lambda}_n(., \theta^*)$ by Theorem 3.17 (p. 32) in Jacod and Shiryaev (2013). □

We also define the log likelihood process of the time-transformed process as

$$\bar{l}_{T,n}(\theta) = \sum_{i=1}^{d} \int_0^{Tn} \log(\overline{\lambda}_n^{(i)}(t, \theta))d\overline{N}_{t,n}^{(i)} - \sum_{i=1}^{d} \int_0^{Tn} \overline{\lambda}_n^{(i)}(t, \theta)dt, \tag{C8}$$

and $\widehat{\widehat{\theta}}_{T,n}$ any maximizer of it. The following lemma states that a.s. the MLE on $[0,T]$ of the standard point process is equal to the MLE on $[0,nT]$ of the time-transformed point process.

Lemma C3. *We have that a.s.*

$$\widehat{\widehat{\theta}}_{T,n} = \widehat{\theta}_{T,n}$$

Proof. By the time-transformed process form and Lemma C2, the result follows. □

C.3 Adaptation of some lemmas in the mixture of generalized gamma kernels case

The following lemma shows that the time-transformed $\overline{\mathbf{F}}_n$-intensity, together with its first three derivatives, are in L^p for any $p \in \mathbb{N}$ with $p \geq 2$. This corresponds to Condition **[A2] (i)** (p. 1804) in Clinet and Yoshida (2017). This extends Lemma A.5 (p. 1833) in Clinet and Yoshida (2017) which is restricted to the exponential kernel case to the mixture of generalized gamma kernels case.

Lemma C4. *We assume that Condition **[A]** holds. For any $i = 1, \dots, p$, the $\overline{\mathbf{F}}_n$-intensity process and their first derivatives satisfy for any $p \in \mathbb{N}$, $p \geq 2$,*

$$\sup_{n\in\mathbb{N}, t\in[0,nT]} \sum_{l=0}^{3} || \sup_{\theta\in\Theta} | \partial_\theta^l \overline{\lambda}_n^{(i)}(t,\theta) | ||_p < +\infty$$

Proof. Without loss of generality, we will show the statement only for integers of the form 2^p with $p \in \mathbb{N}^*$. By Lemma C1, we have that

$$\overline{\lambda}_n^{(i)}(t,\theta) = \nu^{(i)} + \sum_{j=1}^{d} \int_0^{t^-} h^{(i,j)}((t-s), \theta_{ker}^{(i,j)}) d\overline{N}_{s,n}^{(j)}.$$

Given that $\nu^{(i)}$ is bounded above uniformly in $\theta \in \Theta$ by Condition **[A] (vi)**, there is no loss of generality assuming that $\nu^{(i)} = 0$ in the rest of this proof. Thus, it remains to show for

$i, j = 1, \ldots, d$ and $l = 0, \ldots, 3$ that

$$\sup_{n\in\mathbb{N}, t\in[0,nT]} \mathbb{E}\Big[\Big| \int_0^t \sup_{\theta\in\Theta} \mid \partial_\theta^l h^{(i,j)}((t-s), \theta_{ker}^{(i,j)}) d\overline{N}_{s,n}^{(j)} \mid \Big|^{2^p}\Big] < +\infty.$$

Applying the triangular inequality, it is then sufficient to show that

$$\sup_{n\in\mathbb{N}, t\in[0,nT]} \mathbb{E}\Big[\Big| \int_0^t \sup_{\theta\in\Theta} \mid \partial_\theta^l h^{(i,j)}((t-s), \theta_{ker}^{(i,j)}) \mid d\overline{N}_{s,n}^{(j)} \Big|^{2^p}\Big] < +\infty.$$

Because the term inside the integral is positive, it is sufficient to show that uniformly in $n \in \mathbb{N}$ we have

$$\mathbb{E}\Big[\Big| \int_0^{nT} \sup_{\theta\in\Theta} \mid \partial_\theta^l h^{(i,j)}((nT-s), \theta_{ker}^{(i,j)}) \mid d\overline{N}_{s,n}^{(j)} \Big|^{2^p}\Big] < +\infty.$$

In view of Equation (5), this can be rewritten as

$$\mathbb{E}\Big[\Big| \int_0^{nT} \sup_{\theta\in\Theta} \Big| \partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s), \theta_{ker}^{(i,j)}) \exp(-((nT-s)/\beta_k^{(i,j)})^{p_k^{(i,j)}}) \Big| d\overline{N}_{s,n}^{(i)} \Big|^{2^p}\Big] < +\infty,$$

where $P_k^{(i,j)}(t, \theta_{ker}^{(i,j)})$ is defined as

$$P_k^{(i,j)}(t, \theta_{ker}^{(i,j)}) = \alpha_k^{(i,j)} \frac{p_k^{(i,j)} t^{(D_k^{(i,j)}-1)}}{(\beta_k^{(i,j)})^{D_k^{(i,j)}} \Gamma(D_k^{(i,j)}/p_k^{(i,j)})}. \tag{C9}$$

Since Θ is assumed to be bounded, there exists $\beta_+ \in \mathbb{R}_+^*$ such that we have uniformly $\beta_k^{(i,j)} \leq \beta_+$. Then, an application of that inequality along with Condition **[A](iii)** yields

$$\begin{aligned}
&\mathbb{E}\Big[\Big| \int_0^{nT} \sup_{\theta\in\Theta} \Big| \partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s), \theta_{ker}^{(i,j)}) \exp(-((nT-s)/\beta_k^{(i,j)})^{p_k^{(i,j)}}) \Big| d\overline{N}_{s,n}^{(i)} \Big|^{2^p}\Big] \\
&\qquad \leq \mathbb{E}\Big[\Big| \int_0^{nT} e^{-((nT-s)/\beta_+)^{p_-}} \sup_{\theta\in\Theta} \Big| \partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s), \theta_{ker}^{(i,j)}) \Big| d\overline{N}_{s,n}^{(i)} \Big|^{2^p}\Big].
\end{aligned}$$

In what follows, we will show by induction that

$$\begin{aligned}
&\mathbb{E}\Big[\Big| \int_0^{nT} e^{-((nT-s)/\beta_+)^{p_-}} \sup_{\theta\in\Theta} \Big| \partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s), \theta_{ker}^{(i,j)}) \Big| d\overline{M}_{t,n}^{(i)} \Big|^{2^p}\Big] \qquad \text{(C10)} \\
\leq\ & K_p \mathbb{E}\Big[\int_0^{Tn} e^{-p((nT-s)/\beta_+)^{p_-}} \sup_{\theta\in\Theta} \Big| \partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s), \theta_{ker}^{(i,j)}) \Big|^{2p} \overline{\lambda}_n^{(i)}(t, \theta^*) dt\Big] \\
&+ K_p \mathbb{E}\Big[\Big| \int_0^{Tn} e^{-((nT-s)/\beta_+)^{p_-}} \sup_{\theta\in\Theta} \Big| \partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s), \theta_{ker}^{(i,j)}) \Big|^2 dt \Big|^{2^{p-1}}\Big].
\end{aligned}$$

We define

$$f(t) = e^{-((nT-s)/\beta_+)^{p-}} \sup_{\theta\in\Theta} \Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s), \theta_{ker}^{(i,j)})\Big|.$$

We consider the case $p = 1$. We can calculate

$$\begin{aligned}\mathbb{E}\Big[\Big|\int_0^{Tn} f(t)d\overline{M}_{t,n}^{(i)}\Big|^2\Big] &= \mathbb{E}\Big[\int_0^{Tn} f(t)^2 d\langle \overline{M}_n^{(i)}, \overline{M}_n^{(i)}\rangle_t\Big] \\ &= \mathbb{E}\Big[\int_0^{Tn} f(t)^2 \overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big],\end{aligned}$$

where the first equality was obtained with Itô isometry for point process martingales and the second equality is due to Equation (C6). This implies that Inequality (C10) holds in the case $p = 1$. We investigate now the case $p \geq 2$. By the Burkholder-Davis-Gundy inequality (see, e.g., Equation (2.1.32) in Jacod and Protter (2011)), we obtain

$$\begin{aligned}\mathbb{E}\Big[\Big|\int_0^{Tn} f(t)d\overline{M}_{t,n}^{(i)}\Big|^{2^p}\Big] &\leq D_p\mathbb{E}\Big[\Big|\int_0^{Tn} f(t)^2 d\overline{N}_{t,n}^{(i)}\Big|^{2^{p-1}}\Big] \\ &\leq 2^{p-1}D_p\mathbb{E}\Big[\Big|\int_0^{Tn} f(t)^2 d\overline{M}_{t,n}^{(i)}\Big|^{2^{p-1}}\Big] \\ &\quad +2^{p-1}D_p\mathbb{E}\Big[\Big|\int_0^{Tn} f(t)^2 \overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big|^{2^{p-1}}\Big].\end{aligned}$$

Now, an induction argument yields that for some constant $Q_p > 0$:

$$\mathbb{E}\Big[\Big|\int_0^{Tn} f(t)d\overline{M}_{t,n}^{(i)}\Big|^{2^p}\Big] \leq Q_p \sum_{q=1}^{p} \mathbb{E}\Big[\Big|\int_0^{Tn} f(t)^{2q} \overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big|^{2^{p-q}}\Big].$$

If we can show that for any $q = 1, \ldots, p$ we have

$$\begin{aligned}\Big|\int_0^{Tn} f(t)^{2q}\overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big|^{2^{p-q}} &\leq \int_0^{Tn} f(t)^{2p}\overline{\lambda}_n^{(i)}(t,\theta^*)dt \\ &\quad + \Big|f(t)^2 dt\Big|^{2^{p-1}},\end{aligned} \tag{C11}$$

then Inequality (C10) is shown with $K_p = pQ_p$. We prove now that Inequality (C11) holds. We write

$$g_n^{(i)}(t) = \frac{f(t)}{\Big|\int_0^{Tn} f(t)^2\overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big|^{\frac{1}{2}}}. \tag{C12}$$

We then obtain that

$$\Big|\int_0^{Tn} g_n^{(i)}(t)^{2q}\overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big|^{2^{p-q}} = \Big|\int_0^{Tn} g_n^{(i)}(t)^{2q-2}\mu_n^{(i)}(dt)\Big|^{2^{p-q}} \le \Big|\int_0^{Tn} g_n^{(i)}(t)^{2p}\overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big|^{\frac{2^{p-1}-2^{-q}}{2^{p-1}-1}} \le \Big|\int_0^{Tn} g_n^{(i)}(t)^{2p}\overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big|^{\frac{2^{p-1}-2^{-q}}{2^{p-1}-1}} + 1, \tag{C13}$$

where the equality is due to the fact that $\mu_n^{(i)}(dt) := g_n^{(i)}(t)^{2q}\overline{\lambda}_n^{(i)}(t,\theta^*)dt$ is a probability measure on $[0,Tn]$ and the first inequality comes from Jensen's inequality. If we reexpress Inequality (C13) with Definition (C12), we can show Inequality (C10).

