# Estimation of latency for Hawkes processes with a polynomial periodic kernel

Deniz Erdemlioglu* Yoann Potiron† Vladimir Volkov‡ and Taiyu Xu§

**Abstract**

We consider estimation of latency, i.e. the time to learn an event and respond. We study Hawkes mutually exciting processes such that their intensity has a parametric form. We assume that the kernel is polynomial and periodic. We define latency as a known function of kernel parameters. Our parametric inference is based on maximum likelihood estimation. We give one central limit theorem for estimation of parameters, a second central limit theorem for estimation of latency and a third central limit theorem for the joint estimation of latency and some parameters. We propose a Wald test statistic of linear hypothesis jointly for latency and some parameters. A numerical study corroborates the asymptotic theory and shows that we improve latency estimation with this more realistic kernel. Our empirical application examines the heterogeneity in central bank communication in the context of monetary policy.

**Keywords**: latency; Hawkes mutually exciting processes; polynomial kernel; periodicity; parametric inference; monetary policy; heterogeneity

*?. E-mail: ?

†(CONTACT) Faculty of Business and Commerce, Keio University. 2-15-45 Mita, Minato-ku, Tokyo, 108-8345, Japan. E-mail: potiron@fbc.keio.ac.jp website: http://www.fbc.keio.ac.jp/~potiron

‡School of Business and Economics, University of Tasmania. E-mail: vladimir.volkov@utas.edu.au

§?. E-mail: ?

## 1 Introduction

This paper concerns estimation of a latency matrix, i.e. the time to learn an event and respond. The latency can also be called a delay. We assume that the latency is a matrix of dimension $d \times d$. In the finance literature, a common definition of latency is based on datasets that are not necessarily available to the statistician (see [Hasbrouck and Saar, 2013]). An alternative definition of latency is using a statistical model based on point processes which characterizes the event times (see [Potiron and Volkov, 2025]). These models of point processes in which the time between two events is random can be seen as a natural extension of standard time series, in which the time between two events is fixed. We rely on the so-called Hawkes mutually exciting processes (see [Hawkes, 1971b] and [Hawkes, 1971a]). We define the point process of dimension $d$ as $N_t$ and its intensity as $\lambda_t$. Then, a standard definition of Hawkes mutually exciting processes is given by

$$\lambda_t = \nu^* + \int_0^t h(t-s)\, dN_s. \tag{1}$$

Here, $\nu^*$ is a Poisson baseline of dimension $d$ and $h$ is a kernel matrix of dimension $d \times d$. If we define $\theta^*$ as the parameters of the kernel, we rely on the parametric specification

$$\lambda_t = \nu^* + \int_0^t h(t-s, \theta^*)\, dN_s. \tag{2}$$

Moreover, we can define the latency matrix of dimension $d \times d$ as a known function $F$ of kernel parameters

$$L = F(\theta^*). \tag{3}$$

More specifically, latency is defined as the time required to reach the pick of the kernel. Since latency is not well-defined with an exponential kernel, [Potiron and Volkov, 2025] study generalized gamma kernels. The main novelty in this paper is that the kernel is more realistic.

The main application of latency lies in finance. [Gagnon and Karolyi, 2010] show that price parity deviations relate positively to proxies for holding costs that can limit arbitrage. The empirical application from [Hasbrouck and Saar, 2013] suggests that high-frequency trading is beneficial to market quality. In [Hoffmann, 2014], fast traders can revise their quotes quickly after news arrivals to reduce market risks. [Budish et al., 2015], [Biais et al., 2015], [Foucault et al., 2016] and [Pagnotta and Philippon, 2018] also consider trading speed. [Potiron and Volkov, 2025] propose estimation of latency. See also [Erdemlioglu et al., 2025] for the time-dependent latency case.

When seen as a delay, there are also applications in management. [Dong et al., 2019] investigate the impact of delay on the coordination within hospitals. [Gámiz et al., 2022], [Gámiz et al., 2023] and [Schoenberg, 2023] consider nonparametric local estimation of Hawkes processes and applications to pandemic. There are also applications in seismology (see [Nolet and Dahlen, 2000]), insurance (see [Lesage et al., 2022]), criminology (see [Nagin and Pogarsky, 2004]), sociology (see [Lahad, 2012]) and medicine (see [Harris, 1990]).

The main reason why Hawkes processes are popular in statistics is that they target the presence of event clustering in time. The main application of Hawkes processes lies in seismology (see [Rubin, 1972], [Vere-Jones, 1978], [Ozaki, 1979], [Vere-Jones and Ozaki, 1982], [Ogata, 1978], [Ogata, 1988]). The impact of earthquake risk is analyzed in [Ikefuji et al., 2022]. There are applications in financial econometrics (see [Yu, 2004], [Bowsher, 2007], [Embrechts et al., 2011], [Aït-Sahalia et al., 2014] and [Corradi et al., 2020]), finance (see [Large, 2007], [Aït-Sahalia et al., 2015] and [Fulop et al., 2015]) and quantitative finance (see [Chavez-Demoulin et al., 2005], [Bacry et al., 2013], but also the papers [Jaisson and Rosenbaum, 2015] and [Morariu-Patrichi and Pakkanen, 2022]). See the references in [Liniger, 2009] and [Hawkes, 2018]. More recently, spectral parametric estimation for misobserved Hawkes processes is given in [Cheysson and Lang, 2022]. [Cavaliere et al., 2023] develop a bootstrap approach. [Clements et al., 2023] study nonparametric estimation. [Christensen and Kolokolov, 2024] propose an unbounded intensity model for more general point processes.

Most papers with Hawkes processes consider at most exponentially decreasing kernels, which are restrictive for applications. In financial applications, there is empirical evidence that the kernel is polynomial (see [Bacry et al., 2012] and [Hardiman et al., 2013]). Moreover, there is periodicity in the data (see ?). Thus, we consider a kernel which is polynomial and periodic. Namely, we study a periodic log-logistic kernel. This allows the latency to be defined as the time required to reach the pick of the kernel which is not necessarily the maximum of the kernel. This is more realistic than the exponential kernel from [Clinet and Yoshida, 2017], the power law kernel from [Cavaliere et al., 2023] and the generalized gamma kernels used in [Potiron and Volkov, 2025].

Our inferential theory builds on in-fill asymptotics when the final time $T$ is finite and the number of observations on $[0, T]$ increases as $n \to \infty$. These asymptotics are popular with financial applications based on high-frequency data (see [Aït-Sahalia and Jacod, 2014]). The main statistical reason why we use these asymptotics is that we observe time-dependent latency between different days (see our empirical study, Figures 1 and 2 in [Potiron and Volkov, 2025]). Thus, we cannot rely on a final time $T$ that increases to infinity with a constant latency. There already exists work in statistics to accommodate for in-fill asymptotics with Hawkes processes. [Chen and Hall, 2013] use random observation times of order $n$. A single boosting of the baseline, i.e. $\lambda_t = \alpha\nu^* + \int_0^t h(t-s,\theta^*)dN_s$, is considered where $\alpha \to \infty$ is a scaling sequence. [Clinet and Potiron, 2018] introduce a joint boosting of the baseline and the kernel, i.e. $\lambda_t = n\nu^* + \int_0^t na^* \exp(-nb^*(t-s))dN_s$. See also [Kwan et al., 2023], [Christensen and Kolokolov, 2024], [Potiron and Volkov, 2025], [Erdemlioglu et al., 2025], [Potiron et al., 2025b] and [Potiron et al., 2025a].

Our parametric inference procedure builds on maximum likelihood estimation. [Ogata, 1978] shows the central limit theorem of the estimation procedure for an ergodic stationary point process. However, the definition of ergodicity is vague in that paper. Most papers on parametric inference for Hawkes processes make this ergodicity assumption (see [Bowsher, 2007], [Large, 2007] and [Cavaliere et al., 2023]). [Clinet and Yoshida, 2017] exhibit the conditions required, i.e. ergodicity of the Hawkes intensity process and its derivative. They consider general point processes and derive the central limit theorem of the estimation procedure in Theorem 3.11 (p. 1809) under these ergodicity assumptions. They also show these ergodicity assumptions in the case of a Hawkes process with exponential kernel in Theorem 4.6 (p. 1821). [Kwan, 2023] considers the non-exponential kernel case and shows the ergodicity for the Hawkes intensity process itself but not for its derivative. Thus, he can only show the consistency of the inference procedure in Theorem 3.4.3 (p. 73). When the kernel follows a generalized gamma distribution, [Potiron and Volkov, 2025] (Theorem 1) show that the ergodicity assumptions are satisfied and also obtain the central limit theorem of the estimation procedure. [Potiron, 2025] extends [Clinet and Yoshida, 2017] and allow for kernels with power distribution, under some smoothness assumptions on the kernel shape.

All these results are useful, but none of them consider Hawkes processes with a periodic log-logistic kernel. In our Theorem 1, we deliver the central limit theorem of the statistical procedure for parameter estimation. This is the main theoretical result of this paper. This extends [Clinet and Yoshida, 2017] (Theorem 4.6, p. 1821) and [Potiron and Volkov, 2025] (Theorem 1). See also [Cavaliere et al., 2023] (Theorem 2, p. 138), who require stronger conditions. We give the central limit theorem for estimation of latency in Corollary 1. This extends Corollary 3 in [Potiron and Volkov, 2025] to the more realistic case of polynomial periodic kernel. To obtain joint tests based on the latency matrix $L$, the baseline parameters $\nu^*$ and the parameters of the kernel which are not used in the definition of latency, we derive the central limit theorem for estimation of the parameter vector $\kappa^* = (\nu^*, \theta_o^*, \overline{L})$ in Proposition 1. This general result is novel to the literature on latency. Finally, we consider Wald tests based on the parameter vector $\kappa^*$. With this framework, we can jointly test for latency and the parameters of the kernel which are not used in the definition of latency $\theta_o$. Corollary 2 shows that the Wald test statistic converges in distribution to a chi-squared distribution with $q$ degrees of freedom under the null hypothesis and is consistent under the alternative hypothesis. This general result is novel to the literature on latency and extends Corollary 6 in [Potiron and Volkov, 2025], which is restricted to the latency vector.

Our proof strategy follows the general machinery of [Clinet and Yoshida, 2017] and [Potiron, 2025], which consider asymptotics when the final time increases, i.e. $T \to \infty$. To rewrite our problem with in-fill asymptotics as a problem with $T \to \infty$, we consider a time transformation. This was already used in [Clinet and Potiron, 2018], [Kwan et al., 2023] and [Potiron and Volkov, 2025]. The main novelty in the proofs is in showing that Hawkes processes with a periodic log-logistic kernel satisfies Assumption

2 from [Potiron, 2025]. This requires to study deeply some smoothness properties of the log-logistic distribution, when seen as a function of its parameters.

We apply our estimation framework to examine empirically the heterogeneity in central bank communication in the context of monetary policy. Specifically, we focus on time-stamped speeches and statements delivered by FOMC members of the U.S. Federal Reserve System. This application is particularly compelling, as it offers an empirical setting that is characterized by potential heterogeneity in speech *timing*, *tone*, *ambiguity* of the messages, and institutional *role* of the speakers (i.e., Fed Chair versus other members). These features, taken together, allow us to explore how different forms of information release interact with the distributional characteristics of latency and to assess the flexibility and performance of our modeling approach under realistic conditions.

To further motivate the empirical relevance of this heterogeneity, consider the case of a speech delivered by an FOMC member—e.g., the Chair of the Federal Reserve—at a standard, prescheduled time (e.g., 14:30). Suppose the message conveyed is both anticipated and unambiguous. For instance, in his March 16, 2022 press conference, Chair Powell opened by stating: "*Today, in support of these goals, the FOMC raised its policy interest rate by 1/4 percentage point.*" (I WILL INSERT THE REF OF THIS SPEECH HERE.) In such cases, where the communication is timely, relatively clear, and broadly expected, market participants are likely to incorporate the information swiftly and uniformly, resulting in a smooth price adjustment. If the same speaker were to deliver a similarly transparent and expected message at the same time one week or one month later, there would still be little reason to expect substantial variation in the market's reaction. That is, when key dimensions such as timing, speaker identity, message tone, and clarity are held constant, the latency of price response is expected to be relatively stable—reflecting what we may call a "normal" pattern of adjustment. Under such homogeneity, the distribution of latency may exhibit light tails, and classical specifications such as the gamma distribution (TO BE CHECKED/REFINED AGAIN) may provide an adequate fit.

However, we argue that this idealized setting is far from representative of the broader data-generating process observed in practice. Central bank communication exhibits considerable heterogeneity along multiple dimensions, each of which may alter the informational content and perceived uncertainty associated with a speech (I WILL INSERT SOME REF). First, the institutional role of the speaker matters: the market impact of a speech by the Fed Chair is not equivalent to that of a regional Fed president or other committee member. Second, the intraday timing of the speech may influence market sensitivity; speeches occurring at irregular or unscheduled times may convey urgency or unexpected information, leading to differential reactions. Third, the tone of the communication—whether hawkish, dovish, or neutral—affects expectations, particularly when sentiment diverges from prior beliefs of market participants. Fourth, and critically, the clarity or ambiguity of the message influences how quickly and confidently market participants can interpret the signal.

Taken together, these dimensions introduce significant variation in the speed and pattern of price adjustment across events. We argue that this heterogeneity is not merely incidental but has structural implications for the shape of the latency distribution itself. In particular, the presence of latent informational frictions and interpretive uncertainty can generate fat tails in the distribution of latency—that is, some events lead to unusually delayed reactions, reflecting slower diffusion of news. We seek to explore these patterns in our empirical analysis and differentiate the forms of latency characteristics embedded in central bank speeches.

The rest of this paper is organized as follows. The setting is introduced in Section 2. The parametric inference procedure is given in Section 3. The theory is developed in Section 4. Our numerical study is carried in Section **??**. Our empirical application is provided in Section **??**. We conclude in Section 5. The supplementary materials contain all the proofs of the manuscript.

## 2 Setting

In this section, we introduce Hawkes mutually exciting processes such that their intensity has a parametric form, latency and a Wald test of linear hypothesis jointly for latency and some parameters.

We start with an introduction to the multidimensional point process $N_t$. For any index $i = 1, \cdots, d$, each component of the point process $N_t^{(i)}$ counts the number of events between 0 and $t$. Here, we denote the i-th component of a vector $V$ by $V^{(i)}$. Then, we define $N^{(i)}$ as a simple point process on $[0,T]$, i.e. a family $\{N^{(i)}(C)\}_{C\in\mathcal{B}([0,T])}$ of random variables with values in the space of natural integers $\mathbb{N}$. Here, $\mathcal{B}([0,T])$ is the Borel $\sigma$-algebra on the compact space $[0,T]$. In addition, we have $N^{(i)}(C)=\sum_{k\in\mathbb{N}}\mathbf{1}_C(\tau_k^{(i)})$ where $\{\tau_k^{(i)}\}_{k\in\mathbb{N}}$ is a sequence of $\mathbb{R}^+$-valued event times, which are random. We assume that the first time is equal to 0 and the following times are increasing for each process a.s., i.e. $\mathbb{P}\big(\tau_0^{(i)}=0<\tau_1^{(i)}<\cdots<\tau_{N_T^{(i)}}^{(i)}<T<\tau_{N_T^{(i)}+1}^{(i)} \text{ for } i=1,\cdots,d\big)=1$. We also assume that no events happen at the same time for different processes a.s., i.e. $\mathbb{P}\big(\tau_k^{(i)}\neq\tau_l^{(j)} \text{ for } k,l\in\mathbb{N}_* \text{ and } i,j=1,\cdots,d \text{ s.t. } i\neq j\big)=1$. Here, we define the space without zero as $S_*$ for any space $S$ such that $0\in S$.

To deliver the definition of intensity, we introduce some more theoretical tools. We suppose that a complete stochastic basis $\mathcal{B}=(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in[0,T]},\mathbb{P})$ describes the evolution of the point process. More specifically, we assume that, for any $t\in[0,T]$, the canonical filtration of $N_t$ is included in the main filtration, i.e. $\mathcal{F}_t^N\subset\mathcal{F}_t$. Here, the canonical filtration of $X_t$ is defined as $\mathcal{F}_t^X=\sigma\big(X(C),C\in\mathcal{B}([0,T]),C\subset[0,t]\big)$ for any process $X_t$. Any nonnegative $\mathcal{F}_t$-progressively measurable process $\{\lambda_t\}_{t\in[0,T]}$, which is of dimension $d$, such that $\mathbb{E}[N((a,b])\mid\mathcal{F}_a]=\mathbb{E}\big[\int_a^b\lambda_s ds\,\big|\,\mathcal{F}_a\big]$ a.s. for all intervals $(a,b]\subset[0,T]$, is called an $\mathcal{F}_t$-intensity of $N_t$. Intuitively, the intensity corresponds to the expected number of events given the past information, i.e.

$$\lambda_t=\lim_{u\to0}\mathbb{E}\Big[\frac{N_{t+u}-N_t}{u}\mid\mathcal{F}_t\Big] \text{ a.s..}$$

For background on point processes, the reader can consult [Jacod, 1975], [Jacod and Shiryaev, 2003], [Daley and Vere-Jones, 2003], and [Daley and Vere-Jones, 2008].

The present work is concerned with Hawkes mutually exciting processes such that their intensity has a parametric form. For a matrix $\phi$, we denote its component $(i,j)$ as $\phi^{(i,j)}$. We introduce the parameter space $\Xi$, consisting of $m$ parameters. We assume that the parameter $\xi$ has the form $\xi=(\nu,\theta)$, and that they belong to the parameter space $\Xi=(\Phi,\Theta)$ where $\nu\in\Phi$ and $\theta\in\Theta$. For any parameter $\xi\in\Xi$, we introduce the family of intensities

$$\lambda_t(\xi)=\nu+\int_0^t h(t-s,\theta)\,dN_s. \tag{4}$$

Here, $h$ is a kernel matrix $d \times d$. The point processes are mutually exciting in the sense that the diagonal components $h^{(i,i)}$ are self-exciting terms for the $i$-th process and non-diagonal components $h^{(i,j)}$ are cross-exciting terms for the $i$-th process made by events from the $j$-th process. Moreover, $\nu$ consists of $d$ baseline parameters, while $\theta$ consists of $m - d$ kernel parameters. We also assume that $m - d \geq d^2$, since at least one parameter should be used in each component of the kernel matrix. Finally, we assume the existence of the true parameter $\xi^* \in \Xi$ such that

$$\lambda_t = \lambda_t(\xi^*). \tag{5}$$

Here again, we assume that the parameter $\xi^*$ has the form $\xi^* = (\nu^*, \theta^*)$, where $\nu^* \in \Phi$ and $\theta^* \in \Theta$.

This paper targets estimation of latency. The latency is defined as a matrix of dimension $d \times d$ which is a known function $F$ of the kernel parameters $\theta^*$, i.e.

$$L = F(\theta^*). \tag{6}$$

With a latency matrix, we can study the latency of each individual process and the latency between two different processes. More specifically, a latency between events from the $j$th process and its impact on events from the $i$th process is introduced at time $t$ if $L^{(i,j)} > 0$. In this paper, we set $F$ such that the latency $L^{(i,j)}$ is equal to the time required to reach the pick of the kernel which is not necessarily the maximum of the kernel. This is more realistic than the definition introduced by [Potiron and Volkov, 2025], which sets latency as the time required to reach the pick of the kernel which is the maximum. See also [Erdemlioglu et al., 2025] for the time-dependent latency case. Moreover, this definition of latency is in agreement with the finance literature, which defines latency as the time it takes to learn and generate response to a trading event (see [Hasbrouck and Saar, 2013]). An advantage of this definition is that latency can be characterized by parameters $\theta^{*,(i,j)}$ associated with factors affecting latency. Such a structural approach permits identification of different aspects of latency. Finally, there is no latency between events from the $j$th process and its impact on events from the $i$th process at time $t$ when $L_t^{(i,j)} \leq 0$.

Since we want to get a standard normal vector in the limit of the central limit theorem, we rewrite the latency matrix $L$ of dimension $d \times d$ as a latency vector $\overline{L} = (L^{(1,1)}, L^{(1,2)}, \cdots, L^{(d,d)})^T$ of dimension $d^2$. We also introduce the parameters of the kernel used in the definition of latency as $\theta_l$, consisting of $l$ parameters. As latency is equal to a function of kernel parameters, we have by definition that $\theta_l \subset \theta$ and $l \leq m - d$. Moreover, we can rewrite the latency function as $L = F(\theta_l^*)$. We also introduce the parameters of the kernel which are not used in the definition of latency as $\theta_o$, which consists of $m - d - l$ parameters. Then, we reorder any parameter $\xi \in \Xi$ and the parameter space $\Xi$ such that $\xi = (\nu, \theta_o, \theta_l)$. To obtain joint tests based on the latency vector $\overline{L}$, the baseline parameters $\nu^*$ and the parameters of the kernel which are not used in the definition of latency as $\theta_o^*$, we consider estimation of the parameter vector $\kappa^* = (\nu^*, \theta_o^*, \overline{L})$ of dimension denoted by $d_\kappa = d^2 + m - l$.

We finally introduce a Wald test of $q$ linear hypotheses on the parameter vector $\kappa^* = (\nu^*, \theta_o^*, \overline{L})$. This test is based on the matrix $R$ of dimension $q \times d_\kappa$. Namely, we define the null hypothesis as $H_0 : \{R\kappa^* = r\}$ and the alternative hypothesis as $H_1 : \{R\kappa^* \neq r\}$ for a real number $r \in \mathbb{R}$. In general, the Wald test assesses constraints on parameters based on the weighted distance between the unrestricted estimate and its value under the null hypothesis. Intuitively, the larger this weighted distance, the less likely it is that the constraint is true. With this framework, we can jointly test for latency and the parameters of the kernel which are not used in the definition of latency $\theta_o$. This extends the tests proposed in [Potiron and Volkov, 2025], which are restricted to the latency vector.

## 3 Inference

In this section, we introduce the in-fill asymptotics, parametric estimation, latency estimation and the Wald test statistic of linear hypothesis jointly for latency and some parameters.

For inference purposes, we consider in-fill asymptotics with joint boosting of the baseline and the kernel, i.e.

$$\lambda_t = n\nu^* + \int_0^t nh(n(t-s), \theta^*)\, dN_s. \tag{7}$$

Here and in what follows, we prefer most of the time not to write explicitly the dependence on $n$.

Also, in-fill asymptotics are based on random observation times of order $n$ within the time interval $[0, T]$ for a finite final time $T$. These in-fill asymptotics, also based on joint boosting, are used in [Clinet and Potiron, 2018], [Kwan et al., 2023], [Kwan, 2023], [Potiron and Volkov, 2025], but also [Potiron et al., 2025b], [Erdemlioglu et al., 2025] and [Potiron et al., 2025a]. They are different from [Chen and Hall, 2013] in-fill asymptotics which considers no boosting of the kernel. There are compatible with [Christensen and Kolokolov, 2024] in-fill asymptotics. The main statistical reason why we use these asymptotics is that we observe time-dependent latency between different days (see our empirical study, Figures 1 and 2 in [Potiron and Volkov, 2025]). Thus, we cannot rely on a final time $T$ that increases to infinity with a constant latency.

The parametric estimation relies on the log likelihood process (see [Ogata, 1978] as well as the book [Daley and Vere-Jones, 2003])

$$l(\xi) = \sum_{i=1}^{d} \int_0^T \log(\lambda_t^{(i)}(\xi)) dN_t^{(i)} - \sum_{i=1}^{d} \int_0^T \lambda_t^{(i)}(\xi) dt. \tag{8}$$

Here, 0 is the starting time and T is the final time. Then, the maximum likelihood estimator is defined as the maximizing parameter of the log likelihood process between the starting time 0 and the final time T, i.e.

$$\widehat{\xi} \in \operatorname{argmax}_{\xi \in \Xi} l(\xi).$$

Here, we have that the estimator $\widehat{\xi}$ has the form $\widehat{\xi} = (\widehat{\nu}, \widehat{\theta})$, where $\widehat{\nu} \in \Phi$ and $\widehat{\theta} \in \Theta$. Finally, we rely on the latency estimator

$$\widehat{L} = F\big(\widehat{\theta}\big). \tag{9}$$

This estimator was introduced by [Potiron and Volkov, 2025]. See also [Erdemlioglu et al., 2025] for the time-dependent latency case.

In this paper, we rely on a time change from the time interval $[0, T]$ to the time interval $[0, nT]$. The main reason is that our proof strategy follows the general machinery of [Potiron, 2025], which consider asymptotics where the final time diverges, i.e. $T \to \infty$. This is the same strategy that

[Potiron and Volkov, 2025] applied with [Clinet and Yoshida, 2017]. More specifically, we introduce the time-changed point process $\overline{N}_t = N_{\frac{t}{n}}$ for any time $t \in [0, nT]$. We also define the rescaled and time-changed intensity process as $\overline{\lambda}_t(\xi) = \frac{\lambda_{t/n}(\xi)}{n}$ for any parameter $\xi \in \Xi$ and any time $t \in [0, nT]$. Lemma 2 in the supplementary materials show that $\overline{\lambda}_t(\xi^*)$ is the intensity of $\overline{N}_t$. By rescaling the intensity, we can show its stability.

On the way to define the asymptotic covariance matrix, we introduce some more notation. In this paper, we focus on the stochastic processes $X_t = (\overline{\lambda}_t(\xi^*), \overline{\lambda}_t(\xi), \partial_\xi \overline{\lambda}_t(\xi))$ defined for $t \in [0, nT]$ and taking values in the space $E^d$ where $E = \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}^m$. Lemma 13 from the supplementary materials states that $X_t$ is stable for any parameter $\xi \in \Xi$, i.e. for any index $i = 1, \cdots, d$ there exists an $\mathbb{R}_+^*$-valued random variable $\overline{\lambda}_l^{(i)}(\xi)$ such that we have $X_{nT}^{(i)} \xrightarrow{\mathcal{D}} \big(\overline{\lambda}_l^{(i)}(\xi^*), \overline{\lambda}_l^{(i)}(\xi), \partial_\xi \overline{\lambda}_l^{(i)}(\xi)\big)$ as $n \to \infty$. Moreover, Lemma 14 from the supplementary materials establishes that $X_t$ is ergodic for any parameter $\xi \in \Xi$, i.e. for any index $i = 1, \cdots, d$ there exists a function $\pi^{(i)} : C_b(E, \mathbb{R}) \to \mathbb{R}$ such that for any $\psi \in C_b(E, \mathbb{R})$ we have $\frac{1}{nT} \int_0^{nT} \psi(X_s^{(i)}) ds \xrightarrow{\mathbb{P}} \pi^{(i)}(\psi)$ as $n \to \infty$. Here, we denote by $\mathcal{C}_b(E, F)$ the space of bounded and continuous functions from $E$ to $F$. The lemma also derives the more explicit expression of the limit function as $\pi^{(i)}(\psi) = \mathbb{E}\big[\psi(\overline{\lambda}_l^{(i)}(\xi^*), \overline{\lambda}_l^{(i)}(\xi), \partial_\xi \overline{\lambda}_l^{(i)}(\xi))\big]$ for any index $i = 1, \cdots, d$.