We obviously have that

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\int_0^{nT} e^{-((nT-s)/\beta_+)^{p-}} \sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big| d\overline{N}_{s,n}^{(i)}\Big|^{2^p}\Big] \\
&= \mathbb{E}\Big[\Big|\int_0^{nT} e^{-((nT-s)/\beta_+)^{p-}} \sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big| d\overline{M}_{s,n}^{(i)}\Big|^{2^p}\Big] \\
&+\Big(\mathbb{E}\Big[\Big|\int_0^{nT} e^{-((nT-s)/\beta_+)^{p-}} \sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big| d\overline{N}_{s,n}^{(i)}\Big|^{2^p}\Big] \\
&-\mathbb{E}\Big[\Big|\int_0^{nT} e^{-((nT-s)/\beta_+)^{p-}} \sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big| d\overline{M}_{s,n}^{(i)}\Big|^{2^p}\Big]\Big).
\end{aligned}$$

We use now Inequality (C10), and we obtain that

$$\begin{aligned}
&\mathbb{E}\Big[\Big|\int_0^{nT} e^{-((nT-s)/\beta_+)^{p-}} \sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big| d\overline{N}_{t,n}^{(i)}\Big|^{2^p}\Big] \\
\le\ & K_p\mathbb{E}\Big[\int_0^{Tn} e^{-p((nT-s)/\beta_+)^{p-}} \sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big|^{2p}\overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big] \\
&+K_p\mathbb{E}\Big[\Big|\int_0^{Tn} e^{-((nT-s)/\beta_+)^{p-}} \sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big|^2 dt\Big|^{2^{p-1}}\Big] \\
&+\Big(\mathbb{E}\Big[\Big|\int_0^{nT} e^{-((nT-s)/\beta_+)^{p-}} \sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big| d\overline{N}_{s,n}^{(i)}\Big|^{2^p}\Big] \\
&-\mathbb{E}\Big[\Big|\int_0^{nT} e^{-((nT-s)/\beta_+)^{p-}} \sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big| d\overline{M}_{s,n}^{(i)}\Big|^{2^p}\Big]\Big) \\
:=\ & I_n + II_n + III_n.
\end{aligned}$$

We now show that $I_n < \infty$ uniformly in $n \in \mathbb{N}$. We can calculate

$$\begin{aligned}\frac{I_n}{K_p} &= \mathbb{E}\Big[\int_0^{Tn} e^{-p((nT-s)/\beta_+)^{p_-}} \sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big|^{2p}\overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big] \\ &= \int_0^{Tn} e^{-p((nT-s)/\beta_+)^{p_-}} \sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big|^{2p}\mathbb{E}\Big[\overline{\lambda}_n^{(i)}(t,\theta^*)\Big]dt,\end{aligned}$$

where we use the definition of I_n in the first equality, and the second equality is due to Tonelli's theorem along with the fact that $\mathbb{E}[aX] = a\mathbb{E}[X]$ for any random variable X and any nonrandom $a \in \mathbb{R}$. It remains to prove that uniformly in $n \in \mathbb{N}$ we have

$$\int_0^{Tn} e^{-p((nT-s)/\beta_+)^{p_-}} \sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big|^{2p}\mathbb{E}\Big[\overline{\lambda}_n^{(i)}(t,\theta^*)\Big]dt < \infty.$$

By the kernel definition (see Equation (5) and Equation (C9)), since Θ is bounded itself and by Condition **[A](iii)-(iv)** we have that $P_k^{(i,j)}(t,.)$ is in $C^3(\overline{\Theta})$. We can deduce that uniformly in $n \in \mathbb{N}$ we have

$$\sup_{\theta\in\Theta}\Big|\partial_\theta^l \sum_{k=1}^{K^{(i,j)}} P_k^{(i,j)}((nT-s),\theta_{ker}^{(i,j)})\Big|^{2p} \leq C.$$

The proof of $I_n < \infty$ amounts to showing that uniformly in $n \in \mathbb{N}$ we have

$$\int_0^{Tn} e^{-p((nT-s)/\beta_+)^{p_-}} \mathbb{E}\Big[\overline{\lambda}_n^{(i)}(t,\theta^*)\Big]dt < \infty.$$

By Condition **[A] (v)** and Definition (C3), we obtain uniformly in $n \in \mathbb{N}$ and in $s \in [0, nT]$ that

$$\mathbb{E}\Big[\overline{\lambda}_n^{(i)}(t,\theta^*)\Big] \leq C.$$

We thus obtain that

$$I_n \leq C\int_0^{Tn} e^{-p((nT-s)/\beta_+)^{p_-}} dt.$$

By a change of variable in the integral, we can deduce that

$$\int_0^{nT} e^{-p((nT-s)/\beta_+)^{p_-}} dt = \int_0^{nT} e^{-p(u/\beta_+)^{p_-}} du.$$

The obtained term can be dominated uniformly in $n \in \mathbb{N}$ by

$$\begin{aligned} \int_0^{nT} e^{-p(u/\beta_+)^{p_-}} du &\leq \int_0^{\infty} e^{-p(u/\beta_+)^{p_-}} du \\ &= C_1. \end{aligned} \tag{C14}$$

We have thus proven that $I_n < \infty$ uniformly in $n \in \mathbb{N}$. The proof for II_n and III_n follows with the same arguments. □

In what follows, we provide the definition of ergodicity in our time-transformed framework. This extends Definition 3.1 in Clinet and Yoshida (2017) which does not consider any time transformation. See also Kwan (2023) for a similar time-transformed framework.

Definition C1. *(ergodicity) We assume that $(E, \mathbf{B}(E))$ is a Borel space, and $X_n : \Omega \times [0, nT] \to E$ a sequence of stochastic processes adapted to the time-transformed filtration. We say that X_n is ergodic if there exists a mapping $\pi : C_b(E, \mathbb{R}) \to \mathbb{R}$ such that for any $\psi \in C_b(E, \mathbb{R})$ we have*

$$\frac{1}{nT} \int_0^{nT} \psi(X_{s,n}) ds \to^{\mathbb{P}} \pi(\psi).$$

The following definition introduces the notion of mixing to our time-transformed framework. This extends the definition from Section 3.4 in Clinet and Yoshida (2017) which does not consider any time transformation. See also Kwan (2023) for a similar time-transformed framework.

Definition C2. *(mixing) We assume that $(E, \mathbf{B}(E))$ is a Borel space, and $X_n : \Omega \times [0, nT] \to E$ a sequence of stochastic processes adapted to the time-transformed filtration. We say that X_n is C-mixing, for some set of functions C from E to $\mathbb{R}$, if for any ϕ, $\psi \in C$, the following convergence holds*

$$\rho_u := \sup_{n \in \mathbb{N}^*} \sup_{s \in [0, nT-u]} | \operatorname{Cov}[\phi(X_{s,n}), \psi(X_{s+u,n})] | \to 0 \text{ as } u \to +\infty.$$

The following proposition states that $X_n^{(i)}$ is mixing in the sense of Definition C2, stable, and ergodic in the sense of Definition C1. This corresponds to Condition **[A3]** (p. 1805) and Condition **[M1]** (p. 1815) in Clinet and Yoshida (2017). This extends Lemma 3.16 (p. 1815) and Lemma A.6 (p. 1834) in Clinet and Yoshida (2017) which are restricted to the exponential kernel case to the mixture of generalized gamma kernels case. This also extends their general machinery by proving first that $X_n^{(i)}$ is mixing and stable, and then this implies its ergodicity. Finally, this extends Kwan (2023) who considers the non-exponential kernel case but can only show the ergodicity of $(\overline{\lambda}_n^{(i)}(.,\theta^*), \overline{\lambda}_n^{(i)}(.,\theta))$ but not the ergodicity of $X_n^{(i)}$. The stability is a direct consequence of Theorem 1 and Lemma 4 in Brémaud and Massoulié (1996), along with Condition **[A] (v)**.

Proposition C1. *We assume that Condition **[A] (ii) to (vi)** hold. For any $i = 1, \ldots, d$ and any $\theta \in \Theta$, $X_n^{(i)}$ is:*

(i) $C_b(E, \mathbb{R})$-mixing in the sense of Definition C2.

(ii) stable, i.e. there exists an $\mathbb{R}_+^*$-valued random variable $\overline{\lambda}_{lim}^{(i)}(\theta)$ such that

$$X_{nT,n}^{(i)} \to^{\mathcal{D}} (\overline{\lambda}_{lim}^{(i)}(\theta^*), \overline{\lambda}_{lim}^{(i)}(\theta), \partial_\theta \overline{\lambda}_{lim}^{(i)}(\theta)).$$

(iii) ergodic in the sense of Definition C1, i.e. there exists a mapping $\pi^{(i)} : C_b(E, \mathbb{R}) \times \Theta \to \mathbb{R}$ such that for any $(\psi, \theta) \in C_b(E, \mathbb{R}) \times \Theta$ we have $\frac{1}{nT} \int_0^{nT} \psi(X_{s,n}^{(i)}) ds \to^{\mathbb{P}} \pi^{(i)}(\psi, \theta)$, where $\pi^{(i)}(\psi, \theta) = \mathbb{E}[\psi(\overline{\lambda}_{lim}^{(i)}(\theta^*), \overline{\lambda}_{lim}^{(i)}(\theta), \partial_\theta \overline{\lambda}_{lim}^{(i)}(\theta))]$.

Proof. We first show (i). We first define the truncated version of $X_{s,n}^{(i)}$ at time $t \leq s$ as

$$\widetilde{X}_{t,s,n}^{(i)} := (\overline{\lambda}_n^{(i)}(s, \theta^*), \sum_{j=1}^d \int_t^{s^-} h^{(i,j)}(s-u, \theta) d\overline{N}_{u,n}^{(i)}, \sum_{j=1}^d \int_t^{s^-} \partial_\theta h^{(i,j)}(s-u, \theta) d\overline{N}_{u,n}^{(i)}).$$

By considering ϕ, $\psi \in C_b(E, \mathbb{R})$, we can reexpress $\rho_u^{(i)}$ as

$$\begin{aligned}
\rho_u^{(i)} &= \sup_{n \in \mathbb{N}^*} \sup_{s \in [0, nT-u]} | \operatorname{Cov}[\phi(X_{s,n}^{(i)}), \psi(X_{s+u,n}^{(i)})] | \\
&= \sup_{n \in \mathbb{N}^*} \sup_{s \in [0, nT-u]} | \operatorname{Cov}[\phi(X_{s,n}^{(i)}), \psi(X_{s+u,n}^{(i)}) - \psi(\widetilde{X}_{s+v,s+u,n}^{(i)}) + \psi(\widetilde{X}_{s+v,s+u,n}^{(i)})] |,
\end{aligned}$$

where we use Definition C2 in the first equality, and we have $v \leq s - u$ in the second equality. Using the triangular inequality, we can dominate $\rho_u^{(i)}$ as

$$\begin{aligned}
\rho_u^{(i)} &\leq \sup_{n \in \mathbb{N}^*} \sup_{s \in [0, nT-u]} \mid \operatorname{Cov}[\phi(X_{s,n}^{(i)}), \psi(X_{s+u,n}^{(i)}) - \psi(\widetilde{X}_{s+v,s+u,n}^{(i)})] \mid \\
&\quad + \sup_{n \in \mathbb{N}^*} \sup_{s \in [0, nT-u]} \mid \operatorname{Cov}[\phi(X_{s,n}^{(i)}), \psi(\widetilde{X}_{s+v,s+u,n}^{(i)})] \mid \\
&:= I_u + II_u.
\end{aligned}$$

Since Θ is assumed to be bounded, there exists $\beta_+ \in \mathbb{R}_+^*$ such that we have uniformly $\beta_k^{(i,j)} \leq \beta_+$. Then, an application of that inequality along with Condition **[A](ii) and (iii)** yields that the intensity process is decreasing exponentially. Thus, we can deduce by similar arguments from the proof of Lemma A.6 (pp. 1834-1836) in Clinet and Yoshida (2017) that $I_u \to 0$ and $II_u \to 0$. This in turn implies that

$$\rho_u^{(i)} \to 0 \text{ as } u \to +\infty.$$

The stability (ii) is a direct consequence of Theorem 1 and Lemma 4 in Brémaud and Massoulié (1996), along with Condition **[A] (v)**.

We now show the ergodicity (iii). For $\psi \in C_b(E, \mathbb{R})$ we define $V_n^{(i)}(\psi, \theta)$ as

$$V_n^{(i)}(\psi, \theta) = \frac{1}{nT} \int_0^{nT} \psi(X_{s,n}^{(i)}) ds. \tag{C15}$$