Since the functions that will be used in the definition of the asymptotic covariance matrix are not bounded, we need to extend from $C_b(E, \mathbb{R})$ to a bigger space $C_\uparrow(E, \mathbb{R})$ the space of functions in which the ergodicity condition holds. More specifically, we denote by $C_\uparrow(E, \mathbb{R})$ the set of continuous functions $\psi : (u, v, w) \to \psi(u, v, w)$ from $E$ to $\mathbb{R}$ that satisfy $\psi$ is of polynomial growth in $u$, $v$, $w$, $\frac{1}{u}$ and $\frac{1}{v}$. This more or less corresponds to Definition 3.7 (p. 1806) in [Clinet and Yoshida, 2017] and Definition 2 in [Potiron, 2025]. Lemma 15 from the supplementary materials extends the starting space of the limit function $\pi$ from $C_b(E, \mathbb{R})$ to $C_\uparrow(E, \mathbb{R})$ and gives a more explicit form. Namely, it shows that, for any index $i = 1, \cdots, d$ and any parameter $\xi \in \Xi$, there exists a probability measure $\pi_\xi^{(i)}$ on $E$ such that, for any $\psi \in C_\uparrow(E, \mathbb{R})$, we have

$$\pi^{(i)}(\psi) = \int_E \psi(u, v, w) \pi_\xi^{(i)}(du, dv, dw).$$

We have now all the ingredients to derive the form of the asymptotic covariance matrix. For a vector $z \in \mathbb{R}^m$, we define its tensor product as $z^{\otimes 2} = z \times z^T \in \mathbb{R}^{m \times m}$. Then, we define the asymptotic Fisher information matrix $\Gamma$ of dimension $m \times m$ as

$$\Gamma = \sum_{i=1}^{d} \int_E w^{\otimes 2} \frac{1}{u} \Pi_{\xi^*}^{(i)}(du, dv, dw). \tag{10}$$

The Fisher information matrix measures the amount of information that the intensity $\lambda_t$ carries about the parameter $\xi^*$. Formally, it is the expected value of the observed information. The Fisher information matrix is used to calculate the covariance matrices associated with maximum likelihood estimation. In other words, $\Gamma^{-1}$ is the asymptotic covariance matrix. The asymptotic Fisher information matrix is estimated from

$$\widehat{\Gamma} = -\partial_\xi^2 \bar{l}(\widehat{\xi}). \tag{11}$$

Here, we define the log likelihood process of the time-changed point process $\overline{N}_t$ between the starting time 0 and the final time $nT$ as $\bar{l}(\xi) = \sum_{i=1}^d \int_0^{Tn} \log(\overline{\lambda}_t^{(i)}(\xi)) d\overline{N}_t^{(i)} - \sum_{i=1}^d \int_0^{Tn} \overline{\lambda}_t^{(i)}(\xi) dt$ and its Hessian matrix as $\partial_\xi^2 \bar{l}(\xi)$. This is a natural estimator since we can reexpress the Fisher information matrix as $\Gamma = -\lim_{n \to \infty} \frac{1}{Tn} \mathbb{E}\big[\partial_\xi^2 \bar{l}(\xi^*)\big]$. This corresponds exactly to the estimation procedure used in [Potiron and Volkov, 2025]. This is based on [Clinet and Yoshida, 2017] who only considers the case when the final time diverges, i.e. $T \to \infty$.

We also rewrite the latency estimator matrix $\widehat{L}$ of dimension $d \times d$ as a latency estimator vector $\widehat{\overline{L}} = (\widehat{L}^{(1,1)}, \widehat{L}^{(1,2)}, \cdots, \widehat{L}^{(d,d)})^T$ of dimension $d^2$. We denote the kernel estimator $\widehat{\theta}$ restricted to the latency parameter $\theta_l^*$ by $\widehat{\theta}_l$. We also denote the Fisher information matrix of dimension $l \times l$ restricted to the latency parameter $\theta_l^*$ and its estimator by $\Gamma_l$ and $\widehat{\Gamma}_l$. For any $i = 1, \cdots, d$ and $j = 1, \cdots, d$ we define the differential vector of the latency function $F^{(i,j)}$ at the latency parameter $\theta_l$ as $dF^{(i,j)}(\theta_l) = (dF^{(i,j,1)}(\theta_l), \cdots, dF^{(i,j,l)}(\theta_l))$, which is of dimension $l$. Moreover, we introduce the asymptotic covariance matrix $\overline{\Gamma}^{-1}$ of dimension $d^2 \times d^2$ for the latency estimator vector satisfying

$$\big(\overline{\Gamma}^{-1}\big)^{((i-1)d+j,(k-1)d+u)} = \sum_{q=1}^{l} \Big( \sum_{r=1}^{l} dF^{(i,j,r)}(\theta_l^*) \big(\Gamma_l^{-1/2}\big)^{(r,q)} \Big) \Big( \sum_{r=1}^{l} dF^{(k,u,r)}(\theta_l^*) \big(\Gamma_l^{-1/2}\big)^{(r,q)} \Big), \tag{12}$$

for any $i = 1, \cdots, d$, $j = 1, \cdots, d$, $k = 1, \cdots, d$ and $u = 1, \cdots, d$. Finally, we propose estimation of the asymptotic covariance matrix as

$$\big(\widehat{\overline{\Gamma}}^{-1}\big)^{((i-1)d+j,(k-1)d+u)} = \sum_{q=1}^{l}\Big(\sum_{r=1}^{l} dF^{(i,j,r)}(\widehat{\theta}_l)\big(\widehat{\Gamma}_l^{-1/2}\big)^{(r,q)}\Big)\Big(\sum_{r=1}^{l} dF^{(k,u,r)}(\widehat{\theta}_l)\big(\widehat{\Gamma}_l^{-1/2}\big)^{(r,q)}\Big), \quad (13)$$

for any $i = 1, \cdots, d$, $j = 1, \cdots, d$, $k = 1, \cdots, d$ and $u = 1, \cdots, d$.

Moreover, we propose estimation of the vector $\kappa^* = (\nu^*, \theta_o^*, \overline{L})$ with $\widehat{\kappa} = (\widehat{\nu}, \widehat{\theta}_o, \widehat{\overline{L}})$. Here, we denote the kernel estimator $\widehat{\theta}$ restricted to the parameters which are not used in the definition of latency $\theta_o^*$ by $\widehat{\theta}_o$. Then, we introduce the asymptotic covariance matrix $\Gamma_\kappa^{-1}$ of dimension $d_\kappa \times d_\kappa$ for the vector estimator $\widehat{\kappa}$ satisfying

$$\begin{aligned} \big(\Gamma_\kappa^{-1}\big)^{(i,j)} &= \big(\Gamma^{-1}\big)^{(i,j)} \text{ for } i = 1, \cdots, m-l, j = 1, \cdots, m-l, & (14)\\ \big(\Gamma_\kappa^{-1}\big)^{(m-l+i,m-l+j)} &= \big(\overline{\Gamma}^{-1}\big)^{(i,j)} \text{ for } i = 1, \cdots, d^2, j = 1, \cdots, d^2, \\ \big(\Gamma_\kappa^{-1}\big)^{(m-l+(i-1)d+j,k)} &= \big(\Gamma_\kappa^{-1}\big)^{(k,m-l+(i-1)d+j)} = \sum_{q=1}^{l}\Big(\sum_{r=1}^{l} dF^{(i,j,r)}(\theta_l^*)\big(\Gamma_l^{-1/2}\big)^{(r,q)}\Big)\big(\Gamma^{-1/2}\big)^{(m-l+q,k)} \\ &\qquad \text{for } i = 1, \cdots, d, j = 1, \cdots, d, k = 1, \cdots, m-l. \end{aligned}$$

Finally, we propose estimation of the asymptotic covariance matrix $\Gamma_\kappa^{-1}$ as

$$\begin{aligned} \big(\widehat{\Gamma}_\kappa^{-1}\big)^{(i,j)} &= \big(\widehat{\Gamma}^{-1}\big)^{(i,j)} \text{ for } i = 1, \cdots, m-l, j = 1, \cdots, m-l, & (15)\\ \big(\widehat{\Gamma}_\kappa^{-1}\big)^{(m-l+i,m-l+j)} &= \big(\widehat{\overline{\Gamma}}^{-1}\big)^{(i,j)} \text{ for } i = 1, \cdots, d^2, j = 1, \cdots, d^2, \\ \big(\widehat{\Gamma}_\kappa^{-1}\big)^{(m-l+(i-1)d+j,k)} &= \big(\widehat{\Gamma}_\kappa^{-1}\big)^{(k,m-l+(i-1)d+j)} = \sum_{q=1}^{l}\Big(\sum_{r=1}^{l} dF^{(i,j,r)}(\widehat{\theta}_l)\big(\widehat{\Gamma}_l^{-1/2}\big)^{(r,q)}\Big)\big(\widehat{\Gamma}^{-1/2}\big)^{(m-l+q,k)} \\ &\qquad \text{for } i = 1, \cdots, d, j = 1, \cdots, d, k = 1, \cdots, m-l. \end{aligned}$$

Since we have proposed estimation of the asymptotic covariance matrix $\Gamma_\kappa^{-1}$, we can introduce our Wald test statistic

$$S = nT\big(R\widehat{\kappa} - r\big)^T\big(R\widehat{\Gamma}_\kappa^{-1}R^T\big)^{-1}\big(R\widehat{\kappa} - r\big). \quad (16)$$

Here, $R^T$ denotes the transpose matrix of the matrix $R$. The Wald test statistic relies on two approximations, namely the asymptotic covariance matrix estimator $\widehat{\Gamma}_\kappa^{-1}$ and the vector parameter estimator $\widehat{\kappa}$.

## 4 Theory

In this section, we introduce the periodic log-logistic kernel. In our Theorem 1, we deliver the central limit theorem of the statistical procedure. This is the main theoretical result of this paper. In particular, we provide conditions such that Hawkes processes with a periodic log-logistic kernel satisfies Assumption 2 from [Potiron, 2025]. This requires to study deeply some smoothness properties of the log-logistic distribution, when seen as a function of its parameters. Moreover, we give the central limit theorem for estimation of latency in Corollary 1. In addition, we derive the central limit theorem for estimation of the parameter vector $\kappa^*$ in Proposition 1. Finally, Corollary 2 shows that the Wald test statistic converges in distribution to a chi-squared distribution with $q$ degrees of freedom under the null hypothesis and is consistent under the alternative hypothesis. All these results build on in-fill asymptotics when $n \to \infty$ and the final time $T$ is finite.

We first introduce the periodic log-logistic kernel. For any $i = 1, \cdots, d$ and $j = 1, \cdots, d$, we define the component $(i,j)$ of the periodic log-logistic kernel as

$$h^{(i,j)}(t, \theta^{(i,j)}) = \gamma^{(i,j)}(1 + A^{(i,j)} \cos(\pi^{(i,j)} t)) \frac{\beta^{(i,j)} t^{\beta^{(i,j)}-1}}{(\alpha^{(i,j)})^{\beta^{(i,j)}}(1 + (t/\alpha^{(i,j)})^{\beta^{(i,j)}})^2}. \tag{17}$$

Here, $\gamma^{(i,j)} \in \mathbb{R}_+^*$ is the level of intensity, $A^{(i,j)} \in \mathbb{R}_+^*$ is the amplitude of the oscillation and $\pi^{(i,j)} \in \mathbb{R}_+^*$ is the period of the oscillation. Moreover, $\alpha^{(i,j)} \in \mathbb{R}_+^*$ is the scale parameter and $\beta^{(i,j)} \in \mathbb{R}_+^*$ is the shape parameter from the log-logistic distribution. We assume that the kernel parameter is of the form

$$\theta = (\theta^{(i,j)})_{1 \leq i,j \leq d} = (\theta^{(1,1)}, \theta^{(1,2)}, \cdots, \theta^{(d,d-1)}, \theta^{(d,d)}) \tag{18}$$

$$\theta^{(i,j)} = (\gamma^{(i,j)}, A^{(i,j)}, \pi^{(i,j)}, \alpha^{(i,j)}, \beta^{(i,j)}).$$

We study a periodic log-logistic kernel, which is polynomially decreasing. This complements the exponential kernel from [Clinet and Yoshida, 2017], the power law kernel from [Cavaliere et al., 2023] and the mixture of generalized gamma kernel used in [Potiron and Volkov, 2025], which is exponentially decreasing. We consider this more general kernel as there is empirical evidence that the kernel decays as the power distribution and periodicity in finance. It is also adapted when examining the heterogeneity in central bank communication in the context of monetary policy.

From Equations (18), we have that the true kernel parameter is of the form $\theta^* = (\theta^{*,(i,j)})_{1\leq i,j\leq d} = (\theta^{*,(1,1)}, \theta^{*,(1,2)}, \cdots, \theta^{*,(d,d-1)}, \theta^{*,(d,d)})$, where $\theta^{*,(i,j)} = (\gamma^{*,(i,j)}, A^{*,(i,j)}, \pi^{*,(i,j)}, \alpha^{*,(i,j)}, \beta^{*,(i,j)})$. Then, latency is naturally defined as one half of the period, i.e. for any $i = 1, \cdots, d$ and $j = 1, \cdots, d$

$$L^{*,(i,j)} = \frac{\pi^{*,(i,j)}}{2}.$$

In other words, we have $\theta_l = \pi$ and $F^{(i,j)}(\theta_l) = \pi^{(i,j)}/2$. In our model, we necessarily have that latency is positive, i.e. $L^{*,(i,j)} > 0$. This is a technical limitation as the case when the latency is null is degenerate, with a decrease in the number of parameters from the model (18). However, this allows us to derive joint general results on the parameter vector $\kappa^*$. Moreover, the empirical application in [Potiron and Volkov, 2025] documents the presence of positive latency when used on financial data.