We consider $\pi^{(i)}(\psi, \theta) = \mathbb{E}[\psi(\overline{\lambda}_{lim}^{(i)}(\theta^*), \overline{\lambda}_{lim}^{(i)}(\theta), \partial_\theta \overline{\lambda}_{lim}^{(i)}(\theta))]$. Establishing ergodicity amounts to showing the convergence in probability, i.e. $V_n^{(i)}(\psi, \theta) \to^{\mathbb{P}} \pi^{(i)}(\psi, \theta)$. In what follows, we show a stronger statement, i.e. the L^2-convergence. We calculate

$$\begin{aligned}
\mathbb{E}[(V_n^{(i)}(\psi, \theta) - \pi^{(i)}(\psi, \theta))^2] &= \operatorname{Var}[V_n^{(i)}(\psi, \theta)] + (\mathbb{E}[V_n^{(i)}(\psi, \theta)] - \pi^{(i)}(\psi, \theta))^2 \\
&:= I_n + II_n
\end{aligned}$$

where the equality is due to the fact that for any random variable X and any nonrandom $a \in \mathbb{R}$

we have $\mathbb{E}[(X-a)^2] = \text{Var}[X] + (\mathbb{E}[X]-a)^2$. For the first term, we have that

$$\begin{aligned}
I_n &= \text{Var}[V_n^{(i)}(\psi,\theta)] \\
&= \text{Var}\left[\frac{1}{nT}\int_0^{nT} \psi(X_{s,n}^{(i)})ds\right] \\
&= \frac{1}{n^2T^2}\text{Var}\left[\int_0^{nT} \psi(X_{s,n}^{(i)})\right] \\
&= \frac{1}{n^2T^2}\text{Var}\left[\lim_{K\to\infty}\frac{nT}{K}\sum_{k=0}^{K-1}\psi(X_{knT/K,n}^{(i)})\right] \\
&= \frac{1}{n^2T^2}\lim_{K\to\infty}\text{Var}\left[\frac{nT}{K}\sum_{k=0}^{K-1}\psi(X_{knT/K,n}^{(i)})\right] \\
&= \frac{1}{n^2T^2}\lim_{K\to\infty}\frac{n^2T^2}{K^2}\text{Var}\left[\sum_{k=0}^{K-1}\psi(X_{knT/K,n}^{(i)})\right] \\
&= \frac{1}{n^2T^2}\lim_{K\to\infty}\frac{n^2T^2}{K^2}\sum_{k=0}^{K-1}\sum_{l=0}^{K-1}\text{Cov}\left[\psi(X_{knT/K,n}^{(i)}),\psi(X_{lnT/K,n}^{(i)})\right] \\
&= \frac{1}{n^2T^2}\int_0^{nT}\int_0^{nT}\text{Cov}\left[\psi(X_{s,n}^{(i)}),\psi(X_{u,n}^{(i)})\right]dsdu,
\end{aligned}$$

where the second equality is obtained via Definition (C15), the third equality and the sixth equality are due to the fact that for any nonrandom $a \in \mathbb{R}$ and any random variable X we have $\text{Var}[aX] = a^2\,\text{Var}[X]$, we used the approximation of Riemann sum in the fourth equality and eight equality, the fifth equality is an application of dominated convergence theorem, and the seventh equality corresponds to Bienayme's identity. By Definition C1 and results obtained in (i), we can bound I_n as

$$I_n \leq \frac{1}{n^2T^2}\int_0^{nT}\int_0^{nT}\rho_{|s-u|}^{(i)}dsdu.$$

Then, splitting the integral into two terms leads to

$$I_n \leq \frac{1}{n^2T^2}\int_0^{nT}\int_0^{nT}\rho_{|s-u|}^{(i)}\mathbf{1}_{\{|s-u|\leq\sqrt{nT}\}}dsdu + \frac{1}{n^2T^2}\int_0^{nT}\int_0^{nT}\rho_{|s-u|}^{(i)}\mathbf{1}_{\{|s-u|>\sqrt{nT}\}}dsdu$$

Since there exists $\rho^{(i)}_{\max} > 0$ such that for any $t \geq 0$ we have $\rho^{(i)}_t \leq \rho^{(i)}_{\max}$, we can deduce that

$$\begin{aligned}\frac{1}{n^2T^2}\int_0^{nT}\int_0^{nT}\rho^{(i)}_{|s-u|}\mathbf{1}_{\{|s-u|\leq\sqrt{nT}\}}dsdu &\leq \frac{\rho^{(i)}_{\max}}{n^2T^2}\int_0^{nT}\int_0^{nT}\mathbf{1}_{\{|s-u|\leq\sqrt{nT}\}}dsdu \\ &= O(\frac{1}{\sqrt{nT}}) \\ &\to 0.\end{aligned}$$

We also have that

$$\begin{aligned}\frac{1}{n^2T^2}\int_0^{nT}\int_0^{nT}\rho^{(i)}_{|s-u|}\mathbf{1}_{\{|s-u|>\sqrt{nT}\}}dsdu &\leq \frac{\sup_{y>\sqrt{nT}}\rho^{(i)}_y}{n^2T^2}\int_0^{nT}\int_0^{nT}\mathbf{1}_{\{|s-u|>\sqrt{nT}\}}dsdu \\ &\leq \sup_{y>\sqrt{nT}}\rho^{(i)}_y.\end{aligned}$$

Since $\rho_u \to^{u\to\infty} 0$ by (i), we also deduce that

$$\sup_{y>\sqrt{nT}}\rho^{(i)}_y \to 0.$$

This implies that

$$\frac{1}{n^2T^2}\int_0^{nT}\int_0^{nT}\rho^{(i)}_{|s-u|}\mathbf{1}_{\{|s-u|>\sqrt{nT}\}}dsdu \to 0,$$

and thus $I_n \to 0$. For the second term, we know by (ii) that $X^{(i)}_{s,n} \to^{\mathcal{D}} (\overline{\lambda}^{(i)}_{lim}(\theta^*), \overline{\lambda}^{(i)}_{lim}(\theta), \partial_\theta\overline{\lambda}^{(i)}_{lim}(\theta))$. In particular, convergence in distribution implies convergence in expectation of any bounded function, thus we obtain that $\mathbb{E}[\psi(X^{(i)}_{s,n})] \to \mathbb{E}[\psi(\overline{\lambda}^{(i)}_{lim}(\theta^*), \overline{\lambda}^{(i)}_{lim}(\theta), \partial_\theta\overline{\lambda}^{(i)}_{lim}(\theta))]$. This can be reexpressed as $\mathbb{E}[\psi(X^{(i)}_{s,n})] \to \pi^{(i)}(\psi,\theta)$. In particular, this implies that $(\mathbb{E}[\psi(X^{(i)}_{s,n})]-\pi^{(i)}(\psi,\theta))^2 \to 0$, i.e. $II_n \to 0$. □

For a measure μ, we denote by $\mathbb{L}^1(\mu)$ the space of functions that are integrable with respect to μ. Since the functions that we will be using in our proofs will not necessarily be bounded, we need to extend from $C_b(E,\mathbb{R})$ to $C_\uparrow(E,\mathbb{R})$ the space of functions in which the ergodicity holds. We also give a more explicit form to the mapping $\pi^{(i)}(\psi,\theta)$. The following proposition extends Proposition 3.8 (pp. 1806-1807) in Clinet and Yoshida (2017). The proof follows the arguments from the proof of Proposition 3.8 (pp. 1822-1824) in Clinet and Yoshida (2017).

Proposition C2. *We assume that Condition* ***[A] (ii) to (vi)*** *holds. Then, for any $\theta \in \Theta$ and for any $i = 1, \dots, d$, the following properties hold*

(i) The ergodicity, i.e. Proposition C1(iii), still holds for any $\psi \in C_{\uparrow}(E, \mathbb{R})$. In particular, the mapping $\pi^{(i)}_{\theta^*}(., \theta)$ can be extended to $C_{\uparrow}(E, \mathbb{R})$. Moreover, for any $\psi \in C_{\uparrow}(E, \mathbb{R})$ the convergence is uniform in θ.

(ii) There exists a probability measure $\Pi^{(i)}_{\theta}$ on $(E, \mathbf{B}(E))$ such that for any $\psi \in C_{\uparrow}(E, \mathbb{R})$, we have $\pi^{(i)}(\psi, \theta) = \int_E \psi(x) \Pi^{(i)}_{\theta}(dx)$. In particular, $C_{\uparrow}(E, \mathbb{R}) \subset \mathbb{L}^1(\Pi^{(i)}_{\theta})$.

Proof. We can use the arguments from the proof of Proposition 3.8 (pp. 1822-1824) in Clinet and Yoshida (2017). □

C.4 Proofs of consistency

We define

$$\overline{\mathbb{Y}}_n(\theta) = \frac{1}{nT}(\bar{l}_{T,n}(\theta) - \bar{l}_{T,n}(\theta^*)) \tag{C16}$$

and also the asymptotic rescaled of the time-transformed log likelihood as

$$\overline{\mathbb{Y}}(\theta) = \sum_{i=1}^{d} \int_E (\log(\frac{v}{u})u - (v-u))\Pi^{(i)}_{\theta}(du, dv, dw). \tag{C17}$$

In the following lemma, we will prove that $\overline{\mathbb{Y}}_n(\theta)$ goes to $\overline{\mathbb{Y}}(\theta)$ uniformly in $\theta \in \Theta$ and in probability. This extends Lemma 3.10 (p. 1807) in Clinet and Yoshida (2017) which is restricted to the large-T asymptotics to the in-fill asymptotics.

Lemma C5. *We assume that Condition* ***[A]*** *holds. We have that*

$$\sup_{\theta \in \Theta} \mid \overline{\mathbb{Y}}_n(\theta) - \overline{\mathbb{Y}}(\theta) \mid \to^{\mathbb{P}} 0.$$

Proof. We can rewrite $\overline{\mathbb{Y}}_n(\theta)$ as

$$\begin{aligned}
\overline{\mathbb{Y}}_n(\theta) &= \frac{1}{nT}(\bar{l}_{T,n}(\theta) - \bar{l}_{T,n}(\theta^*)) \\
&= \frac{1}{nT}\Big(\sum_{i=1}^{d}\int_0^{Tn} \log(\overline{\lambda}_n^{(i)}(t,\theta))d\overline{N}_{t,n}^{(i)} - \sum_{i=1}^{d}\int_0^{Tn} \overline{\lambda}_n^{(i)}(t,\theta)dt \\
&\quad - \sum_{i=1}^{d}\int_0^{Tn} \log(\overline{\lambda}_n^{(i)}(t,\theta^*))d\overline{N}_{t,n}^{(i)} + \sum_{i=1}^{d}\int_0^{Tn} \overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big) \\
&= \frac{1}{nT}\sum_{i=1}^{d}\int_0^{Tn} \log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta^*)}\Big)d\overline{N}_{t,n}^{(i)} - \frac{1}{nT}\sum_{i=1}^{d}\int_0^{Tn} (\overline{\lambda}_n^{(i)}(t,\theta) - \overline{\lambda}_n^{(i)}(t,\theta^*))dt \\
&= \frac{1}{nT}\sum_{i=1}^{d}\int_0^{Tn} \log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta^*)}\Big)d\overline{M}_{t,n}^{(i)} \\
&\quad - \frac{1}{nT}\sum_{i=1}^{d}\int_0^{Tn} \Big(\overline{\lambda}_n^{(i)}(t,\theta) - \overline{\lambda}_n^{(i)}(t,\theta^*) - \log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta^*)}\Big)\overline{\lambda}_n^{(i)}(t,\theta^*)\Big)dt \\
&:= \sum_{i=1}^{d} I_n^{(i)}(nT,\theta) + \sum_{i=1}^{d} II_n^{(i)}(nT,\theta),
\end{aligned}$$

where we use Equation (C16) in the first equality, Equation (C8) in the second equality, algebraic manipulation in the third equality, Equation (C6) and algebraic manipulation in the fourth equality.

We first show that the martingale term disappears uniformly asymptotically in probability, i.e. that

$$\sup_{\theta\in\Theta}\Big|\sum_{i=1}^{d} I_n^{(i)}(nT,\theta)\Big| \to^{\mathbb{P}} 0.$$

As an application of Lemma C4 along with Condition **[A]**, for any $i = 1, \ldots, d$ and any $\theta \in \Theta$ we can deduce that $I_n^{(i)}(t,\theta)$ is an L^p-integrable martingale for any $p \in \mathbb{N}$, with $p \geq 2$. By Condition **[A] (vi)**, we can apply Sobolev's inequality, and for some big enough $p \in \mathbb{N}$ we obtain

$$\mathbb{E}\Big[\Big|\sup_{\theta\in\Theta} I_n^{(i)}(nT,\theta)\Big|^p\Big] \leq C\Big(\int_\Theta d\theta\mathbb{E}\Big[\Big|I_n^{(i)}(nT,\theta)\Big|^p\Big] + \int_\Theta d\theta\mathbb{E}\Big[\Big|\partial\theta I_n^{(i)}(nT,\theta)\Big|^p\Big]\Big). \tag{C18}$$