Before introducing the conditions, we first need to introduce some notation. We define $\overline{\Xi}$ as the closure space of the parameter space $\Xi$. For a matrix $\phi$, we denote its spectral radius as $\rho(\phi)$. For any time $t \in \mathbb{R}^+$, we denote by $\theta_t^+$ the maximum argument parameter of the kernel spectral radius $\rho(h(t,\theta))$. It is defined through

$$h(t, \theta_t^+) = \sup_{\theta\in\Theta} \rho(h(t,\theta)). \tag{19}$$

Then, we define the matrix $\phi$ of dimension $d \times d$ as the integral of $h(t, \theta_t^+)$ over time, i.e.

$$\phi = \int_0^\infty h(s, \theta_s^+) ds.$$

For a vector or a matrix $V$ of dimension $k$, we denote its $L^1$ norm as $|V| = \sum_{i=1}^k |V^{(i)}|$. Moreover, $\partial_\theta G(\theta)$ denotes the vector of partial derivatives for any function $G(\theta)$. For any $i = 1, \cdots, d$, any $j = 1, \cdots, d$ and any time $t \in \mathbb{R}^+$, we denote by $k_{t,3}^{(i,j)}$ the maximum index argument for the $L^1$ norm of the kernel partial derivatives $\left|\partial_\theta h^{(i,j)}(t,\theta)^{(k)}\right|$. It is defined through

$$\left|\partial_\theta h^{(i,j)}(t,\theta)^{(k_{t,3}^{(i,j)})}\right| = \sup_{k=1,\cdots,d} \left|\partial_\theta h^{(i,j)}(t,\theta)^{(k)}\right|. \tag{20}$$

Then, we define the matrix $\phi_3(\theta)$ of dimension $d\times d$ as as the integral of $\left|\partial_\theta h^{(i,j)}(t,\theta)^{(k_{t,3}^{(i,j)})}\right|$ over time, i.e.

$$\phi_3^{(i,j)}(\theta) = \int_0^\infty \left|\partial_\theta h^{(i,j)}(s,\theta)^{(k_{t,3}^{(i,j)})}\right| ds,$$

for any $i = 1, \cdots, d$ and any $j = 1, \cdots, d$. For any $i = 1, \cdots, d$, any $j = 1, \cdots, d$ and any $t \in \mathbb{R}^+$, we denote by $(k_{t,4}^{(i,j)}, l_{t,4}^{(i,j)})$ the maximum component argument for the $L^1$ norm of the kernel Hessian matrix $\left|\partial_\theta^2 h^{(i,j)}(t,\theta)^{(k,l)}\right|$. It is defined through

$$\left|\partial_\theta^2 h^{(i,j)}(t,\theta)^{(k_{t,4}^{(i,j)}, l_{t,4}^{(i,j)})}\right| = \sup_{k,l=1,\cdots,d} \left|\partial_\theta^2 h^{(i,j)}(t,\theta)^{(k,l)}\right|. \tag{21}$$

Then, we define the matrix $\phi_4(\theta)$ of dimension $d \times d$ as the integral of $\left|\partial_\theta^2 h^{(i,j)}(t,\theta)^{(k_{t,4}^{(i,j)}, l_{t,4}^{(i,j)})}\right|$ over time, i.e.

$$\phi_4^{(i,j)}(\theta) = \int_0^\infty \left|\partial_\theta^2 h^{(i,j)}(s,\theta)^{(k_{t,4}^{(i,j)}, l_{t,4}^{(i,j)})}\right| ds,$$

for any $i = 1, \cdots, d$ and $j = 1, \cdots, d$.

We make the following set of conditions for the central limit theorem of the statistical procedure for parameter estimation.

*Condition* 1. (a) The parameter space $\Xi \subset \mathbb{R}^m$ is such that its closure $\overline{\Xi}$ is a compact space.

(b) There exists $\nu_- > 0$ such that for any $\nu \in \Phi$ and any $i = 1, \cdots, d$ we have $\nu^{(i)} > \nu_-$.

(c) For any $\theta \in \Theta$, we have that the kernel $h(t,\theta)$ follows (17) and its parameter $\theta$ satisfies (18).

(d) There exists $A_+ \in (0,1)$ such that for any $i = 1, \cdots, d$ and any $j = 1, \cdots, d$ we have $A^{(i,j)} < A_+$.

(e) There exists $\alpha_- > 0$ such that for any $i = 1, \cdots, d$ and any $j = 1, \cdots, d$ we have $\alpha^{(i,j)} > \alpha_-$.

(f) There exists $\beta_- > 0$ such that for any $i = 1, \cdots, d$ and any $j = 1, \cdots, d$ we have $\beta^{(i,j)} > \beta_-$.

(g) We have the spectral norm of the matrix $\phi$ is smaller than unity, i.e. $\rho(\phi) < 1$.

(h) For any parameter $\theta \in \Theta$, we have $\rho(\phi_3(\theta)) < 1$ and $\rho(\phi_4(\theta)) < 1$.

Condition 1 (a) corresponds exactly to Assumption 2 (a) in [Potiron, 2025] and is weaker than the framework from [Clinet and Yoshida, 2017] and [Potiron and Volkov, 2025] where the parameter space satisfies the assumptions from the Sobolev embedding theorem. It is often necessary to restrict on a compact space to obtain consistency of the statistical procedure based on maximum likelihood estimation. Condition 1 (b) imply that the point processes are well-defined and are also required in the simpler case of heterogeneous Poisson processes without a kernel (see [Daley and Vere-Jones, 2003]). Condition 1 (c) restricts to Hawkes processes with log-logistic kernel, which is more realistic than the case with exponential Hawkes processes from [Clinet and Yoshida, 2017], the power law kernel case from [Cavaliere et al., 2023] and the case with mixture of generalized gamma kernels (see Condition [A] (ii) in [Potiron and Volkov, 2025]). Conditions 1 (d), (e) and (f) put restriction on the parameter space.

The remaining two conditions are a bit more abstract. Condition 1 (g) states that the spectral radius of the kernel integral when evaluated at the maximum argument parameter of $\rho(h(t,\theta))$ is strictly smaller than unity. This is slightly stronger than the condition which is necessary to obtain a stationary intensity with finite first moment (see Lemma 1 (p. 495) in [Hawkes and Oakes, 1974] and Theorem 1 (p. 1567) in [Brémaud and Massoulié, 1996]). Nonetheless, Condition 1 (g) is very close to the stationary condition in practice since our parameter space is included in a compact space. Moreover, it is much weaker than Assumption 2 (d) in [Potiron, 2025].

Finally, the case $\rho(\phi_3(\theta)) < \phi_+$ in Condition 1 (h) states that the spectral radius of the kernel derivative integral when evaluated at the maximum argument of $\left|\partial_\theta h^{(i,j)}(t,\theta)^{(k)}\right|$ is strictly smaller than $\phi_+$ uniformly in the space parameter value. The case $\rho(\phi_4(\theta)) < \phi_+$ in Condition 1 (h) ensures that the spectral radius of the kernel second derivative integral when evaluated at the maximum argument of $\left|\partial_\theta^2 h^{(i,j)}(t,\theta)^{(k,l)}\right|$ is strictly smaller than $\phi_+$ uniformly in the space parameter value. Condition 1 (h) ensures that the kernel shape is smooth enough uniformly in the parameter space and corresponds exactly to Assumption 2 (g) in [Potiron, 2025]. In practice, Condition 1 (h) does not limit more than Condition 1 (g) since the kernel is polynomially decreasing.

The theorem that follows is the main theoretical result of this paper. It delivers the central limit theorem of the statistical procedure for parameter estimation. This is based on Hawkes processes with parametric intensity and polynomial periodic kernel. The parametric inference relies on maximum likelihood estimation. We consider in-fill asymptotics when $n \to \infty$ and the final time $T$ is finite. This extends [Clinet and Yoshida, 2017] (Theorem 4.6, p. 1821) and [Potiron and Volkov, 2025] (Theorem 1). See also [Cavaliere et al., 2023] (Theorem 2, p. 138), who require stronger conditions. In the theorem and what follows, $\upsilon$ is defined as standard normal vector of dimension m.

**Theorem 1.** *We assume that Condition 1 holds. We have the central limit theorem of the statistical procedure based on Hawkes processes with parametric intensity and polynomial periodic kernel, i.e.*

$$\sqrt{n}(\widehat{\xi} - \xi^*) \xrightarrow{\mathcal{D}} \sqrt{T}\Gamma^{-1/2}\upsilon. \tag{22}$$

*Moreover, we have the feasible normalized central limit theorem*

$$\widehat{\Gamma}^{1/2}\sqrt{nT}(\widehat{\xi} - \xi^*) \xrightarrow{\mathcal{D}} \upsilon. \tag{23}$$

We introduce the matrix of dimension $d^2 \times l$

$$M^{((i-1)d+j,q)} = \sum_{r=1}^{l} dF^{(i,j,r)}(\theta_l^*)\big(\Gamma_l(\theta_l^*)^{-1/2}\big)^{(r,q)},$$

for any $i = 1, \cdots, d$, any $j = 1, \cdots, d$ and any $q = 1, \cdots, m - d$. We denote the space of latency parameters by $\Theta_l$. We now give a set of conditions required for the central limit theorem of latency estimation.

*Condition* 2. (a) The latency function $F : \Theta_l \to \mathbb{R}^{d\times d}$ is twice continuously differentiable.

(b) The matrix $M$ has rank $d^2$.

Condition 2 (a) puts some regular smoothness restrictions on $F$ that are needed to apply Taylor expansions in the proofs. Condition 2 (b) ensures the existence of a standard normal vector in the limit of the central limit theorem. In practice, this implies that $d^2 \leq l$. However, this condition is automatically satisfied since we use at least one parameter for each component of the latency matrix. Condition 2 already appears in [Potiron and Volkov, 2025] (Conditions [B] and [C]).

We now turn our attention to the central limit theorem for estimation of latency. The results are obtained with in-fill asymptotics when $n \to \infty$ and the final time $T$ is finite. This extends Corollary 3 in [Potiron and Volkov, 2025] to the more realistic case of polynomial periodic kernel. Before stating the corollary, we introduce the standard normal vector $\overline{\upsilon}$ of dimension $m^2$.

**Corollary 1.** *We assume that Conditions 1 and 2 hold. We have the central limit theorem for latency estimation based on Hawkes processes with parametric intensity and polynomial periodic kernel, i.e.*

$$\sqrt{n}(\widehat{L} - L) \xrightarrow{\mathcal{D}} \sqrt{T}\overline{\Gamma}^{-1/2}\overline{\upsilon}. \tag{24}$$

*Moreover, we have the feasible normalized central limit theorem*

$$\widehat{\overline{\Gamma}}^{1/2}\sqrt{nT}(\widehat{L} - L) \xrightarrow{\mathcal{D}} \overline{\upsilon}. \tag{25}$$

To obtain joint tests based on the latency vector $\overline{L}$, the baseline parameters $\nu^*$ and the parameters of the kernel which are not used in the definition of latency $\theta_o^*$, we provide the central limit theorem for estimation of the parameter vector $\kappa^* = (\nu^*, \theta_o^*, \overline{L})$. The results are obtained with in-fill asymptotics when $n \to \infty$ and the final time $T$ is finite. This general result is novel to the literature on latency. Before stating the proposition, we introduce the standard normal vector $\upsilon_\kappa$ of dimension $d_\kappa$.

**Proposition 1.** *We assume that Conditions 1 and 2 hold. We have the central limit theorem for estimation of the vector $\kappa^* = (\nu^*, \theta_o^*, \overline{L})$ based on Hawkes processes with parametric intensity and polynomial periodic kernel, i.e.*

$$\sqrt{n}(\widehat{\kappa} - \kappa^*) \xrightarrow{\mathcal{D}} \sqrt{T}\Gamma_\kappa^{-1/2}\upsilon_\kappa. \tag{26}$$

*Moreover, we have the feasible normalized central limit theorem*

$$\widehat{\Gamma}_\kappa^{1/2}\sqrt{nT}(\widehat{\kappa} - \kappa^*) \xrightarrow{\mathcal{D}} \upsilon_\kappa. \tag{27}$$

Finally, we consider joint tests based on the latency vector $\overline{L}$, the baseline parameters $\nu^*$ and the parameters of the kernel which are not used in the definition of latency $\theta_o^*$. The following corollary shows that the Wald test statistic $S$ converges in distribution to a chi-squared distribution with $q$ degrees of freedom under the null hypothesis and is consistent under the alternative hypothesis. We consider in-fill asymptotics when $n \to \infty$ and the final time $T$ is finite. This general result is novel to the literature on latency and extends Corollary 6 in [Potiron and Volkov, 2025], which is restricted to the latency vector. In the corollary, we define $Q(u)$ as the quantile function of the chi-squared distribution with $q$ degrees of freedom.

**Corollary 2.** *We assume that Conditions 1 and 2 hold. Then, the test statistic $S$ converges in distribution to a chi-squared random variable with $q$ degrees of freedom under the null hypothesis $H_0$. The test statistic $S$ is consistent under the alternative hypothesis $H_1$, i.e. we have $\mathbb{P}(S > Q(u) \mid H_1) \to 1$ for any $0 < u < 1$.*

## 5 Conclusion

In this paper, we have studied estimation of latency. We have considered Hawkes mutually exciting processes such that their intensity has a parametric form. We also assumed that that the kernel is polynomial and periodic. We defined latency as a known function of kernel parameters. Our parametric inference was based on maximum likelihood estimation. We have given three central limit theorems for the estimation procedure. We have proposed a Wald test statistic. A numerical study corroborated the asymptotic theory and showed that we improved latency estimation with this more realistic kernel. Our empirical application examined.

## Supplementary materials

All proofs of the theory can be found in the supplementary materials. These proofs are based on

## Funding

Financial support from the Japanese Society for the Promotion of Science under grant 23H00807 (Potiron) is gratefully acknowledged.

## References

[Aït-Sahalia et al., 2015] Aït-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117:585–606.

[Aït-Sahalia and Jacod, 2014] Aït-Sahalia, Y. and Jacod, J. (2014). *High-frequency financial econometrics*. Princeton University Press.

[Aït-Sahalia et al., 2014] Aït-Sahalia, Y., Laeven, R., and Pelizzon, L. (2014). Mutual excitation in eurozone sovereign CDS. *Journal of Econometrics*, 183:151–167.

[Bacry et al., 2012] Bacry, E., Dayri, K., and Muzy, J.-F. (2012). Non-parametric kernel estimation for symmetric Hawkes processes. Application to high frequency financial data. *The European Physical Journal B*, 85(5):1–12.

[Bacry et al., 2013] Bacry, E., Delattre, S., Hoffmann, M., and Muzy, J.-F. (2013). Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Processes and their Applications*, 123(7):2475–2499.

[Biais et al., 2015] Biais, B., Foucault, T., and Moinas, S. (2015). Equilibrium fast trading. *Journal of Financial Economics*, 116(2):292–313.

[Bowsher, 2007] Bowsher, C. G. (2007). Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics*, 141(2):876–912.

[Brémaud and Massoulié, 1996] Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *Annals of Probability*, pages 1563–1588.

[Budish et al., 2015] Budish, E., Cramton, P., and Shim, J. (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics*, 130(4):1547–1621.

[Cavaliere et al., 2023] Cavaliere, G., Lu, Y., Rahbek, A., and Stærk-Østergaard, J. (2023). Bootstrap inference for Hawkes and general point processes. *Journal of Econometrics*, 235:133–165.

[Chavez-Demoulin et al., 2005] Chavez-Demoulin, V., Davison, A., and McNeil, A. (2005). Estimating value-at-risk: a point process approach. *Quantitative Finance*, 5:227–234.

[Chen and Hall, 2013] Chen, F. and Hall, P. (2013). Inference for a nonstationary self-exciting point process with an application in ultra-high frequency financial data modeling. *Journal of Applied Probability*, 50(4):1006–1024.

[Cheysson and Lang, 2022] Cheysson, F. and Lang, G. (2022). Spectral estimation of Hawkes processes from count data. *Annals of Statistics*, 50(3):1722–1746.

[Christensen and Kolokolov, 2024] Christensen, K. and Kolokolov, A. (2024). An unbounded intensity model for point processes. *Journal of Econometrics*, 244(1):105840.

[Clements et al., 2023] Clements, A., Hurn, A., Lindsay, K., and Volkov, V. (2023). Estimating a non-parametric memory kernel for mutually-exciting point processes. *Journal of Financial Econometrics*, 21(5):1759–1790.

[Clinet and Potiron, 2018] Clinet, S. and Potiron, Y. (2018). Statistical inference for the doubly stochastic self-exciting process. *Bernoulli*, 24(4B):3469–3493.

[Clinet and Yoshida, 2017] Clinet, S. and Yoshida, N. (2017). Statistical inference for ergodic point processes and application to limit order book. *Stochastic Processes and their Applications*, 127:1800–1839.

[Corradi et al., 2020] Corradi, V., Distaso, W., and Fernandes, M. (2020). Testing for jump spillovers without testing for jumps. *Journal of the American Statistical Association*, 115:1214–1226.

[Daley and Vere-Jones, 2003] Daley, D. J. and Vere-Jones, D. (2003). *An introduction to the theory of point processes: Elementary theory and methods*, volume 1. Springer New York, NY, 2nd edition.

[Daley and Vere-Jones, 2008] Daley, D. J. and Vere-Jones, D. (2008). *An introduction to the theory of point processes: General theory and structure*, volume 2. Springer New York, NY, 2nd edition.

[Dong et al., 2019] Dong, J., Yom-Tov, E., and Yom-Tov, G. B. (2019). The impact of delay announcements on hospital network coordination and waiting times. *Management Science*, 65(5):1969–1994.

[Embrechts et al., 2011] Embrechts, P., Liniger, T., and Lin, L. (2011). Multivariate Hawkes processes: an application to financial data. *Journal of Applied Probability*, 48:367–378.

[Erdemlioglu et al., 2025] Erdemlioglu, D., Potiron, Y., Volkov, V., and Xu, T. (2025). Estimation of time-dependent latency with locally stationary Hawkes processes. *Working paper available at https://www.fbc.keio.ac.jp/$\sim$ potiron/Erdemlioglu2025workingpaper.pdf*.

[Foucault et al., 2016] Foucault, T., Hombert, J., and Roşu, I. (2016). News trading and speed. *The Journal of Finance*, 71(1):335–382.

[Fulop et al., 2015] Fulop, A., Li, J., and Yu, J. (2015). Self-exciting jumps, learning, and asset pricing implications. *The Review of Financial Studies*, 28(3):876–912.

[Gagnon and Karolyi, 2010] Gagnon, L. and Karolyi, G. (2010). Multi-market trading and arbitrage. *Journal of Financial Economics*, 97:53–80.

[Gámiz et al., 2022] Gámiz, M., Mammen, E., Martínez-Miranda, M., and Nielsen, J. (2022). Missing link survival analysis with applications to available pandemic data. *Computational Statistics & Data Analysis*, 169:107405.

[Gámiz et al., 2023] Gámiz, M., Mammen, E., Martínez-Miranda, M., and Nielsen, J. (2023). Monitoring a developing pandemic with available data. *arXiv preprint arXiv:2308.09919*.

[Hardiman et al., 2013] Hardiman, S. J., Bercot, N., and Bouchaud, J.-P. (2013). Critical reflexivity in financial markets: a Hawkes process analysis. *The European Physical Journal B*, 86(10):1–9.

[Harris, 1990] Harris, J. (1990). Reporting delays and the incidence of aids. *Journal of the American Statistical Association*, 85:915–924.

[Hasbrouck and Saar, 2013] Hasbrouck, J. and Saar, G. (2013). Low-latency trading. *Journal of Financial Markets*, 16:646–679.

[Hawkes, 2018] Hawkes, A. (2018). Hawkes processes and their applications to finance: a review. *Quantitative Finance*, 18(2):193–198.

[Hawkes, 1971a] Hawkes, A. G. (1971a). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 438–443.

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

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

[Hoffmann, 2014] Hoffmann, P. (2014). A dynamic limit order market with fast and slow traders. *Journal of Financial Economics*, 113:156–169.

[Ikefuji et al., 2022] Ikefuji, M., Laeven, R., Magnus, J., and Yue, Y. (2022). Earthquake risk embedded in property prices: Evidence from five japanese cities. *Journal of the American Statistical Association*, 117:82–93.

[Jacod, 1975] Jacod, J. (1975). Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 31(3):235–253.

[Jacod and Shiryaev, 2003] Jacod, J. and Shiryaev, A. (2003). *Limit theorems for stochastic processes*. Springer Berlin, Heidelberg, 2nd edition.

[Jaisson and Rosenbaum, 2015] Jaisson, T. and Rosenbaum, M. (2015). Limit theorems for nearly unstable Hawkes processes. *Annals of Applied Probability*, 25(2):600–631.

[Kwan, 2023] Kwan, T. (2023). *Asymptotic analysis and ergodicity of the Hawkes process and its extensions*. PhD thesis, UNSW Sydney.

[Kwan et al., 2023] Kwan, T., Chen, F., and Dunsmuir, W. (2023). Alternative asymptotic inference theory for a nonstationary Hawkes process. *Journal of Statistical Planning and Inference*, 227:75–90.

[Lahad, 2012] Lahad, K. (2012). Singlehood, waiting, and the sociology of time 1. In *Sociological Forum*, volume 27, pages 163–186. Wiley Online Library.

[Large, 2007] Large, J. (2007). Measuring the resiliency of an electronic limit order book. *Journal of Financial Markets*, 10:1–25.

[Lesage et al., 2022] Lesage, L., Deaconu, M., Lejay, A., Meira, J., Nichil, G., and State, R. (2022). Hawkes processes framework with a gamma density as excitation function: application to natural disasters for insurance. *Methodology and Computing in Applied Probability*, 24:2509–2537.

[Liniger, 2009] Liniger, T. (2009). *Multivariate Hawkes processes*. PhD thesis, ETH Zurich.

[Morariu-Patrichi and Pakkanen, 2022] Morariu-Patrichi, M. and Pakkanen, M. S. (2022). State-dependent Hawkes processes and their application to limit order book modelling. *Quantitative Finance*, 22(3):563–583.

[Nagin and Pogarsky, 2004] Nagin, D. S. and Pogarsky, G. (2004). Time and punishment: Delayed consequences and criminal behavior. *Journal of Quantitative Criminology*, 20:295–317.

[Nolet and Dahlen, 2000] Nolet, G. and Dahlen, F. (2000). Wave front healing and the evolution of seismic delay times. *Journal of Geophysical Research: Solid Earth*, 105(B8):19043–19054.

[Ogata, 1978] Ogata, Y. (1978). The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics*, 30(1):243–261.

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

[Ozaki, 1979] Ozaki, T. (1979). Maximum likelihood estimation of Hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, 31(1):145–155.

[Pagnotta and Philippon, 2018] Pagnotta, E. S. and Philippon, T. (2018). Competing on speed. *Econometrica*, 86(3):1067–1115.

[Potiron, 2025] Potiron, Y. (2025). Inference for Hawkes processes with a general kernel. *Working paper available at https://www.fbc.keio.ac.jp/∼ potiron/Potiron2025inferenceworkingpaper.pdf*.

[Potiron et al., 2025a] Potiron, Y., Scaillet, O., Volkov, V., and Yu, S. (2025a). Estimation of branching ratio for Hawkes processes with Itô semimartingale baseline. *Working paper available at https://www.fbc.keio.ac.jp/∼ potiron/Potiron2025estimationworkingpaper.pdf*.

[Potiron et al., 2025b] Potiron, Y., Scaillet, O., Volkov, V., and Yu, S. (2025b). High-frequency estimation of Itô semimartingale baseline for Hawkes processes. *Working paper available at https://www.fbc.keio.ac.jp/∼ potiron/Potiron2025workingpaper.pdf*.

[Potiron and Volkov, 2025] Potiron, Y. and Volkov, V. (2025+). Mutually exciting point processes with latency. *To appear in Journal of the American Statistical Association*.

[Rubin, 1972] Rubin, I. (1972). Regular point processes and their detection. *IEEE Transactions on Information Theory*, 18:547–557.

[Schoenberg, 2023] Schoenberg, F. (2023). Estimating Covid-19 transmission time using Hawkes point processes. *The Annals of Applied Statistics*, 17(4):3349–3362.

[Vere-Jones, 1978] Vere-Jones, D. (1978). Earthquake prediction-a statistician's view. *Journal of Physics of the Earth*, 26:129–146.

[Vere-Jones and Ozaki, 1982] Vere-Jones, D. and Ozaki, T. (1982). Some examples of statistical estimation applied to earthquake data: I. cyclic poisson and self-exciting models. *Annals of the Institute of Statistical Mathematics*, 34:189–207.

[Yu, 2004] Yu, J. (2004). Empirical characteristic function estimation and its applications. *Econometric Reviews*, 23(2):93–123.

# Supplementary materials

This part corresponds to the supplementary materials of "Estimation of latency for Hawkes processes with a polynomial periodic kernel" by Deniz Erdemlioglu, Yoann Potiron, Vladimir Volkov and Taiyu Xu submitted to the Journal of the American Statistical Association. All the proofs of the theory can be found in Section 6.

## 6 Proofs

We first introduce some notations that we will be using throughout this section. We use $C$ for any constant, and the value of the constant can change from one line to the next. Any operation with two vectors of the same size means the operation component by component. When $Y$ and $Z$ are two sequences of real numbers, we define the notation big tau as $Y = O(Z)$, which means that $\frac{Y}{Z}\mathbf{1}_{\{Z\neq 0\}}$ is bounded. Finally, any convergence refers to the convergence when $n \to \infty$ and the final time $T$ is finite.

Our proof strategy follows the general machinery of [Clinet and Yoshida, 2017] and [Potiron, 2025], which consider asymptotics where $T \to \infty$. To rewrite our problem with in-fill asymptotics as a problem with large-T asymptotics, we consider a time change as in [Clinet and Potiron, 2018], [Kwan et al., 2023], [Potiron and Volkov, 2025] and [Erdemlioglu et al., 2025]. Namely, we define the time-changed filtration as $(\overline{\mathcal{F}}_t)_{t\in[0,nT]}$, where $\overline{\mathcal{F}}_t = \mathcal{F}_{\frac{t}{n}}$. For any $i = 1, \cdots, d$ the $i$th process of the time-changed point process $\overline{N}_t^{(i)}$ has events at times $(\overline{\tau}_1^{(i)}, \cdots, \overline{\tau}_{N^{(i)}}^{(i)})$, defined as $\overline{\tau}_k^{(i)} = n\tau_k^{(i)}$ for $k = 1, \cdots, N^{(i)}$. We also define the rescaled time-changed $\overline{\mathcal{F}}_t$-intensity process as $\overline{\lambda}_t = \frac{\lambda_{t/n}}{n}$ for $t \in [0, nT]$.

In this first lemma, we rewrite the rescaled time-changed intensity in terms of the time-changed point process. This corresponds exactly to Lemma C1 in Supplement C from [Potiron and Volkov, 2025].