The first term in the right hand-side of Equation (C18) can be bounded by

$$
\begin{aligned}
\int_{\Theta} d\theta \mathbb{E}\Big[\Big|I_n^{(i)}(nT,\theta)\Big|^p\Big] &\leq C \sup_{\theta\in\Theta} \mathbb{E}\Big[\Big|I_n^{(i)}(nT,\theta)\Big|^p\Big] \\
&= C \sup_{\theta\in\Theta} \mathbb{E}\Big[\Big|\frac{1}{nT}\int_0^{Tn} \log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta^*)}\Big) d\overline{M}_{t,n}^{(i)}\Big|^p\Big] \\
&= C \sup_{\theta\in\Theta} \mathbb{E}\Big[\frac{1}{(nT)^p}\Big|\int_0^{Tn} \log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta^*)}\Big) d\overline{M}_{t,n}^{(i)}\Big|^p\Big] \\
&\leq C \sup_{\theta\in\Theta} \mathbb{E}\Big[\frac{1}{(nT)^p}\Big|\int_0^{Tn} \log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta^*)}\Big)^2 \overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big|^{\frac{p}{2}}\Big] \\
&\leq C \sup_{\theta\in\Theta} \mathbb{E}\Big[\frac{1}{(nT)^{\frac{p}{2}-1}}\int_0^{Tn} \log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta^*)}\Big)^p \overline{\lambda}_n^{(i)}(t,\theta^*)^{\frac{p}{2}}dt\Big] \\
&\leq C \sup_{\theta\in\Theta} \mathbb{E}\Big[\frac{1}{(nT)^{\frac{p}{2}-1}}\int_0^{Tn} \log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\nu_-}\Big)^p \overline{\lambda}_n^{(i)}(t,\theta^*)^{\frac{p}{2}}dt\Big],
\end{aligned}
$$

where the first equality is obtained with $I_n(t,\theta)$ definition, the second inequality is a consequence to Burkholder-Davis-Gundy inequality, the third inequality comes from Jensen's inequality, and the fourth inequality is due to Condition **[A](i)-(ii)**. We can continue to bound the first term in the right hand-side of Equation (C18) by

$$
\begin{aligned}
\int_{\Theta} d\theta \mathbb{E}\Big[\Big|I_n^{(i)}(nT,\theta)\Big|^p\Big] &\leq \sup_{\theta\in\Theta} \mathbb{E}\Big[\frac{1}{(nT)^{\frac{p}{2}-1}}\int_0^{Tn} \log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\nu_-}\Big)^p \overline{\lambda}_n^{(i)}(t,\theta^*)^{\frac{p}{2}}dt\Big] \\
&= \frac{1}{(nT)^{\frac{p}{2}-1}} \sup_{\theta\in\Theta} \mathbb{E}\Big[\int_0^{Tn} \log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\nu_-}\Big)^p \overline{\lambda}_n^{(i)}(t,\theta^*)^{\frac{p}{2}}dt\Big] \\
&= \frac{1}{(nT)^{\frac{p}{2}-1}} \sup_{\theta\in\Theta} \int_0^{Tn} \mathbb{E}\Big[\log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\nu_-}\Big)^p \overline{\lambda}_n^{(i)}(t,\theta^*)^{\frac{p}{2}}\Big]dt \\
&= \frac{1}{(nT)^{\frac{p}{2}-2}} \sup_{\theta\in\Theta} \sup_{n\in\mathbb{N},t\in[0,nT]} \mathbb{E}\Big[\log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\nu_-}\Big)^p \overline{\lambda}_n^{(i)}(t,\theta^*)^{\frac{p}{2}}\Big] \\
&\leq \frac{1}{(nT)^{\frac{p}{2}-2}} \sup_{\theta\in\Theta} \sup_{n\in\mathbb{N},t\in[0,nT]} \sqrt{\mathbb{E}\Big[\log\Big(\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\nu_-}\Big)^{2p}\Big]}\sqrt{\mathbb{E}\Big[\overline{\lambda}_n^{(i)}(t,\theta^*)^p\Big]} \\
&\leq \frac{1}{(nT)^{\frac{p}{2}-2}} \sup_{\theta\in\Theta} \sup_{n\in\mathbb{N},t\in[0,nT]} \sqrt{\mathbb{E}\Big[\Big(1+\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\nu_-}\Big)^{2p}\Big]}\sqrt{\mathbb{E}\Big[\overline{\lambda}_n^{(i)}(t,\theta^*)^p\Big]} \\
&\leq \frac{1}{(nT)^{\frac{p}{2}-2}} \sup_{n\in\mathbb{N},t\in[0,nT]} \sqrt{\mathbb{E}\Big[\sup_{\theta\in\Theta}\Big(1+\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\nu_-}\Big)^{2p}\Big]}\sqrt{\mathbb{E}\Big[\sup_{\theta\in\Theta}\overline{\lambda}_n^{(i)}(t,\theta^*)^p\Big]},
\end{aligned}
$$

where the first equality corresponds to the fact that $\mathbb{E}[aX] = a\mathbb{E}[X]$ for any random variable X and any nonrandom $a \in \mathbb{R}$, the second equality is explained by Tonelli's theorem, the second inequality is a consequence of Cauchy-Schwarz inequality, we used the fact that $\sup \mathbb{E}[.] \leq \mathbb{E}[\sup .]$ in the fourth inequality. Using the arguments from the proof of Lemma C4 along with Condition **[A]**, we can show that

$$\frac{1}{(nT)^{\frac{p}{2}-2}} \sup_{n\in\mathbb{N}, t\in[0,nT]} \sqrt{\mathbb{E}\Big[\sup_{\theta\in\Theta}\Big(1+\frac{\overline{\lambda}_n^{(i)}(t,\theta)}{\nu_-}\Big)^{2p}\Big]}\sqrt{\mathbb{E}\Big[\sup_{\theta\in\Theta}\overline{\lambda}_n^{(i)}(t,\theta^*)^p\Big]} \to 0,$$

which implies that for $i = 1, \ldots, d$ we have

$$\int_\Theta d\theta \mathbb{E}\Big[\Big|I_n^{(i)}(nT,\theta)\Big|^p\Big] \to 0.$$

With the same arguments, we can also show that for $i = 1, \ldots, d$ we have

$$\int_\Theta d\theta \mathbb{E}\Big[\Big|\partial\theta I_n^{(i)}(nT,\theta)\Big|^p\Big] \to 0.$$

Thus, we can deduce by Equation (C18) for $i = 1, \ldots, d$ that

$$\mathbb{E}\Big[\Big|\sup_{\theta\in\Theta} I_n^{(i)}(nT,\theta)\Big|^p\Big] \to 0. \tag{C19}$$

We can deduce that

$$\begin{aligned}
\mathbb{E}\Big[\Big|\sup_{\theta\in\Theta}\Big|\sum_{i=1}^d I_n^{(i)}(nT,\theta)\Big|^p\Big] &\leq C\mathbb{E}\Big[\sup_{\theta\in\Theta}\sum_{i=1}^d\Big|I_n^{(i)}(nT,\theta)\Big|^p\Big] \\
&\leq C\mathbb{E}\Big[\sum_{i=1}^d\sup_{\theta\in\Theta}\Big|I_n^{(i)}(nT,\theta)\Big|^p\Big] \\
&= C\sum_{i=1}^d\mathbb{E}\Big[\sup_{\theta\in\Theta}\Big|I_n^{(i)}(nT,\theta)\Big|^p\Big] \\
&= C\sum_{i=1}^d\mathbb{E}\Big[\Big|\sup_{\theta\in\Theta}I_n^{(i)}(nT,\theta)\Big|^p\Big] + o_{\mathbb{P}}(1) \\
&\to 0,
\end{aligned}$$

where the first inequality is a consequence of the fact that $|\sum_{i=1}^d a_i|^p \leq C\sum_{i=1}^d |a_i|^p$, the second inequality follows as $\sup\sum \leq \sum\sup$, the first equality corresponds to the fact that

$\mathbb{E}[aX] = a\mathbb{E}[X]$ for any random variable X and any nonrandom $a \in \mathbb{R}$, the second equality comes from Equation (C6) and the martingaleness of $\overline{M}_{t,n}$, and the convergence is due to Equation (C19). To prove that $|\sum_{i=1}^{d} II_n^{(i)}(nT,\theta) - \overline{\mathbb{Y}}(\theta)| \to^{\mathbb{P}} 0$, we can use Proposition C2 along with Condition **[A]**. □

We provide now the following lemma, which is the classical nondegeneracy condition on $\overline{\mathbb{Y}}$. This corresponds to Condition **[A](iv)** (p. 1807) in Clinet and Yoshida (2017). This complements Lemma A.7 (p. 1836) in Clinet and Yoshida (2017) which is restricted to the large-T asymptotics.

Lemma C6. *We assume that Condition **[A] (ii) to (vii)** hold. For any $\theta \in \overline{\Theta} - \{\theta^*\}$, we have that $\overline{\mathbb{Y}}(\theta) \neq 0$.*

Proof. We assume that $\theta \in \overline{\Theta}$ and that $\overline{\mathbb{Y}}(\theta) = 0$. In view of Equation (C17), we can deduce that

$$0 = \sum_{i=1}^{d} \int_E (\log(\frac{v}{u})u - (v-u))\Pi_\theta^{(i)}(du, dv, dw).$$

By Proposition C1 along with Condition **[A] (ii) to (vi)**, this equation can be reexpressed as

$$0 = \sum_{i=1}^{d} \mathbb{E}\Big[(\log(\frac{\overline{\lambda}_{lim}^{(i)}(\theta)}{\overline{\lambda}_{lim}^{(i)}(\theta^*)})\overline{\lambda}_{lim}^{(i)}(\theta^*) - (\overline{\lambda}_{lim}^{(i)}(\theta) - \overline{\lambda}_{lim}^{(i)}(\theta^*)))\Big].$$

For any $i = 1, \dots, d$ we also have that

$$0 \geq (\log(\frac{\overline{\lambda}_{lim}^{(i)}(\theta)}{\overline{\lambda}_{lim}^{(i)}(\theta^*)})\overline{\lambda}_{lim}^{(i)}(\theta^*) - (\overline{\lambda}_{lim}^{(i)}(\theta) - \overline{\lambda}_{lim}^{(i)}(\theta^*))).$$

This yields that for any $i = 1, \dots, d$ a.s.

$$0 = (\log(\frac{\overline{\lambda}_{lim}^{(i)}(\theta)}{\overline{\lambda}_{lim}^{(i)}(\theta^*)})\overline{\lambda}_{lim}^{(i)}(\theta^*) - (\overline{\lambda}_{lim}^{(i)}(\theta) - \overline{\lambda}_{lim}^{(i)}(\theta^*))).$$

We can then deduce that for any $i = 1, \dots, d$ a.s.

$$\overline{\lambda}_{lim}^{(i)}(\theta^*) = \overline{\lambda}_{lim}^{(i)}(\theta).$$

By Condition **[A] (ii) and (vii)**, we can get that $\theta^* = \theta$. □

We provide the proof of consistency in what follows. This extends Theorem 3.9 (p. 1807) in Clinet and Yoshida (2017) which is restricted to the large-T asymptotics to the in-fill asymptotics.

Proof of Equation (17) in Theorem 1. We have by Lemma C3 a.s. $\widehat{\widehat{\theta}}_{T,n} = \widehat{\theta}_{T,n}$. Since the consistency is a convergence in probability, we can replace $\widehat{\theta}_{T,n}$ by $\widehat{\widehat{\theta}}_{T,n}$ in the rest of this proof. In view of the expression $\overline{\mathbb{Y}}(\theta)$, we can see that $\overline{\mathbb{Y}}(\theta) \leq 0$ for any $\theta \in \Theta$ and $\overline{\mathbb{Y}}(\theta^*) = 0$. As an application of Lemma C6 along with Condition **[A]**, we can deduce that θ^* is a global maximum of $\overline{\mathbb{Y}}$. By Lemma C5 along with Condition **[A]**, the consistency is directly proven. □

C.5 Proofs of the CLT

We start with the following lemma. This extends Lemma A.1 (p. 1824) in Clinet and Yoshida (2017) which is restricted to the large-T asymptotics to the in-fill asymptotics.

Lemma C7. *We assume that Condition **[A]** holds. For any $\theta \in \Theta$, we have that $\bar{l}_{T,n}(\theta)$ is a.s. finite and admits a derivative in θ that satisfies*

$$\partial_\theta \bar{l}_{T,n}(\theta) = \sum_{i=1}^{d} \int_0^{Tn} \frac{\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta)} d\overline{N}_{t,n}^{(i)} - \sum_{i=1}^{d} \int_0^{Tn} \partial_\theta \overline{\lambda}_n^{(i)}(t,\theta) dt.$$

Moreover, we have that $\bar{l}_{T,n}(\theta)$ is twice differentiable and that its Hessian matrix satisfies

$$\begin{aligned} \partial_\theta^2 \bar{l}_{T,n}(\theta) &= \sum_{i=1}^{d} \int_0^{Tn} \partial_\theta \frac{\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta)} d\overline{M}_{t,n}^{(i)} - \sum_{i=1}^{d} \int_0^{Tn} (\partial_\theta \overline{\lambda}_n^{(i)})^{\otimes 2}(t,\theta)) \overline{\lambda}_n^{(i)}(t,\theta)^{-2} \overline{\lambda}_n^{(i)}(t,\theta^*) dt \\ &\quad + \sum_{i=1}^{d} \int_0^{Tn} (\partial_\theta^2 \overline{\lambda}_n^{(i)})(t,\theta)) \overline{\lambda}_n^{(i)}(t,\theta)^{-1} (\overline{\lambda}_n^{(i)}(t,\theta) - \overline{\lambda}_n^{(i)}(t,\theta^*)) dt. \end{aligned}$$

Proof. From Equation (C8), we define $\bar{l}_{T,n}(\theta) := \sum_{i=1}^{d} \bar{l}_{T,n}^{(i),I}(\theta) - \bar{l}_{T,n}^{(i),II}(\theta)$. First, we show that for any $\theta \in \Theta$ and any $i = 1, \ldots, d$ we have that $\bar{l}_{T,n}^{(i),I}(\theta) - \bar{l}_{T,n}^{(i),II}(\theta)$ is a.s. finite and admits a derivative in θ satisfying

$$\partial_\theta \int_0^{nT} \log(\overline{\lambda}_n^{(i)}(t,\theta)) d\overline{N}_{t,n}^{(i)} = \int_0^{nT} \partial_\theta \log(\overline{\lambda}_n^{(i)}(t,\theta)) d\overline{N}_{t,n}^{(i)}, \tag{C20}$$

$$\partial_\theta \int_0^{nT} \overline{\lambda}_n^{(i)}(t,\theta) dt = \int_0^{Tn} \partial_\theta \overline{\lambda}_n^{(i)}(t,\theta) dt. \tag{C21}$$

By Equation (C1), $\bar{l}_{T,n}^{(i),I}(\theta)$ can be reexpressed as

$$\bar{l}_{T,n}^{(i),I}(\theta) = \sum_{k \in \mathbb{N}_* \text{ s.t. } \overline{T}_k^{(i)} < nT} \log(\overline{\lambda}_n^{(i)}(\overline{T}_k^{(i)}, \theta)). \tag{C22}$$

As $\bar{l}_{T,n}^{(i),I}(\theta)$ is a finite sum, and since $\overline{\lambda}_n^{(i)}(\overline{T}_k^{(i)}, \theta) > 0$ by Condition **[A] (i)-(ii)** , it is a.s. finite. In addition, since log and $\overline{\lambda}_n^{(i)}(\overline{T}_k^{(i)}, .)$ admit derivative in $\theta \in \Theta$ by Lemma C4 along with Condition **[A]**, then $\bar{l}_{T,n}^{(i),I}$ also admits a derivative by the chain rule. As the sum is finite and by linearity of the derivative operator, we deduce that

$$\partial_\theta \sum_{k \in \mathbb{N}_* \text{ s.t. } \overline{T}_k^{(i)} < nT} \log(\overline{\lambda}_n^{(i)}(\overline{T}_k^{(i)}, \theta)) = \sum_{k \in \mathbb{N}_* \text{ s.t. } \overline{T}_k^{(i)} < nT} \partial_\theta \log(\overline{\lambda}_n^{(i)}(\overline{T}_k^{(i)}, \theta)).$$

By Equation (C22), this equality can be reexpressed as

$$\partial_\theta \int_0^{nT} \log(\overline{\lambda}_n^{(i)}(t,\theta)) d\overline{N}_{t,n}^{(i)} = \int_0^{nT} \partial_\theta \log(\overline{\lambda}_n^{(i)}(t,\theta)) d\overline{N}_{t,n}^{(i)}.$$

The term $\bar{l}_{T,n}^{(i),II}(\theta)$ will be a.s. finite if we can show that its L^1-norm is finite. We have that its L^1-norm can be bounded as

$$\begin{aligned}
\mathbb{E}\Big[\Big|\bar{l}_{T,n}^{(i),II}(\theta)\Big|\Big] &= \mathbb{E}\Big[\Big|\int_0^{nT} \log(\overline{\lambda}_n^{(i)}(t,\theta)) dt\Big|\Big] \\
&\leq \mathbb{E}\Big[\int_0^{nT} \Big|\log(\overline{\lambda}_n^{(i)}(t,\theta))\Big| dt\Big] \\
&\leq \mathbb{E}\Big[\int_0^{nT} |\overline{\lambda}_n^{(i)}(t,\theta)|^{-1} dt\Big] + \mathbb{E}\Big[\int_0^{nT} \Big|\overline{\lambda}_n^{(i)}(t,\theta)) - 1\Big| dt\Big] \\
&\leq \frac{nT}{\nu_-} + \mathbb{E}\Big[\int_0^{nT} \Big|\overline{\lambda}_n^{(i)}(t,\theta)) - 1\Big| dt\Big] \\
&\leq \frac{nT}{\nu_-} + C,
\end{aligned}$$

where we use the definition of $\bar{l}_{T,n}^{(i),II}(\theta)$ in the equality, the triangular inequality in the first inequality, the fact that $|\log(z)| \leq z^{-1} + |z-1|$ for any $z \in \mathbb{R}_*^+$ together with the linearity of the expectation operator in the second inequality, the third inequality is explained by Condition **[A] (i)-(ii)** and the fourth inequality by Lemma C4 along with Condition **[A]**. We have thus shown that the L^1-norm of $\bar{l}_{T,n}^{(i),II}(\theta)$ is finite, so that $\bar{l}_{T,n}^{(i),II}(\theta)$ is a.s. finite. By extending the arguments, we can prove that

$$\begin{aligned}\int_0^{Tn} |\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta)| dt &\leq \int_0^{Tn} |\sup_{\theta\in\Theta} \partial_\theta \overline{\lambda}_n^{(i)}(t,\theta)| dt \\ &\leq C.\end{aligned}$$

Now, an application of the dominated convergence theorem yields Equation (C21). We can prove the case $\partial_\theta^2 \bar{l}_{T,n}(\theta)$ with the same arguments.

□

We note that by Equation (C6), $\partial_\theta \bar{l}_{T,n}(\theta^*)$ can be reexpressed as

$$\partial_\theta \bar{l}_{T,n}(\theta^*) = \sum_{i=1}^{d} \int_0^{Tn} \frac{\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta^*)}{\overline{\lambda}_n^{(i)}(t,\theta^*)} d\overline{M}_{t,n}^{(i)}. \tag{C23}$$

We provide the proof of the CLT in what follows. This extends Theorem 3.11 (p. 1809) in Clinet and Yoshida (2017) which is restricted to the large-T asymptotics to the in-fill asymptotics.

Proof of Equations (18), (19) and (20) in Theorem 1. First, we have a.s. $\widehat{\widehat{\theta}}_{T,n} = \widehat{\theta}_{T,n}$ by Lemma C3 and thus we can replace $\widehat{\theta}_{T,n}$ by $\widehat{\widehat{\theta}}_{T,n}$ in the rest of this proof since it will not affect convergence in distribution. We obtain by a Taylor expansion that

$$\partial_\theta \bar{l}_{T,n}(\widehat{\widehat{\theta}}_{T,n}) = \partial_\theta \bar{l}_{T,n}(\theta^*) + \partial_\theta^2 \bar{l}_{T,n}(\zeta_n)(\widehat{\widehat{\theta}}_{T,n} - \theta^*),$$

where ζ_n is between $\widehat{\widehat{\theta}}_{T,n}$ and θ^*. Since $\widehat{\widehat{\theta}}_{T,n}$ is defined as the maximizer of $\bar{l}_{T,n}(.)$, we deduce that $\partial_\theta \bar{l}_{T,n}(\widehat{\widehat{\theta}}_{T,n}) = 0$. This yields that

$$0 = \partial_\theta \bar{l}_{T,n}(\theta^*) + \partial_\theta^2 \bar{l}_{T,n}(\zeta_n)(\widehat{\widehat{\theta}}_{T,n} - \theta^*).$$

If we multiply by $\frac{-\Gamma^{-1}}{\sqrt{nT}}$, we obtain that

$$0 = \frac{-\Gamma^{-1}}{\sqrt{nT}} \partial_\theta \bar{l}_{T,n}(\theta^*) + \frac{-\Gamma^{-1}}{\sqrt{nT}} \partial_\theta^2 \bar{l}_{T,n}(\zeta_n)(\widehat{\overline{\theta}}_{T,n} - \theta^*).$$

This equation can be reexpressed as

$$0 = \frac{-\Gamma^{-1}}{\sqrt{nT}} \partial_\theta \bar{l}_{T,n}(\theta^*) + \frac{-\Gamma^{-1}}{nT} \partial_\theta^2 \bar{l}_{T,n}(\zeta_n) \sqrt{nT}(\widehat{\overline{\theta}}_{T,n} - \theta^*).$$

To prove Equation (18), it remains to show that

$$\frac{-\Gamma^{-1}}{\sqrt{nT}} \partial_\theta \bar{l}_{T,n}(\theta^*) \quad \to^{\mathcal{D}} \quad \Gamma^{-1/2}\xi., \tag{C24}$$

$$\frac{-\Gamma^{-1}}{nT} \partial_\theta^2 \bar{l}_{T,n}(\zeta_n) \quad \to^{\mathbb{P}} \quad 1. \tag{C25}$$

Then, Equation (18) easily follows using Slutsky's theorem. We prove now Equation (C24). By Equation (C23), we have that

$$\frac{-\Gamma^{-1}}{\sqrt{nT}} \partial_\theta \bar{l}_{T,n}(\theta^*) = \frac{-\Gamma^{-1}}{\sqrt{nT}} \sum_{i=1}^{d} \int_0^{Tn} \frac{\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta^*)}{\overline{\lambda}_n^{(i)}(t,\theta^*)} d\overline{M}_{t,n}^{(i)}.$$

For $u \in [0,1]$, we define $S_{u,n}$ as

$$S_{u,n} = \frac{-\Gamma^{-1}}{\sqrt{nT}} \sum_{i=1}^{d} \int_0^{uTn} \frac{\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta^*)}{\overline{\lambda}_n^{(i)}(t,\theta^*)} d\overline{M}_{t,n}^{(i)}. \tag{C26}$$

We use Theorem $VIII.3.24$ in Jacod and Shiryaev (2013). We can calculate that

$$\begin{aligned} \langle S_n, S_n \rangle_u \quad &= \quad \frac{\Gamma^{-2}}{nT} \sum_{i=1}^{d} \int_0^{uTn} \frac{\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta^*)^2}{\overline{\lambda}_n^{(i)}(t,\theta^*)^2} dt. \\ &\to^{\mathbb{P}} \quad u\Gamma^{-1}. \end{aligned}$$

We prove now that Lindeberg's condition is satisfied. For any $a > 0$ we have

$$\begin{aligned} \mathbb{E}\Big[\sum_{s \le u} |\Delta S_{s,n}|^2 \mathbf{1}_{|\Delta S_{s,n}| > a}\Big] \quad &\le \quad \mathbb{E}\Big[\frac{1}{a} \sum_{s \le u} |\Delta S_{s,n}|^3\Big] \\ &= \quad \mathbb{E}\Big[\frac{1}{a} \sum_{s \le u} \Big| \frac{-\Gamma^{-1}}{\sqrt{nT}} \sum_{i=1}^{d} \int_0^{uTn} \frac{\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta^*)}{\overline{\lambda}_n^{(i)}(t,\theta^*)} d\overline{M}_{t,n}^{(i)} \Big|^3\Big] \\ &= \quad \mathbb{E}\Big[\frac{1}{a} \sum_{s \le u} \Big| \frac{-\Gamma^{-1}}{\sqrt{nT}} \sum_{i=1}^{d} \int_0^{uTn} \frac{\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta^*)}{\overline{\lambda}_n^{(i)}(t,\theta^*)} d\overline{N}_{t,n}^{(i)} \Big|^3\Big] \end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\Big[\frac{1}{a}\sum_{i=1}^{d}\int_0^{uTn}\Big|\frac{-\Gamma^{-1}}{\sqrt{nT}}\frac{\partial_\theta\overline{\lambda}_n^{(i)}(t,\theta^*)}{\overline{\lambda}_n^{(i)}(t,\theta^*)}\Big|^3 d\overline{N}_{t,n}^{(i)}\Big] \\
&= \mathbb{E}\Big[\frac{1}{a}\sum_{i=1}^{d}\int_0^{uTn}\Big|\frac{-\Gamma^{-1}}{\sqrt{nT}}\frac{\partial_\theta\overline{\lambda}_n^{(i)}(t,\theta^*)}{\overline{\lambda}_n^{(i)}(t,\theta^*)}\Big|^3 \overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big] \\
&\leq \mathbb{E}\Big[\frac{1}{a}\Big|\frac{-\Gamma^{-1}}{\sqrt{nT}}\Big|^3\sum_{i=1}^{d}\int_0^{uTn}\Big|\frac{\partial_\theta\overline{\lambda}_n^{(i)}(t,\theta^*)}{\overline{\lambda}_n^{(i)}(t,\theta^*)}\Big|^3 \overline{\lambda}_n^{(i)}(t,\theta^*)dt\Big] \\
&\leq \mathbb{E}\Big[\frac{1}{a}\Big|\frac{-\Gamma^{-1}}{\sqrt{nT}}\Big|^3\sum_{i=1}^{d}\int_0^{uTn}\sup_{\theta\in\Theta}\Big|\frac{\partial_\theta\overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta)}\Big|^3 \overline{\lambda}_n^{(i)}(t,\theta)dt\Big],
\end{aligned}
$$