**Lemma 1.** *We have that*

$$\overline{\lambda}_t = \nu^* + \int_0^t h(t-s, \theta^*) d\overline{N}_s. \tag{28}$$

*Proof of Lemma 1.* If we substitute the definitions of $\overline{\lambda}_t$ and $\overline{N}_t$ into Definition (2), we obtain the lemma. □

We define the compensated rescaled time-changed point process as

$$\overline{M}_t = \overline{N}_t - \int_0^t \overline{\lambda}_s ds. \tag{29}$$

The next lemma shows that the rescaled time-changed point process $\overline{M}_t$ is an $\overline{\mathcal{F}}_t$-local martingale and that the time-changed point process $\overline{N}_t$ is a parametric Hawkes process with the same kernel $h$, baseline parameter $\nu^*$ and kernel parameter $\theta^*$. This corresponds to Lemma C2 in [Potiron and Volkov, 2025].

**Lemma 2.** *We have that the rescaled time-changed point process $\overline{M}_t$ is an $\overline{\mathcal{F}}_t$-local martingale with intensity $\overline{\lambda}_t$. Moreover, the time-changed point process $\overline{N}_t$ is a parametric Hawkes process with the same kernel $h$, baseline parameter $\nu^*$ and kernel parameter $\theta^*$.*

*Proof of Lemma 2.* By definition of a compensator, we have that

$$M_t = N_t - \int_0^t \lambda_s ds \tag{30}$$

is an $\mathcal{F}_t$-local martingale. First, we will show that $\overline{M}_t$ is an $\overline{\mathcal{F}}_t$-local martingale. We have

$$\begin{aligned}
\overline{M}_t &= \overline{N}_t - \int_0^t \overline{\lambda}_s ds \\
&= N_{t/n} - \int_0^t \frac{\lambda_{s/n}}{n} ds \\
&= N_{t/n} - \int_0^{\frac{t}{n}} \lambda_y dy \\
&= M_{t/n}.
\end{aligned}$$

Here, we use Equation (29) in the first equality, the definitions of $\overline{N}_t$ and $\overline{\lambda}_t$ in the second equality, integral change of variable in the third equality and Equation (30) in the fourth equality. As $M_t$ is an $\mathcal{F}_t$-local martingale, we also have that the time-changed local martingale $M_{t/n}$ is an $\overline{\mathcal{F}}_t$-local martingale. Since $\overline{M}_t = M_{t/n}$, it means that $\overline{M}_t$ is an $\overline{\mathcal{F}}_t$-local martingale. Then, we can deduce that $\overline{N}_t$ is a parametric Hawkes process with the same kernel $h$, parameters $\xi^*$, and $\overline{\mathcal{F}}_t$-intensity $\overline{\lambda}_t$ by Theorem 3.17 (p. 32) in [Jacod and Shiryaev, 2003]. □

For any parameter $\xi \in \Xi$ and any time $t \in [0, nT]$, we define the rescaled time-changed intensity at the parameter value $\xi$ as $\overline{\lambda}_t(\xi) = \frac{\lambda_{t/n}(\xi)}{n}$. We also define the log likelihood process of the rescaled time-changed point process $\overline{N}$ on the time interval $[0, nT]$ as

$$\bar{l}(\xi) = \sum_{i=1}^{d} \int_0^{nT} \log(\overline{\lambda}_t^{(i)}(\xi)) d\overline{N}_t^{(i)} - \sum_{i=1}^{d} \int_0^{nT} \overline{\lambda}_t^{(i)}(\xi) dt. \tag{31}$$

Then, the maximum likelihood estimator is defined as the maximizing parameter of the log likelihood process, i.e.

$$\widehat{\widehat{\xi}} \in \operatorname{argmax}_{\xi \in \Xi} \bar{l}(\xi).$$

Here, we have that the estimator $\widehat{\widehat{\xi}}$ has the form $\widehat{\widehat{\xi}} = (\widehat{\widehat{\nu}}, \widehat{\widehat{\theta}})$, where $\widehat{\widehat{\nu}} \in \Phi$ and $\widehat{\widehat{\theta}} \in \Theta$.

The following lemma states that a.s. the maximum likelihood estimator of the point process on the interval $[0, T)$ is equal to the maximum likelihood estimator of the rescaled time-changed point process on the interval $[0, nT)$. This corresponds to Lemma C3 in Supplement C from [Potiron and Volkov, 2025].

**Lemma 3.** *We have that*

$$\mathbb{P}(\widehat{\xi} = \widehat{\widehat{\xi}}) = 1.$$

*Proof of Lemma 3.* By Definition (31), the definition of $\widehat{\widehat{\xi}}$ and Lemma 2, the lemma follows. □

In what follows, we verify that Assumption 2 from [Potiron, 2025] is satisfied. In the following lemma, we show that Assumption 2(a) from [Potiron, 2025] holds.

**Lemma 4.** *Under Condition 1 (a), we have that the parameter space $\Xi \subset \mathbb{R}^m$ is such that its closure $\overline{\Xi}$ is a compact space.*

*Proof of Lemma 4.* This can be obtained by Condition 1 (a) with Lemma 2. □

In the following lemma, we consider Assumption 2(b) from [Potiron, 2025].

**Lemma 5.** *Under Condition 1 (b), there exists a positive constant $\nu_- > 0$ such that for any baseline parameter $\nu \in \Phi$ and any index $i = 1, \cdots, d$ we have $\nu^{(i)} > \nu_-$.*

*Proof of Lemma 5.* This can be deduced from Condition 1 (b) with Lemma 2. □

We introduce now a lemma which proves Assumption 2(c) from [Potiron, 2025]. This lemma is about the positivity of the kernel $h$.

**Lemma 6.** *We assume that Conditions 1 (c) and (d) hold. Then, we have that the kernel is positive, i.e. $h(t,\theta) > 0$, for any kernel parameter $\theta \in \Theta$ and any time $t \in \mathbb{R}^+$.*

*Proof of Lemma 6.* From Equation (17) and Condition 1 (c), we get for any line index $i = 1, \cdots, d$ and any column index $j = 1, \cdots, d$ that

$$h^{(i,j)}(t, \theta^{(i,j)}) = \gamma^{(i,j)}(1 + A^{(i,j)} \cos(\pi^{(i,j)} t)) \frac{\beta^{(i,j)} t^{\beta^{(i,j)}-1}}{(\alpha^{(i,j)})^{\beta^{(i,j)}} (1 + (t/\alpha^{(i,j)})^{\beta^{(i,j)}})^2}.$$

Finally, we can deduce that $h(t,\theta) > 0$ from its above expression and Condition 1 (d). □

Before we turn to the case of Assumption 2(d) from [Potiron, 2025], we deliver a couple of lemmas. For that purpose, we need some more notation. For any line index $i = 1, \cdots, d$ and any column index $j = 1, \cdots, d$, we define the component $(i,j)$ of the log-logistic kernel as

$$\widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)}) = \gamma^{(i,j)} \frac{\beta^{(i,j)} t^{\beta^{(i,j)}-1}}{(\alpha^{(i,j)})^{\beta^{(i,j)}} (1 + (t/\alpha^{(i,j)})^{\beta^{(i,j)}})^2}. \tag{32}$$

Here, we assume that the kernel parameter $\widetilde{\theta}$ is of the form

$$\widetilde{\theta} = (\widetilde{\theta}^{(i,j)})_{1 \leq i,j \leq d} = (\widetilde{\theta}^{(1,1)}, \widetilde{\theta}^{(1,2)}, \cdots, \widetilde{\theta}^{(d,d-1)}, \widetilde{\theta}^{(d,d)}), \tag{33}$$
$$\widetilde{\theta}^{(i,j)} = (\gamma^{(i,j)}, \alpha^{(i,j)}, \beta^{(i,j)}).$$

We denote the parameter space of $\widetilde{\theta}$ by $\widetilde{\Theta}$. The next lemma shows that the log-logistic kernel is continuously differentiable with respect to its parameter and gives its partial derivatives.

**Lemma 7.** *For any line index $i = 1, \cdots, d$, any column index $j = 1, \cdots, d$ and any time $t \geq 0$, we have that the log-logistic kernel $\widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)})$ is continuously differentiable with respect to its parameter*

$\widetilde{\theta}^{(i,j)}$ *and its partial derivatives are equal to*

$$\frac{\partial \widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)})}{\partial \gamma^{(i,j)}} = \frac{\beta^{(i,j)} t^{\beta^{(i,j)}-1}}{(\alpha^{(i,j)})^{\beta^{(i,j)}}(1 + (t/\alpha^{(i,j)})^{\beta^{(i,j)}})^2}, \tag{34}$$

$$\frac{\partial \widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)})}{\partial \alpha^{(i,j)}} = \gamma^{(i,j)} \frac{(\beta^{(i,j)})^2 t^{\beta^{(i,j)}-1} \left( \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} - 1 \right) (\alpha^{(i,j)})^{-\beta^{(i,j)}-1}}{\left( \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} + 1 \right)^3}, \tag{35}$$

$$\begin{aligned} \frac{\partial \widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)})}{\partial \beta^{(i,j)}} = \gamma^{(i,j)} \Bigg( &- \frac{t^{\beta^{(i,j)}-1} \left( \left( \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} \left( 2 \ln \left( \frac{t}{\alpha^{(i,j)}} \right) - \ln(t) + \ln \left( \alpha^{(i,j)} \right) \right) \right) \beta^{(i,j)} \right)}{(\alpha^{(i,j)})^{\beta^{(i,j)}} \left( \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} + 1 \right)^3} \\ &- \frac{t^{\beta^{(i,j)}-1} \left( \left( -\ln(t) + \ln \left( \alpha^{(i,j)} \right) \right) \beta^{(i,j)} - \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} - 1 \right)}{(\alpha^{(i,j)})^{\beta^{(i,j)}} \left( \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} + 1 \right)^3} \Bigg). \end{aligned} \tag{36}$$

*Proof of Lemma 7.* By differentiating Definition (32) with respect to $\gamma^{(i,j)}$, we get Equation (34) for any line index $i = 1, \cdots, d$, any column index $j = 1, \cdots, d$ and any time $t \geq 0$. Then, we obtain by differentiating Definition (32) with respect to $\alpha^{(i,j)}$, for any line index $i = 1, \cdots, d$, any column index $j = 1, \cdots, d$ and any time $t \geq 0$ that

$$\begin{aligned} \frac{\partial \widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)})}{\partial \alpha^{(i,j)}} = \gamma^{(i,j)} \Bigg( &\frac{2(\beta^{(i,j)})^2 t^{\beta^{(i,j)}-1} \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} (\alpha^{(i,j)})^{-\beta^{(i,j)}-1}}{\left( \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} + 1 \right)^3} \\ &- \frac{(\beta^{(i,j)})^2 t^{\beta^{(i,j)}-1} (\alpha^{(i,j)})^{-\beta^{(i,j)}-1}}{\left( \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} + 1 \right)^2} \Bigg). \end{aligned}$$

After some algebraic manipulation, this can be reexpressed as Equation (35). Finally, we can deduce by differentiating Definition (32) with respect to $\beta^{(i,j)}$, for any line index $i = 1, \cdots, d$, any column index $j = 1, \cdots, d$ and any time $t \geq 0$, that

$$\begin{aligned} \frac{\partial \widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)})}{\partial \beta^{(i,j)}} = \gamma^{(i,j)} \Bigg( &\frac{\ln \left( \alpha^{(i,j)} \right) t^{\beta^{(i,j)}-1} \beta^{(i,j)}}{(\alpha^{(i,j)})^{\beta^{(i,j)}} \left( \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} + 1 \right)^2} + \frac{t^{\beta^{(i,j)}-1} \ln(t) \beta^{(i,j)}}{(\alpha^{(i,j)})^{\beta^{(i,j)}} \left( \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} + 1 \right)^2} \\ &- \frac{2 t^{\beta^{(i,j)}-1} \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} \ln \left( \frac{t}{\alpha^{(i,j)}} \right) \beta^{(i,j)}}{(\alpha^{(i,j)})^{\beta^{(i,j)}} \left( \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} + 1 \right)^3} + \frac{t^{\beta^{(i,j)}-1}}{(\alpha^{(i,j)})^{\beta^{(i,j)}} \left( \left( \frac{t}{\alpha^{(i,j)}} \right)^{\beta^{(i,j)}} + 1 \right)^2} \Bigg). \end{aligned}$$

After some algebraic manipulation, this can be reexpressed as Equation (36). □

The next lemma shows that the integral for the product of the log-logistic kernel with its partial derivatives is finite uniformly in the parameter value $\widetilde{\theta} \in \widetilde{\Theta}$.

**Lemma 8.** *We assume that Conditions 1 (a), (c), (e) and (f) hold. For any line index $i = 1, \cdots, d$ and any column index $j = 1, \cdots, d$, we have that the integral for the product of the log-logistic kernel with its partial derivatives is finite uniformly in the parameter value $\widetilde{\theta} \in \widetilde{\Theta}$, i.e.*

$$\sup_{\widetilde{\theta} \in \widetilde{\Theta}} \left| \int_0^{\infty} \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)})}{\partial \gamma^{(i,j)}} ds \right| < +\infty, \tag{37}$$

$$\sup_{\widetilde{\theta} \in \widetilde{\Theta}} \left| \int_0^{\infty} \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)})}{\partial \alpha^{(i,j)}} ds \right| < +\infty, \tag{38}$$

$$\sup_{\widetilde{\theta} \in \widetilde{\Theta}} \left| \int_0^{\infty} \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)})}{\partial \beta^{(i,j)}} ds \right| < +\infty. \tag{39}$$

*Proof of Lemma 8.* To prove Expression (37), supremum properties give for any line index $i = 1, \cdots, d$ and any column index $j = 1, \cdots, d$ that

$$\sup_{\widetilde{\theta} \in \widetilde{\Theta}} \left| \int_0^{\infty} \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)})}{\partial \gamma^{(i,j)}} ds \right| \leq \int_0^{\infty} \sup_{\widetilde{\theta} \in \widetilde{\Theta}} \left| \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)})}{\partial \gamma^{(i,j)}} \right| ds. \tag{40}$$

Moreover, we get by Definition (32), Equation (34) from Lemma 7, Conditions 1 (a) and (e) that

$$\sup_{\widetilde{\theta} \in \widetilde{\Theta}} \left| \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)})}{\partial \gamma^{(i,j)}} \right| = O\left( \sup_{\widetilde{\theta} \in \widetilde{\Theta}} t^{-2(\beta^{(i,j)}+1)} \right).$$

Then, we can deduce by Condition 1 (f) that

$$\sup_{\widetilde{\theta} \in \widetilde{\Theta}} \left| \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)})}{\partial \gamma^{(i,j)}} \right| = O((t^{-2(\beta_{-}+1)})). \tag{41}$$

Finally, we obtain Expression (37) by Expressions (40), (41) and the criteria for finiteness of integrals.