where we used the fact that $\mathbf{1}_{|\Delta S_{s,n}|>a} \leq \frac{1}{a}|\Delta S_{s,n}|$ in the first inequality, the first equality is due to Definition (C26), the second equality is explained by the fact that the drift part does not jump, the third and fourth equality are a consequence of the form of $d\overline{N}_{t,n}^{(i)}$. We can continue to bound the Linderberg's term by

$$
\begin{aligned}
\mathbb{E}\Big[\sum_{s\leq u}|\Delta S_{s,n}|^2\mathbf{1}_{|\Delta S_{s,n}|>\epsilon}\Big] &\leq \mathbb{E}\Big[\frac{1}{a}\Big|\frac{-\Gamma^{-1}}{\sqrt{nT}}\Big|^3\sum_{i=1}^{d}\int_0^{uTn}\sup_{\theta\in\Theta}\Big|\frac{\partial_\theta\overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta)}\Big|^3 \overline{\lambda}_n^{(i)}(t,\theta)dt\Big] \\
&\leq \mathbb{E}\Big[\frac{1}{a}\Big|\frac{-\Gamma^{-1}}{\sqrt{nT}}\Big|^3\sum_{i=1}^{d}\int_0^{uTn}\sup_{\theta\in\Theta}\frac{|\partial_\theta\overline{\lambda}_n^{(i)}(t,\theta)|^3}{\overline{\lambda}_n^{(i)}(t,\theta)^2}dt\Big] \\
&\leq \mathbb{E}\Big[\frac{1}{a\nu_-^2}\Big|\frac{-\Gamma^{-1}}{\sqrt{nT}}\Big|^3\sum_{i=1}^{d}\int_0^{uTn}\sup_{\theta\in\Theta}|\partial_\theta\overline{\lambda}_n^{(i)}(t,\theta)|^3 dt\Big] \\
&= \frac{1}{a\nu_-^2}\Big|\frac{-\Gamma^{-1}}{\sqrt{nT}}\Big|^3\sum_{i=1}^{d}\int_0^{uTn}\mathbb{E}\Big[\sup_{\theta\in\Theta}|\partial_\theta\overline{\lambda}_n^{(i)}(t,\theta)|^3\Big]dt \\
&\leq \frac{CunT}{a\nu_-^2}\Big|\frac{-\Gamma^{-1}}{\sqrt{nT}}\Big|^3 \\
&\to 0,
\end{aligned}
$$

where the third inequality is due to Condition **[A] (i)-(ii)**, and Lemma C4 along with Condition **[A]** is used for the fourth inequality. We have thus shown that Lindeberg's condition holds, so that Equation (C24) is satisfied. We prove now Equation (C25), i.e. that $\frac{-\Gamma^{-1}}{nT}\partial_\theta^2\overline{l}_{T,n}(\zeta_n) \to^{\mathbb{P}} 1$. It is sufficient to prove that

$$
|\Gamma + (nT)^{-1}\partial_\theta^2\overline{l}_{T,n}(\zeta_n)| \to^{\mathbb{P}} 0.
$$

If we define V_n as a shrinking ball centered on θ^* it is sufficient to show that

$$\sup_{\theta\in V_n} |\Gamma + (nT)^{-1}\partial_\theta^2 \bar{l}_{T,n}(\theta)| \to^{\mathbb{P}} 0. \tag{C27}$$

We can reexpress Equation (C27) as the sum of a martingale term and a drift term. For the martingale term, we can notice that $\partial_\theta \frac{\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta)}$ and $\partial_\theta^2 \frac{\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta)}$ are dominated by polynoms in $\partial_\theta^k \overline{\lambda}_n^{(i)}(t,\theta)$ and $\frac{1}{\overline{\lambda}_n^{(i)}(t,\theta)}$ for $k = 0, 1, 2, 3$. By an application of Sobolev's inequality, Lemma C4 along with Condition **[A]**, we obtain for p big enough that

$$\mathbb{E}\Big|\sup_{\theta\in\Theta} \frac{1}{nT}\int_0^{nT} \partial_\theta \frac{\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta)} d\overline{M}_{t,n}^{(i)}\Big|^p = O((nT)^{-p/2}).$$

Given that we have L^p convergence implies convergence in probability, we can deduce that

$$\sum_{i=1}^d \sup_{\theta\in\Theta} \frac{1}{nT}\int_0^{nT} \partial_\theta \frac{\partial_\theta \overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta)} d\overline{M}_{t,n}^{(i)} \to^{\mathbb{P}} 0.$$

We have that

$$\begin{aligned}
\mathbb{E}\sup_{\theta\in V_n}\Big|\frac{1}{nT}\int_0^{nT} \frac{\partial_\theta^2 \overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta)}(\overline{\lambda}_n^{(i)}(t,\theta) \\
-\overline{\lambda}_n^{(i)}(t,\theta^*))dt\Big| &\leq \frac{1}{nT}\int_0^{nT} \mathbb{E}\sup_{\theta\in V_n}\Big|\frac{\partial_\theta^2 \overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta)}(\overline{\lambda}_n^{(i)}(t,\theta) - \\
&\quad \overline{\lambda}_n^{(i)}(t,\theta^*))\Big|dt \\
&\leq \frac{1}{nT\nu_-}\int_0^{nT} \mathbb{E}\sup_{\theta\in V_n}\Big|\partial_\theta^2 \overline{\lambda}_n^{(i)}(t,\theta)(\overline{\lambda}_n^{(i)}(t,\theta) \\
&\quad -\overline{\lambda}_n^{(i)}(t,\theta^*))\Big|dt \\
&\leq \frac{1}{nT\nu_-}\int_0^{nT} \sqrt{\mathbb{E}\sup_{\theta\in V_n}\Big|\partial_\theta^2 \overline{\lambda}_n^{(i)}(t,\theta)\Big|^2} \\
&\quad \times\sqrt{\mathbb{E}\sup_{\theta\in V_n}\Big|\overline{\lambda}_n^{(i)}(t,\theta) - \overline{\lambda}_n^{(i)}(t,\theta^*)\Big|^2} dt
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{1}{nT\nu_-}\int_0^{nT}\sqrt{\sup_{n\in\mathbb{N},t\in[0,nT]}\mathbb{E}\Big[\sup_{\theta\in\Theta}\Big|\partial_\theta^2\overline{\lambda}_n^{(i)}(t,\theta)\Big|^2\Big]} \\
&\quad\times\sqrt{\mathbb{E}\sup_{\theta\in V_n}\Big|\overline{\lambda}_n^{(i)}(t,\theta)-\overline{\lambda}_n^{(i)}(t,\theta^*)\Big|^2}dt \\
&\leq \frac{C}{nT\nu_-}\int_0^{nT} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (C28)\\
&\quad\sqrt{\mathbb{E}\sup_{\theta\in V_n}\Big|\overline{\lambda}_n^{(i)}(t,\theta)-\overline{\lambda}_n^{(i)}(t,\theta^*)\Big|^2}dt,
\end{aligned}$$

where we use the triangular inequality and linearity of expectation operator in the first inequality, the second inequality is due to Condition **[A] (i)-(ii)**, the third inequality corresponds to Cauchy-Schwarz inequality, and the fifth inequality comes from Lemma C4 along with Condition **[A]**. Now, we obtain by a Taylor expansion that

$$\overline{\lambda}_n^{(i)}(t,\theta)-\overline{\lambda}_n^{(i)}(t,\theta^*)=\partial_\theta\overline{\lambda}_n^{(i)}(t,\widetilde{\theta})(\theta-\theta^*),$$

where $\widetilde{\theta}$ is between θ and θ^*. Applying square operator on both sides of the equation, we obtain that

$$\Big|\overline{\lambda}_n^{(i)}(t,\theta)-\overline{\lambda}_n^{(i)}(t,\theta^*)\Big|^2=\Big|\partial_\theta\overline{\lambda}_n^{(i)}(t,\widetilde{\theta})\Big|^2\Big|\theta-\theta^*\Big|^2.$$

We can easily deduce that

$$\sup_{\theta\in V_n}\Big|\overline{\lambda}_n^{(i)}(t,\theta)-\overline{\lambda}_n^{(i)}(t,\theta^*)\Big|^2\leq\sup_{\theta\in\Theta}\Big|\partial_\theta\overline{\lambda}_n^{(i)}(t,\theta)\Big|^2\Big|\theta-\theta^*\Big|^2. \tag{C29}$$

We have that

$$\begin{aligned}
\mathbb{E}\sup_{\theta\in V_n}\Big|\frac{1}{nT}\int_0^{nT}\frac{\partial_\theta^2\overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta)}(\overline{\lambda}_n^{(i)}(t,\theta) \\
-\overline{\lambda}_n^{(i)}(t,\theta^*))dt\Big| &\leq \frac{C}{nT\nu_-}\int_0^{nT}\sqrt{\mathbb{E}\sup_{\theta\in V_n}\Big|\overline{\lambda}_n^{(i)}(t,\theta)-\overline{\lambda}_n^{(i)}(t,\theta^*)\Big|^2}dt \\
&\leq \frac{C}{nT\nu_-}\int_0^{nT}\sqrt{\mathbb{E}\sup_{\theta\in\Theta}\Big|\partial_\theta\overline{\lambda}_n^{(i)}(t,\theta)\Big|^2\Big|\theta-\theta^*\Big|^2}dt
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{C}{\nu_-}\Big|\theta-\theta^*\Big|\sqrt{\sup_{n\in\mathbb{N},t\in[0,nT]}\mathbb{E}\sup_{\theta\in\Theta}\Big|\partial_\theta\overline{\lambda}_n^{(i)}(t,\theta)\Big|^2} \\
&\leq \frac{CK}{\nu_-}\Big|\theta-\theta^*\Big| \\
&\to 0.
\end{aligned}
$$

where the first inequality comes from Equation (C28), the second inequality is due to Equation (C29), the fourth inequality is deduced by Lemma C4 along with Condition **[A]**, and the convergence is due to the fact that $\theta \in V_n$ with V_n shrinking to θ. Since L^1 convergence implies convergence in probability we obtain that

$$
\sup_{\theta\in V_n}\Big|\frac{1}{nT}\int_0^{nT}\frac{\partial_\theta^2\overline{\lambda}_n^{(i)}(t,\theta)}{\overline{\lambda}_n^{(i)}(t,\theta)}(\overline{\lambda}_n^{(i)}(t,\theta)-\overline{\lambda}_n^{(i)}(t,\theta^*))dt\Big| \to^{\mathbb{P}} 0.
$$

For the drift term, we define the process as

$$
U_n^{(i)}(\theta) = \frac{1}{nT}\int_0^{nT}(\partial_\theta)^{\otimes 2}\overline{\lambda}_n^{(i)}(t,\theta)\overline{\lambda}_n^{(i)}(t,\theta)^{-2}\overline{\lambda}_n^{(i)}(t,\theta^*)dt.
$$

When evaluated at θ^*, this process is equal to

$$
U_n^{(i)}(\theta^*) = \frac{1}{nT}\int_0^{nT}(\partial_\theta)^{\otimes 2}\overline{\lambda}_n^{(i)}(t,\theta^*)\overline{\lambda}_n^{(i)}(t,\theta^*)^{-1}dt.
$$

Using the arguments from the proof of the martingale case, we can show for any $i = 1,\dots,d$ that

$$
|U_n^{(i)}(\theta^*) - U_n^{(i)}(\theta)| \to^{\mathbb{P}} 0.
$$

Then the conclusion follows by writing Γ as the limit of $\sum_{i=1}^d U_n^{(i)}(\theta^*)$ and an application of Proposition C2 along with Condition **[A]**. Finally, the consistency of the asymptotic variance estimator, i.e. Equation (19), follows given its definition (16), the definition of the Fisher Hessian matrix in Equation (15), along with the consistency of $\widehat{\theta}_{T,n}$ (see Equation (17) in Theorem 1). The feasible CLT, i.e. Equation (20), is obtained via the standard CLT (see Equation (18)) together with Slutsky's theorem. □

C.6 Proofs of CLT for latency

We first give the proof of Proposition 2.

Proof of Proposition 2. By Equation (7) and Equation (3), we can reexpress $\widehat{L}_{T,n} - L$ as

$$\widehat{L}_{T,n} - L \quad = \quad F(\widehat{\theta}_{T,n,l}) - F(\theta_l^*). \tag{C30}$$

By Condition **[B]**, we obtain for any $i = 1, \ldots, d$ and any $j = 1, \ldots, d$ by componentwise Taylor expansion that

$$\begin{aligned} F^{(i,j)}(\widehat{\theta}_{T,n,l}) - F^{(i,j)}(\theta_l^*) = dF^{(i,j)}(\theta_l^*)(\widehat{\theta}_{T,n,l} - \theta_l^*) \\ +(\widehat{\theta}_{T,n,l} - \theta_l^*)^T d^2 F^{(i,j)}(\tilde{\theta})(\widehat{\theta}_{T,n,l} - \theta_l^*), \end{aligned} \tag{C31}$$

where $d^2F^{(i,j)}(\theta_l)$ corresponds to the $(m-d)\times(m-d)$-dimensional Hessian matrix of the (i,j)-index of F at point θ_l, and $\tilde{\theta}$ is between $\widehat{\theta}_{T,n,l}$ and θ_l^*. To show the consistency, i.e. Equation (24), we can calculate for any $i = 1, \ldots, d$ and any $j = 1, \ldots, d$ that

$$\begin{aligned} \widehat{L}_{T,n}^{(i,j)} - L^{(i,j)} &= F^{(i,j)}(\widehat{\theta}_{T,n,l}) - F^{(i,j)}(\theta_l^*) \\ &= dF^{(i,j)}(\theta_l^*)(\widehat{\theta}_{T,n,l} - \theta_l^*) + (\widehat{\theta}_{T,n,l} - \theta_l^*)^T d^2F^{(i,j)}(\tilde{\theta})(\widehat{\theta}_{T,n,l} - \theta_l^*) \\ &= O_{\mathbb{P}}(||\widehat{\theta}_{T,n,l} - \theta_l^*||) + O_{\mathbb{P}}(||\widehat{\theta}_{T,n,l} - \theta_l^*||^2) \\ &\to^{\mathbb{P}} 0, \end{aligned}$$

where we use Equation (C30) in the first equality, the second equality is due to Equation (C31), the third equality is a consequence to the fact that $\overline{\Theta}$ is a compact set and F is twice continuously differentiable by Condition **[B]** so that $dF^{(i,j)}(\theta_l^*)$ and $d^2F^{(i,j)}(\tilde{\theta})$ are bounded, and the convergence is obtained via the consistency of $\widehat{\theta}_{T,n,l}$ (see Equation (17) in Theorem 1 along with Condition **[A]**). To prove the CLT, i.e. Equation (25), we can calculate

$$\begin{aligned}
\sqrt{nT}(\widehat{L}_{T,n}^{(i,j)} - L^{(i,j)})_{i=1,\dots,d}^{j=1,\dots,d} &= \sqrt{nT}(F^{(i,j)}(\widehat{\theta}_{T,n,l}) - F^{(i,j)}(\theta_l^*))_{i=1,\dots,d}^{j=1,\dots,d} \\
&= \sqrt{nT}\big(dF^{(i,j)}(\theta_l^*)(\widehat{\theta}_{T,n,l}) - \theta_l^*)\big)_{i=1,\dots,d}^{j=1,\dots,d} \\
&\quad + \sqrt{nT}\big((\widehat{\theta}_{T,n,l} - \theta_l^*)^T d^2 F^{(i,j)}(\tilde{\theta})(\widehat{\theta}_{T,n,l} - \theta_l^*)\big)_{i=1,\dots,d}^{j=1,\dots,d} \\
&= \sqrt{nT}\big(dF^{(i,j)}(\theta_l^*)(\widehat{\theta}_{T,n,l}) - \theta_l^*)\big)_{i=1,\dots,d}^{j=1,\dots,d} \\
&\quad + \sqrt{nT} O_{\mathbb{P}}(||\widehat{\theta}_{T,n,l} - \theta_l^*||^2) \\
&= \sqrt{nT}\big(dF^{(i,j)}(\theta_l^*)(\widehat{\theta}_{T,n,l}) - \theta_l^*)\big)_{i=1,\dots,d}^{j=1,\dots,d} \\
&\quad + O_{\mathbb{P}}(||\widehat{\theta}_{T,n,l} - \theta_l^*||) \\
&\to^{\mathcal{D}} (dF^{(i,j)}(\theta_l^*)\Gamma_l^{-1/2}\xi_l)_{i=1,\dots,d}^{j=1,\dots,d}
\end{aligned}$$

where we use Equation (C30) in the first equality, the second equality is a consequence to Equation (C31), the third equality is due to the fact that $\overline{\Theta}$ is a compact set and F is twice continuously differentiable by Condition **[B]** so that $d^2F^{(i,j)}(\tilde{\theta})$ is bounded, the fourth equality is a consequence to the CLT of $\widehat{\theta}_{T,n,l}$ (see Equation (18) in Theorem 1 along with Condition **[A]**), and the convergence is obtained via the consistency and CLT of $\widehat{\theta}_{T,n,ker}$ (see Equation (17) and Equation (18) in Theorem 1 along with Condition **[A]**). Finally, we have that the limit is not null by Condition **[B]**). We show now Equation (26), i.e. we reexpress $\eta^{(i,j)}$ as

$$\begin{aligned}
\eta^{(i,j)} &= dF^{(i,j)}(\theta_l^*)\Gamma_l^{-1/2}\xi_l \\
&= dF^{(i,j)}(\theta_l^*)\Big(\sum_{q=1}^{l}\big(\Gamma_l^{-1/2}\big)^{(1,q)}\xi_l^{(q)}, \dots, \sum_{q=1}^{l}\big(\Gamma_l^{-1/2}\big)^{(m-d,q)}\xi_l^{(q)}\Big)^T \\
&= \sum_{r=1}^{l}\Big(dF^{(i,j,r)}(\theta_l^*)\sum_{q=1}^{l}\big(\Gamma_l^{-1/2}\big)^{(r,q)}\xi_l^{(q)}\Big) \\
&= \sum_{q=1}^{l}\Big(\sum_{r=1}^{l} dF^{(i,j,r)}(\theta_l^*)\big(\Gamma_l^{-1/2}\big)^{(r,q)}\Big)\xi_l^{(q)},
\end{aligned}$$

where the first equality is due to Equation (21), the second and third equalities are matrix calculation, the fourth equality is from algebraic manipulations by inverting sums. Equation

(27) can be deduced directly by using the fact that ξ_l follows an l-dimensional standard normal vector. The consistency of the covariance estimator, i.e. Equation (28), is due to the consistency of $\widehat{\theta}_{T,n,l}$ (see Equation (17) in Theorem 1 along with Condition **[A]**), the consistency of $\widehat{\Gamma}_{T,n,l}$ (see Equation (19) in Theorem 1), along with Condition **[B]**. The normalized feasible CLT, i.e. Equation (29) and Equation (30), is deduced via the standard CLT (see Equation (25)), the consistency of the covariance estimator (see Equation (28)), together with Slutsky's theorem. Finally, we can show Equation (31) since

$$
\begin{aligned}
\widehat{\operatorname{Cor}\left[\widetilde{\xi}^{(i,j)}, \widetilde{\xi}^{(k,u)}\right]} &= \frac{\widehat{\operatorname{Cov}[\eta^{(i,j)}, \eta^{(k,u)}]}}{\sqrt{\widehat{\operatorname{Var}\left[\eta^{(i,j)}\right]}\,\widehat{\operatorname{Var}\left[\eta^{(k,u)}\right]}}} \\
&= \frac{\sum_{q=1}^{l}\Big(\sum_{r=1}^{l} dF^{(i,j,r)}(\widehat{\theta}_l)\big(\widehat{\Gamma}_l^{-1/2}\big)^{(r,q)}\Big)\Big(\sum_{r=1}^{l} dF^{(k,u,r)}(\widehat{\theta}_l)\big(\widehat{\Gamma}_l^{-1/2}\big)^{(r,q)}\Big)}{\sqrt{\widehat{\operatorname{Var}\left[\eta^{(i,j)}\right]}\,\widehat{\operatorname{Var}\left[\eta^{(k,u)}\right]}}} \\
&\to \frac{\sum_{q=1}^{l}\Big(\sum_{r=1}^{l} dF^{(i,j,r)}(\theta_l^*)\big(\Gamma_l^{-1/2}\big)^{(r,q)}\Big)\Big(\sum_{r=1}^{l} dF^{(k,u,r)}(\theta_l^*)\big(\Gamma_{ker}^{-1/2}\big)^{(r,q)}\Big)}{\sqrt{\operatorname{Var}\left[\eta^{(i,j)}\right]\,\operatorname{Var}\left[\eta^{(k,u)}\right]}} \\
&= \frac{\operatorname{Cov}[\eta^{(i,j)}, \eta^{(k,u)}]}{\sqrt{\operatorname{Var}\left[\eta^{(i,j)}\right]\,\operatorname{Var}\left[\eta^{(k,u)}\right]}} \\
&= \operatorname{Cor}\left[\widetilde{\xi}^{(i,j)}, \widetilde{\xi}^{(k,u)}\right],
\end{aligned}
$$

where the first equality is due to Equation (23), the second equality comes from Equation (22), the convergence is due to the consistency of $\widehat{\theta}_{T,n,l}$ (see Equation (17) in Theorem 1 along with Condition **[A]**) together with the consistency of $\widehat{\Gamma}_{T,n,l}$ (see Equation (19) in Theorem 1 along with Condition **[A]**) in the numerator and to Equation (28) in the denominator, the third equality corresponds to Equation (27), and the fourth equality is obtained by Equation (30). □

We provide now the proof of Corollary 3.

Proof of Corollary 3. We obtain by the CLT of the latency estimator (see Equation (25) in Proposition 2 along with Condition **[A]** and Condition **[B]**) the asymptotic matrix $(\eta^{(i,j)})_{i=1,\dots,d}^{j=1,\dots,d}$, which can be reexpressed as in Equation (26). We can then deduce Equation (34) by Condition **[C]**. The consistency of the asymptotic covariance matrix inverse, i.e. Equation (35), follows directly from the consistency of the covariance estimator in Equation (28). The feasible normalized CLT, i.e. Equation (36), is deduced via the standard CLT (see Equation (34)), the consistency of the asymptotic covariance matrix inverse (see Equation (35)) together with Slutsky's theorem. □

C.7 Proofs of the tests related to latency

We finally give the proofs of the corollaries related to latency tests. We start with the proof of Corollary 4.

Proof of Corollary 4. The size of the first Wald test statistic $W(\widetilde{L})$, i.e. Equation (37), can be shown converging in distribution to a chi-squared distribution with one degree of freedom using its definition (see Equation (8)), the CLT of the latency matrix estimator and the consistency of the covariance estimator (see Equations (25) and (28) in Proposition 2 along with Condition **[A]** and Condition **[B]**), and the form of the chi-squared distribution with one degree of freedom. The power of the first Wald test statistic $W(\widetilde{L})$, i.e. Equation (38), goes to 1 as an application of the CLT of the latency matrix estimator and the consistency of the covariance estimator (see Equations (25) and (28) in Proposition 2 along with Condition **[A]** and Condition **[B]**) along with its definition (see Equation (8)). □

We provide the proof of Corollary 5 in what follows.