To prove Expression (38), supremum properties for any line index $i = 1, \cdots, d$ and any column index $j = 1, \cdots, d$ yield

$$\sup_{\widetilde{\theta} \in \widetilde{\Theta}} \left| \int_0^{\infty} \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)})}{\partial \alpha^{(i,j)}} ds \right| \leq \int_0^{\infty} \sup_{\widetilde{\theta} \in \widetilde{\Theta}} \left| \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)})}{\partial \alpha^{(i,j)}} \right| ds. \tag{42}$$

Moreover, we get by Definition (32), Equation (35) from Lemma 7, Conditions 1 (a) and (e) that

$$\sup_{\widetilde{\theta} \in \widetilde{\Theta}} \left| \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)})}{\partial \alpha^{(i,j)}} \right| = O\left( \sup_{\widetilde{\theta} \in \widetilde{\Theta}} t^{-2(\beta^{(i,j)}+1)} \right).$$

Then, we can deduce by Condition 1 (f) that

$$\sup_{\widetilde{\theta}\in\widetilde{\Theta}} \left| \widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(t,\widetilde{\theta}^{(i,j)})}{\partial \alpha^{(i,j)}} \right| = O(t^{-2(\beta_- +1)}). \tag{43}$$

Finally, we obtain Expression (38) by Expressions (42), (43) and the criteria for finiteness of integrals.

To prove Expression (39), supremum properties for any line index $i = 1, \cdots, d$ and any column index $j = 1, \cdots, d$ deliver

$$\sup_{\widetilde{\theta}\in\widetilde{\Theta}} \left| \int_0^\infty \widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)})}{\partial \beta^{(i,j)}} ds \right| \leq \int_0^\infty \sup_{\widetilde{\theta}\in\widetilde{\Theta}} \left| \widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)})}{\partial \beta^{(i,j)}} \right| ds. \tag{44}$$

Moreover, we obtain by Definition (32), Equation (36) from Lemma 7, Conditions 1 (a) and (e) that

$$\sup_{\widetilde{\theta}\in\widetilde{\Theta}} \left| \widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(t,\widetilde{\theta}^{(i,j)})}{\partial \beta^{(i,j)}} \right| = O\left( \sup_{\widetilde{\theta}\in\widetilde{\Theta}} t^{-2(\beta^{(i,j)}+1)} \ln(t) \right).$$

Then, we can deduce by Condition 1 (f) that

$$\sup_{\widetilde{\theta}\in\widetilde{\Theta}} \left| \widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}) \frac{\partial \widetilde{h}^{(i,j)}(t,\widetilde{\theta}^{(i,j)})}{\partial \beta^{(i,j)}} \right| = O(t^{-2(\beta_- +1)} \ln(t)). \tag{45}$$

Finally, we get Expression (39) by Expressions (42), (43) and the criteria for finiteness of integrals. □

For any time $t \in \mathbb{R}^+$, we denote by $\theta_{t,2}^+$ the maximum argument parameter of the squared kernel spectral radius $\rho(h^2(t,\theta))$. It is defined through

$$h^2(t,\theta_{t,2}^+) = \sup_{\theta\in\Theta} \rho(h^2(t,\theta)). \tag{46}$$

Then, we define the matrix $\phi_2$ of dimension $d \times d$ as the integral of $h^2(t,\theta_{t,2}^+)$ over time, i.e.

$$\phi_2 = \int_0^\infty h^2(s,\theta_{s,2}^+) ds. \tag{47}$$

We deliver in the following a lemma which shows Assumption 2(d) from [Potiron, 2025]. More specifically, we have the spectral radius of the matrix $\phi$ is smaller than unity, i.e. $\rho(\phi) < 1$, and the spectral radius of the matrix $\phi_2$ is finite, i.e. $\rho(\phi_2) < +\infty$.

**Lemma 9.** *We assume that Conditions 1 (a), (c), (d), (e), (f) and (g) hold. Then, we have the spectral radius of the matrix $\phi$ is smaller than unity, i.e. $\rho(\phi) < 1$, and the spectral radius of the matrix $\phi_2$ is finite, i.e. $\rho(\phi_2) < +\infty$.*

*Proof of Lemma 9.* First, we have $\rho(\phi) < 1$ by Condition 1 (g). Second, we prove in what follows that $\rho(\phi_2) < +\infty$. By Definition (17) and Condition 1 (d), it is sufficient to prove that

$$\rho(\widetilde{\phi}_2) < +\infty. \tag{48}$$

Here, we denote for any $t \in \mathbb{R}^+$ by $\widetilde{\theta}^+_{t,2}$ the maximum argument parameter of the squared kernel spectral radius $\rho(\widetilde{h}^2(t, \widetilde{\theta}))$. It is defined through

$$\widetilde{h}^2(t, \widetilde{\theta}^+_{t,2}) = \sup_{\widetilde{\theta} \in \widetilde{\Theta}} \rho(\widetilde{h}^2(t, \widetilde{\theta})). \tag{49}$$

Then, we define the matrix $\widetilde{\phi}_2$ of dimension $d \times d$ as the integral of $\widetilde{h}^2(t, \widetilde{\theta}^+_{t,2})$ over time, i.e.

$$\widetilde{\phi}_2 = \int_0^\infty \widetilde{h}^2(s, \widetilde{\theta}^+_{s,2}) ds. \tag{50}$$

By Definition (50), we have that

$$\rho(\widetilde{\phi}_2) = \rho\Big( \int_0^\infty \widetilde{h}^2(s, \widetilde{\theta}^+_{s,2}) ds \Big). \tag{51}$$

We can rewrite $\rho(\widetilde{\phi}_2)$ as

$$\rho(\widetilde{\phi}_2) = \rho\Big( \int_0^\infty \widetilde{h}^2(s, \widetilde{\theta}) ds \Big) + \rho\Big( \int_0^\infty \widetilde{h}^2(s, \widetilde{\theta}^+_{s,2}) ds \Big) - \rho\Big( \int_0^\infty \widetilde{h}^2(s, \widetilde{\theta}) ds \Big), \tag{52}$$

where $\widetilde{\theta} \in \widetilde{\Theta}$.

To show that $\rho\big( \int_0^\infty \widetilde{h}^2(s, \widetilde{\theta}) ds \big) < +\infty$, it is sufficient to prove for any line index $i = 1, \cdots, d$ and any column index $j = 1, \cdots, d$ that $\int_0^\infty \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)})^2 ds < +\infty$. Definition (32) yields

$$\int_0^\infty \widetilde{h}^{(i,j)}(s, \widetilde{\theta}^{(i,j)})^2 ds = \int_0^\infty (\gamma^{(i,j)})^2 \frac{(\beta^{(i,j)})^2 s^{2(\beta^{(i,j)}-1)}}{(\alpha^{(i,j)})^{2\beta^{(i,j)}} (1 + (t/\alpha^{(i,j)})^{\beta^{(i,j)}})^4} ds. \tag{53}$$

As $s \to +\infty$, we have by Condition 1 (e) that

$$(\gamma^{(i,j)})^2 \frac{(\beta^{(i,j)})^2 s^{2(\beta^{(i,j)}-1)}}{(\alpha^{(i,j)})^{2\beta^{(i,j)}} (1 + (t/\alpha^{(i,j)})^{\beta^{(i,j)}})^4} = O(s^{-2\beta^{(i,j)}-2}). \tag{54}$$

Then, the criteria for finiteness of integrals with Condition 1 (f) yield

$$\int_0^\infty (\gamma^{(i,j)})^2 \frac{(\beta^{(i,j)})^2 s^{2(\beta^{(i,j)}-1)}}{(\alpha^{(i,j)})^{2\beta^{(i,j)}} (1 + (t/\alpha^{(i,j)})^{\beta^{(i,j)}})^4} ds < +\infty. \tag{55}$$

By Equation (53), we can deduce that

$$\int_0^\infty \widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)})^2 ds < +\infty. \tag{56}$$

Thus, we get

$$\rho\bigg(\int_0^\infty \widetilde{h}^2(s,\widetilde{\theta})ds\bigg) < +\infty. \tag{57}$$

To show that $\rho\big(\int_0^\infty \widetilde{h}^2(s,\widetilde{\theta}^+_{s,2})ds\big) - \rho\big(\int_0^\infty \widetilde{h}^2(s,\widetilde{\theta})ds\big) < +\infty$, it is sufficient to prove for any line index $i = 1,\cdots,d$ and any column index $j = 1,\cdots,d$ that

$$\int_0^\infty \big(\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{+,(i,j)}_{s,2})^2 - \widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)})^2\big)ds < +\infty.$$

By an application of Lemma 7, we have that $\widetilde{h}^{(i,j)}(t,\widetilde{\theta}^{(i,j)})$ is continuously differentiable with respect to its parameter $\widetilde{\theta}^{(i,j)}$ for any line index $i = 1,\cdots,d$ and any column index $j = 1,\cdots,d$. Thus, we can use a Taylor expansion

$$\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{+,(i,j)}_{s,2})^2 - \widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)})^2 = \big(\widetilde{\theta}^{+,(i,j)}_{s,2} - \widetilde{\theta}^{(i,j)}\big)\nabla\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}_e)^2. \tag{58}$$

Here, $\nabla F$ is the gradient of a function $F$ and $\widetilde{\theta}^{(i,j)}_e$ is between $\widetilde{\theta}^{(i,j)}$ and $\widetilde{\theta}^{+,(i,j)}_{s,2}$. The right side of Equation (58) can be rewritten as

$$\begin{aligned}\big(\widetilde{\theta}^{+,(i,j)}_{s,2} - \widetilde{\theta}^{(i,j)}\big)\nabla\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}_e)^2 = {} & 2\big(\gamma^{+,(i,j)}_{s,2} - \gamma^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}_e)\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}_e)}{\partial\gamma^{(i,j)}} \\ & + 2\big(\alpha^{+,(i,j)}_{s,2} - \alpha^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}_e)\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}_e)}{\partial\alpha^{(i,j)}} \\ & + 2\big(\beta^{+,(i,j)}_{s,2} - \beta^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}_e)\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}_e)}{\partial\beta^{(i,j)}}.\end{aligned} \tag{59}$$

For the first term on the right side of Equation (59), we have by supremum properties that

$$\begin{aligned}&\int_0^\infty 2\big(\gamma^{+,(i,j)}_{s,2} - \gamma^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}_e)\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}_e)}{\partial\gamma^{(i,j)}}ds \\ &\leq \sup_{\widetilde{\theta}_+\in\widetilde{\Theta}}\bigg|\int_0^\infty 2\big(\gamma^{(i,j)}_+ - \gamma^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}_+)\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)}_+)}{\partial\gamma^{(i,j)}}ds\bigg|.\end{aligned} \tag{60}$$

By Condition 1 (a), we can deduce that

$$\sup_{\widetilde{\theta}_+\in\widetilde{\Theta}}\left|\int_0^\infty 2\big(\gamma_+^{(i,j)}-\gamma^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})}{\partial\gamma^{(i,j)}}ds\right| \le C\sup_{\widetilde{\theta}_+\in\widetilde{\Theta}}\left|\int_0^\infty \widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})}{\partial\gamma^{(i,j)}}ds\right|. \tag{61}$$

By Expression (37) from Lemma 8, we obtain

$$\sup_{\widetilde{\theta}_+\in\widetilde{\Theta}}\left|\int_0^\infty \widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})}{\partial\gamma^{(i,j)}}ds\right| < +\infty. \tag{62}$$

By Expressions (60), (61) and (62), we get

$$\int_0^\infty 2\big(\gamma_{s,2}^{+,(i,j)}-\gamma^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_e^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_e^{(i,j)})}{\partial\gamma^{(i,j)}}ds < +\infty. \tag{63}$$

For the second term on the right side of Equation (59), we have by supremum properties that

$$\int_0^\infty 2\big(\alpha_{s,2}^{+,(i,j)}-\alpha^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_e^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_e^{(i,j)})}{\partial\alpha^{(i,j)}}ds \le \sup_{\widetilde{\theta}_+\in\widetilde{\Theta}}\left|\int_0^\infty 2\big(\alpha_+^{(i,j)}-\alpha^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})}{\partial\alpha^{(i,j)}}ds\right|. \tag{64}$$

By Condition 1 (a), we can deduce that

$$\sup_{\widetilde{\theta}_+\in\widetilde{\Theta}}\left|\int_0^\infty 2\big(\alpha_+^{(i,j)}-\alpha^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})}{\partial\alpha^{(i,j)}}ds\right| \le C\sup_{\widetilde{\theta}_+\in\widetilde{\Theta}}\left|\int_0^\infty \widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})}{\partial\alpha^{(i,j)}}ds\right|. \tag{65}$$