Proof of Corollary 5. Under the null hypothesis $H_0' : \{L^{(i,j)} = L^{(k,u)}\}$, we can calculate that

$$\begin{aligned}\sqrt{nT}(\widehat{L}_T^{(i,j)} - \widehat{L}_T^{(k,u)}) &= \sqrt{nT}(\widehat{L}_T^{(i,j)} - L^{(i,j)}) + \sqrt{nT}(L^{(k,u)} - \widehat{L}_T^{(k,u)}) + \sqrt{nT}(L^{(i,j)} - L^{(k,u)}) \\ &= \sqrt{nT}(\widehat{L}_T^{(i,j)} - L^{(i,j)}) + \sqrt{nT}(L^{(k,u)} - \widehat{L}_T^{(k,u)}) \\ &\to^{\mathcal{D}} \eta^{(i,j)} - \eta^{(k,u)},\end{aligned}$$

where the first equality corresponds to algebraic manipulation, the second equality is due to the fact that under H_0' we have that $L^{(i,j)} = L^{(k,l)}$, and the convergence comes from the CLT of the latency matrix estimator (see Equation (25) in Proposition 2 along with Condition **[A]** and Condition **[B]**). If we write

$$\widetilde{\eta} = \eta^{(i,j)} - \eta^{(k,u)}, \tag{C32}$$

we know that $\widetilde{\eta}$ is normally distributed since we assume that $(\eta^{(i,j)}, \eta^{(k,u)})$ is a two-dimensional random vector. We have that the mean of $\widetilde{\eta}$ is null by its definition in Equation (21). It remains to calculate its variance. We obtain that

$$\begin{aligned} \operatorname{Var}\left[\widetilde{\eta}\right] &= \operatorname{Var}\left[\eta^{(i,j)} - \eta^{(k,u)}\right] \\ &= \operatorname{Var}\left[\eta^{(i,j)}\right] + \operatorname{Var}\left[\eta^{(k,u)}\right] - 2\operatorname{Cov}\left[\eta^{(i,j)}, \eta^{(k,u)}\right] \end{aligned}$$

where the first equality comes from Equation (C32), and the second equality corresponds to a well-known variance-covariance property. By the consistency of the covariance estimator (see Equation (28) in Proposition 2 along with Condition **[A]** and Condition **[B]**), $\operatorname{Var}\left[\widetilde{\eta}\right]$ can be consistently estimated as

$$\widehat{\operatorname{Var}\left[\widetilde{\eta}\right]} = \widehat{\operatorname{Var}\left[\eta^{(i,j)}\right]} + \widehat{\operatorname{Var}\left[\eta^{(k,l)}\right]} - 2\widehat{\operatorname{Cov}\left[\eta^{(i,j)}, \eta^{(k,l)}\right]}.$$

As a consequence, we obtain that W' converges in distribution to a chi-square with one degree of freedom. Under the alternative hypothesis $H_1' : \{L^{(i,j)} \neq L^{(k,l)}\}$, we can calculate that

$$\begin{aligned} \sqrt{nT}(\widehat{L}_T^{(i,j)} - \widehat{L}_T^{(k,l)}) &= \sqrt{nT}(\widehat{L}_T^{(i,j)} - L^{(i,j)}) + \sqrt{nT}(L^{(k,l)} - \widehat{L}_T^{(k,l)}) + \sqrt{nT}(L^{(i,j)} - L^{(k,l)}) \\ &= O_{\mathbb{P}}(1) + \sqrt{nT}(L^{(i,j)} - L^{(k,l)}), \end{aligned}$$

where the first equality corresponds to algebraic manipulation, and the second equality comes from the CLT of the latency matrix estimator (see Equation (25) in Proposition 2 along with

Condition **[A]** and Condition **[B]**). Finally, we have by the consistency of the covariance estimator (see Equation (28) in Proposition 2 along with Condition **[A]** and Condition **[B]**) and by the fact that Ω is bounded that $\text{Var}\left[\widetilde{\eta}\right]$ is uniformly bounded. Thus, we can deduce that the second Wald statistic W' diverges under H_1'. □

Finally, we give the proof of Corollary 6.

Proof of Corollary 6. The size of the third Wald test statistic $\overline{W}(r)$, i.e. Equation (41), can be shown converging in distribution to a chi-squared distribution with q degrees of freedom using its definition (see Equation (10)), the CLT of the latency vector estimator and the consistency of the asymptotic covariance matrix inverse (see Equations (34) and (35) in Corollary 3 along with Condition **[A]**, Condition **[B]** and Condition **[C]**), and the form of the chi-squared distribution with q degrees of freedom. The power of the third Wald test statistic $\overline{W}(r)$, i.e. Equation (42), goes to 1 as an application of the CLT of the latency vector estimator and the consistency of the asymptotic covariance matrix inverse (see Equations (34) and (35) in Corollary 3 along with Condition **[A]**, Condition **[B]** and Condition **[C]**) along with its definition (see Equation (10)).

□

D Additional empirical results

The parameter estimates $\widehat{\alpha_T^{(i,j)}}$ are presented in Figure D1. Recall that α captures a size of jump associated with trading or quoting intensity. The size of jump for self-exciting effects in plots (a) and (c) is larger in periods of intensive trading - March 2020 when traders utilized the market shift due to the start of COVID-19 pandemic. Interestingly the cross-exciting effects from quotes to trades in the US, plots (f), and in Canada, plot (h), dropped during this period showing a change towards a trade driven market.

The parameter estimates $\widehat{\beta^{(i,j)}}$ are presented in Figure D2. At the beginning of COVID-19 pandemic in March 2020, decay captured by β changed from 1 to almost 2 reflecting acceleration in trading reactions for self-excitation parts in the US and Canada (plots (a) and (c)). Cross-exciting parameters $\widehat{\beta_T^{(4,3)}}$ and $\widehat{\beta_T^{(2,1)}}$ for quotes due to trades are stable during this period.

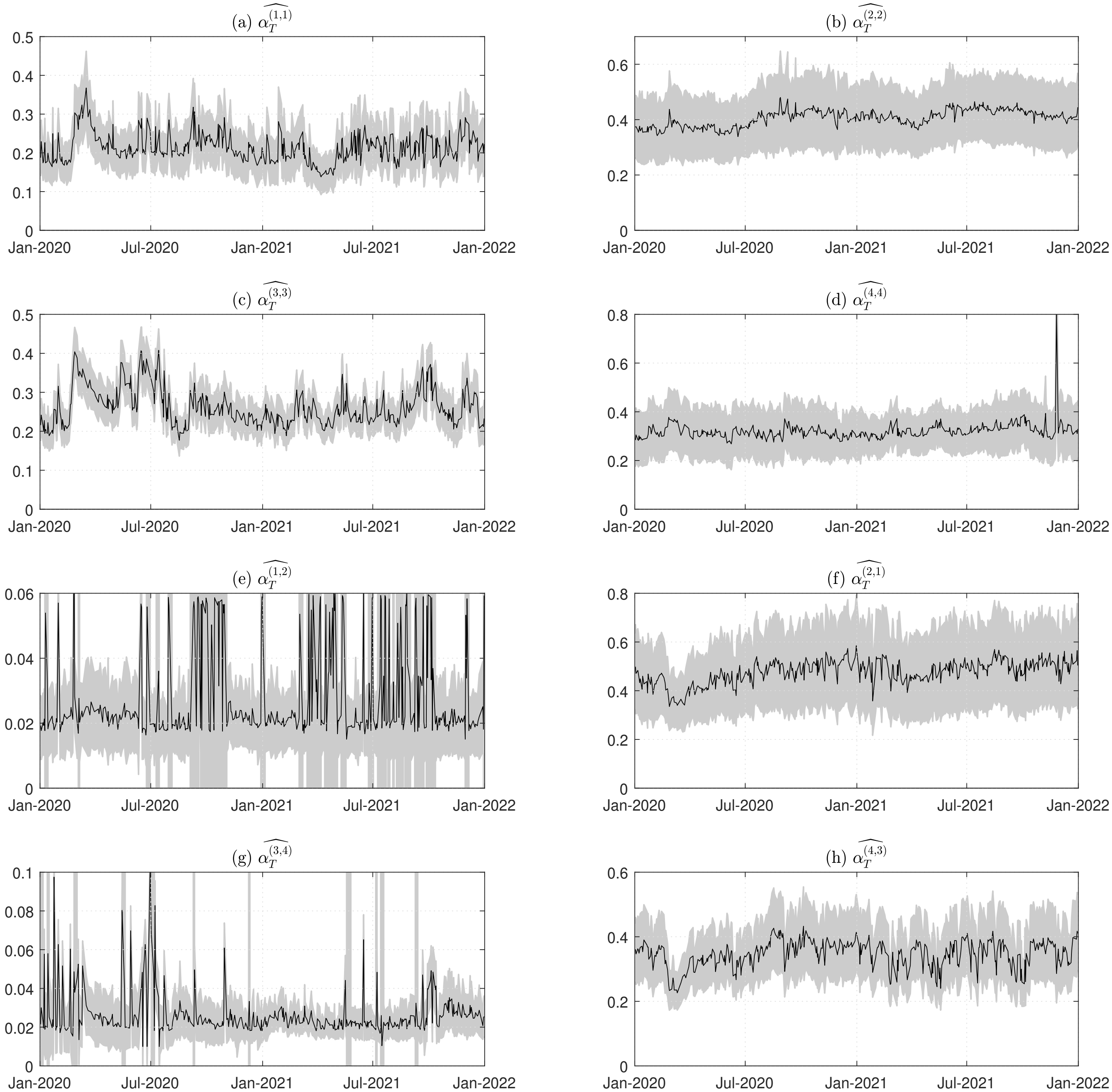

Figure D1: Parameter estimates $\widehat{\alpha_T^{(i,j)}}$ for each day for events in the NYSE and TSX. 90% confidence intervals are reported.

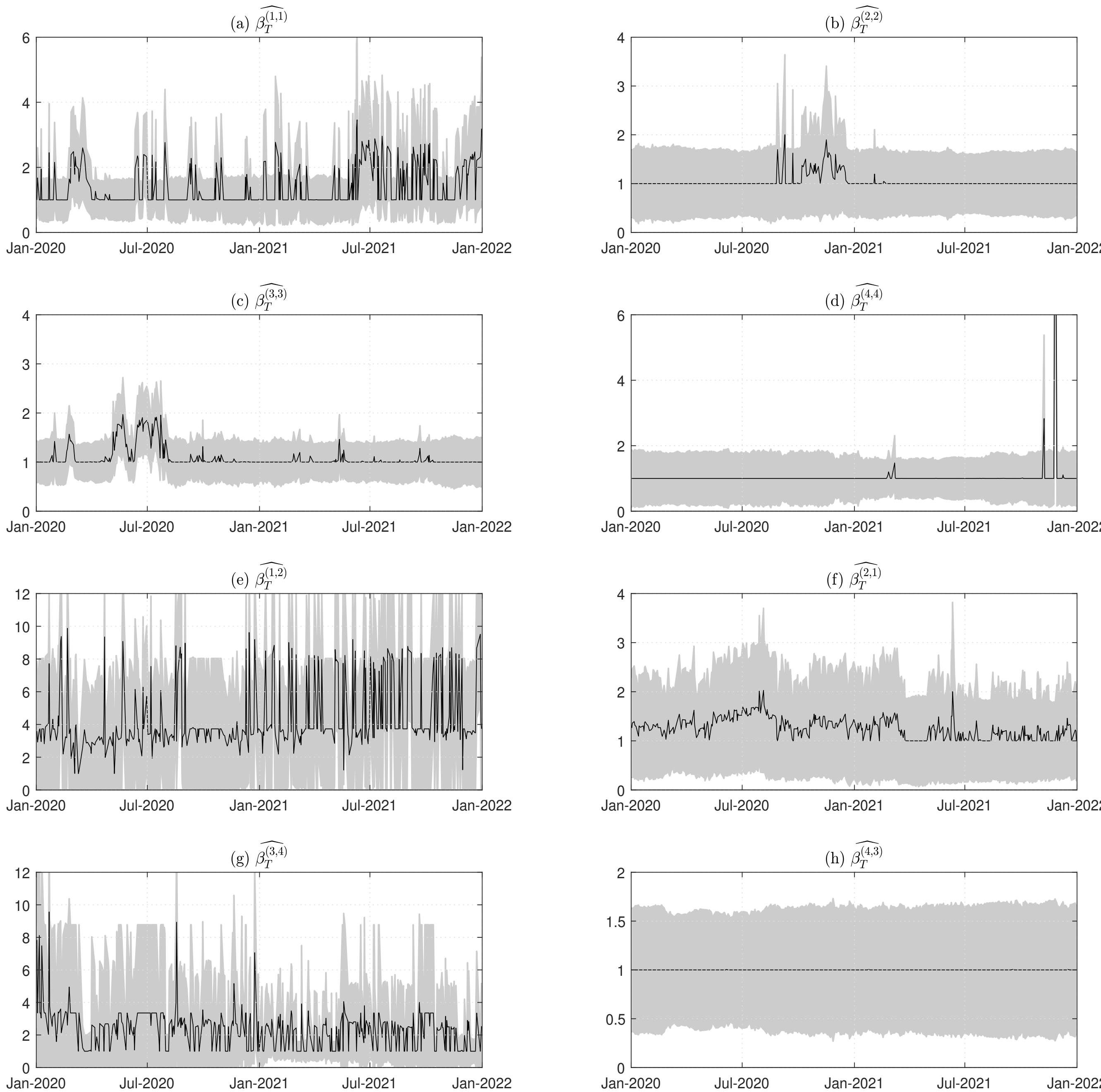

Figure D2: Parameter estimates $\widehat{\beta_T^{(i,j)}}$ for each day for events in the NYSE and TSX. 90% confidence intervals are reported.