By Expression (38) from Lemma 8, we obtain

$$\sup_{\widetilde{\theta}_+\in\widetilde{\Theta}}\left|\int_0^\infty \widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})}{\partial\alpha^{(i,j)}}ds\right| < +\infty. \tag{66}$$

By Expressions (64), (65) and (66), we get

$$\int_0^\infty 2\big(\alpha_{s,2}^{+,(i,j)}-\alpha^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_e^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_e^{(i,j)})}{\partial\alpha^{(i,j)}}ds < +\infty. \tag{67}$$

For the third term on the right side of Equation (59), we have by supremum properties that

$$\int_0^\infty 2\big(\beta_{s,2}^{+,(i,j)}-\beta^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_e^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_e^{(i,j)})}{\partial\beta^{(i,j)}}ds \le \sup_{\widetilde{\theta}_+\in\widetilde{\Theta}}\left|\int_0^\infty 2\big(\beta_+^{(i,j)}-\beta^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})}{\partial\beta^{(i,j)}}ds\right|. \tag{68}$$

By Condition 1 (a), we can deduce that

$$\sup_{\widetilde{\theta}_+\in\widetilde{\Theta}}\Big|\int_0^\infty 2\big(\beta_+^{(i,j)}-\beta^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})}{\partial\beta^{(i,j)}}ds\Big| \\ \leq C\sup_{\widetilde{\theta}_+\in\widetilde{\Theta}}\Big|\int_0^\infty \widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})}{\partial\beta^{(i,j)}}ds\Big|. \tag{69}$$

By Expression (39) from Lemma 8, we obtain

$$\sup_{\widetilde{\theta}_+\in\widetilde{\Theta}}\Big|\int_0^\infty \widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_+^{(i,j)})}{\partial\beta^{(i,j)}}ds\Big| < +\infty. \tag{70}$$

By Expressions (68), (69) and (70), we get

$$\int_0^\infty 2\big(\beta_{s,2}^{+,(i,j)}-\beta^{(i,j)}\big)\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_e^{(i,j)})\frac{\partial\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_e^{(i,j)})}{\partial\beta^{(i,j)}}ds < +\infty. \tag{71}$$

Moreover, Expressions (58), (59), (63), (67) and (71) yield

$$\int_0^\infty \big(\widetilde{h}^{(i,j)}(s,\widetilde{\theta}_{s,2}^{+,(i,j)})^2-\widetilde{h}^{(i,j)}(s,\widetilde{\theta}^{(i,j)})^2\big)ds < +\infty. \tag{72}$$

Thus, we get

$$\rho\Big(\int_0^\infty \widetilde{h}^2(s,\widetilde{\theta}_{s,2}^{+})ds\big)-\rho\big(\int_0^\infty \widetilde{h}^2(s,\widetilde{\theta})ds\Big) < +\infty. \tag{73}$$

Finally, Expressions (52), (57) and (73) give $\rho(\phi_2) < +\infty$. □

We introduce a lemma which proves Assumption 2(e) from [Potiron, 2025], i.e. some smoothness assumptions on the kernel $h$.

**Lemma 10.** *We assume that Conditions 1 (a) and (c) hold. For any time $s\in\mathbb{R}^+$ a.e., we have the kernel function $\theta\to h(s,\theta)$ is continuously differentiable twice from the kernel parameter space $\Theta$ to the space $\mathbb{R}_+^{d\times d}$ and there exists a continuous extension to $\overline{\Theta}$.*

*Proof of Lemma 10.* By Condition 1 (c), it is sufficient to show that, for any time $s\in\mathbb{R}^+$ a.e., any line index $i=1,\cdots,d$ and any column index $j=1,\cdots,d$, we have $\theta^{(i,j)}\to h^{(i,j)}(s,\theta^{(i,j)})$ is continuously differentiable twice from the kernel parameter space $\Theta^{(i,j)}$ to the space $\mathbb{R}_{+}^{d\times d}$ and there exists a continuous extension to $\overline{\Theta}^{(i,j)}$. Here, $\Theta^{(i,j)}$ denotes the parameter space $\Theta$ restricted to the parameters of the component $(i, j)$. By Definition (17), we have

$$h^{(i,j)}(t, \theta^{(i,j)}) = \gamma^{(i,j)}(1 + A^{(i,j)} \cos(\pi^{(i,j)} t)) \frac{\beta^{(i,j)} t^{\beta^{(i,j)}-1}}{(\alpha^{(i,j)})^{\beta^{(i,j)}}(1 + (t/\alpha^{(i,j)})^{\beta^{(i,j)}})^2}.$$

By Definition (32), we can deduce that

$$h^{(i,j)}(t, \theta^{(i,j)}) = (1 + A^{(i,j)} \cos(\pi^{(i,j)} t))\widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)}).$$

First, we have that $h^{(i,j)}(t, \theta^{(i,j)})$ is the product of $(1 + A^{(i,j)} \cos(\pi^{(i,j)} t))$ and $\widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)})$. Secondly, we directly get that $\theta^{(i,j)} \to (1 + A^{(i,j)} \cos(\pi^{(i,j)} t))$ is continuously differentiable twice from the kernel parameter space $\Theta^{(i,j)}$ to the space $\mathbb{R}_{+}^{d\times d}$ and there exists a continuous extension to $\overline{\Theta}^{(i,j)}$ by Condition 1 (a) for any time $s \in \mathbb{R}^{+}$ a.e., any line index $i = 1, \cdots, d$ and column index $j = 1, \cdots, d$. Thirdly, we obtain by an application of Lemma 7 that $\theta^{(i,j)} \to \widetilde{h}^{(i,j)}(t, \widetilde{\theta}^{(i,j)})$ is continuously differentiable twice from the kernel parameter space $\Theta^{(i,j)}$ to the space $\mathbb{R}_{+}^{d\times d}$ and there exists a continuous extension to $\overline{\Theta}^{(i,j)}$ by Condition 1 (a) for any time $s \in \mathbb{R}^{+}$ a.e., any line index $i = 1, \cdots, d$ and any column index $j = 1, \cdots, d$. Thus, we can deduce that $\theta^{(i,j)} \to h^{(i,j)}(t, \theta^{(i,j)})$ is continuously differentiable twice from the kernel parameter space $\Theta^{(i,j)}$ to the space $\mathbb{R}_{+}^{d\times d}$ and there exists a continuous extension to $\overline{\Theta}^{(i,j)}$ for any time $s \in \mathbb{R}^{+}$ a.e., any line index $i = 1, \cdots, d$ and any column index $j = 1, \cdots, d$. □

The next lemma shows that Assumption 2(f) from [Potiron, 2025] is satisfied. Namely, we have the spectral radius of the matrix $\phi_3$ is smaller than $\phi^{+}$, i.e. $\rho(\phi_3(\theta)) < \phi^{+}$, and the spectral radius of the matrix $\phi_4$ is smaller than $\phi^{+}$, i.e. $\rho(\phi_4(\theta)) < \phi^{+}$.

**Lemma 11.** *We assume that Condition 1 (h) holds. For any kernel parameter $\theta \in \Theta$, we have the spectral radius of the matrix $\phi_3(\theta)$ is smaller than $\phi^{+}$, i.e. $\rho(\phi_3(\theta)) < \phi^{+}$, and the spectral radius of the matrix $\phi_4(\theta)$ is smaller than $\phi^{+}$, i.e. $\rho(\phi_4(\theta)) < \phi^{+}$.*

*Proof of Lemma 11.* This is exactly Condition 1 (h). □

We define the matrix $\phi_5$ of dimension $(n-d) \times (n-d)$ as the integral of $\left|\partial_\theta h^{(i,j)}(s,\theta^*)^{(k_{t,3}^{(i,j)})}\right|^2$ over time, i.e.

$$\phi_5^{(i,j)} = \int_0^\infty \left|\partial_\theta h^{(i,j)}(s,\theta^*)^{(k_{t,3}^{(i,j)})}\right|^2 ds,$$

for any line index $i = 1, \cdots, d$ and any column index $j = 1, \cdots, d$. Finally, we define the matrix $\phi_6$ of dimension $d \times d$ as the integral of $\left|\partial_\theta^2 h^{(i,j)}(s,\theta^*)^{(k_{t,4}^{(i,j)}, l_{t,4}^{(i,j)})}\right|^2$ over time, i.e.

$$\phi_6^{(i,j)} = \int_0^\infty \left|\partial_\theta^2 h^{(i,j)}(s,\theta^*)^{(k_{t,4}^{(i,j)}, l_{t,4}^{(i,j)})}\right|^2 ds,$$

for any $i = 1, \cdots, d$ and $j = 1, \cdots, d$.

Moreover, we give the following lemma proving Assumption 2(g) from [Potiron, 2025]. Namely, we have the spectral norm of the matrix $\phi_5$ is finite, i.e. $\rho(\phi_5) < +\infty$, and the spectral norm of the matrix $\phi_6$ is finite, i.e. $\rho(\phi_6) < +\infty$.

**Lemma 12.** *We assume that Conditions 1 (a), (c), (d), (e), (f) and (g) hold. Then, we have the spectral norm of the matrix $\phi_5$ is finite, i.e. $\rho(\phi_5) < +\infty$ and the spectral norm of the matrix $\phi_6$ is finite, i.e.$\rho(\phi_6) < +\infty$.*

*Proof of Lemma 12.* This can be proven by extending the arguments from the proof of Lemma 9. □

To prove Assumption 2(h) from [Potiron, 2025], we give two prior lemmas. First, the following lemma states that $X_t$ is stable. This corresponds to Lemma 13 in [Potiron, 2025]. See also Lemma A.6 (p. 1834) in [Clinet and Yoshida, 2017] and Proposition C1 (ii) in Supplement C of [Potiron and Volkov, 2025]. Its proof is based on [Brémaud and Massoulié, 1996] (see Theorem 1 and Lemma 4).

**Lemma 13.** *We assume that Conditions 1 (b), (c) and (g) hold. $X_t$ is stable for any $\xi \in \Xi$, i.e. for any $i = 1, \cdots, d$ there exists an $\mathbb{R}_+^*$-valued random variable $\overline{\lambda}_l^{(i)}(\xi)$ such that we have*

$$X_{nT}^{(i)} \xrightarrow{\mathcal{D}} \left(\overline{\lambda}_l^{(i)}(\xi^*), \overline{\lambda}_l^{(i)}(\xi), \partial_\xi \overline{\lambda}_l^{(i)}(\xi)\right).$$

*Proof of Lemma 13.* The proof is obtained by [Brémaud and Massoulié, 1996] (see Theorem 1 and Lemma 4) with Conditions 1 (b), (c) and (g). □

In what follows, we provide the definition of ergodicity. This corresponds to Definition 3.1 (p. 1805) in [Clinet and Yoshida, 2017]. See also Definition C1 in Supplement C of [Potiron and Volkov, 2025].

*Definition* 1. We say that $X$ is ergodic if for any index $i = 1, \cdots, d$ there exists a function $\pi^{(i)} : C_b(E, \mathbb{R}) \to \mathbb{R}$ such that for any function $\psi \in C_b(E, \mathbb{R})$ we have

$$\frac{1}{nT} \int_0^{nT} \psi(X_s^{(i)}) ds \xrightarrow{\mathbb{P}} \pi^{(i)}(\psi).$$

The following lemma states that $X_t$ is ergodic in the sense of Definition 1. This is a direct consequence to Lemma 14 in [Potiron, 2025]. See also Lemma 3.16 (p. 1815) in [Clinet and Yoshida, 2017] and Proposition C1 (iii) in Supplement C of [Potiron and Volkov, 2025].

**Lemma 14.** *We assume that Condition 1 holds. For any parameter $\xi \in \Xi$, the process $X_t$ is ergodic in the sense of Definition 1. Moreover, for any index $i = 1, \cdots, d$ we have*

$$\pi^{(i)}(\psi) = \mathbb{E}\big[\psi(\overline{\lambda}_l^{(i)}(\xi^*), \overline{\lambda}_l^{(i)}(\xi), \partial_\xi \overline{\lambda}_l^{(i)}(\xi))\big].$$

*Proof of Lemma 14.* This is a direct consequence to Lemma 14 in [Potiron, 2025] with Condition 1. □

Since the functions that are used in the definition of the Fisher information matrix (10) are not 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 condition holds. We also give a more explicit form to the functions $\pi(\psi)$. The following lemma is Proposition 3.8 (pp. 1806-1807) in [Clinet and Yoshida, 2017]. See also Lemma 1 in [Potiron, 2025]. The proof follows the arguments from the proof of Proposition 3.8 (pp. 1822-1824) in [Clinet and Yoshida, 2017].

**Lemma 15.** *We assume that Condition 1 holds. For any $\xi \in \Xi$, we have*

*(a) The starting space of the limit function $\pi$ in Definition 1 of ergodicity can be extended from $C_b(E, \mathbb{R})$ to $C_\uparrow(E, \mathbb{R})$.*

*(b) For any $i = 1, \cdots, d$, there exists a probability measure $\pi_\xi^{(i)}$ on $E$ such that, for any $\psi \in C_\uparrow(E, \mathbb{R})$, we have $\pi^{(i)}(\psi) = \int_E \psi(u, v, w)\pi_\xi^{(i)}(du, dv, dw)$.*

*Proof of Lemma 15.* We can use the arguments from the proof of Proposition 3.8 (pp. 1822-1824) in [Clinet and Yoshida, 2017] with Conditions 1 (a), (b), (c), (d), (e), (f), (g) and (h). □

Finally, we introduce the lemma proving Assumption 2(h) from [Potiron, 2025], which corresponds to the last condition of the statistical inference theory.

**Lemma 16.** *We assume that Condition 1 holds. Then, we have $\mathbb{P}\big(\lambda_l(\xi^*) = \lambda_l(\xi)\big) = 1$ implies that $\xi^* = \xi$.*

*Proof of Lemma 16.* We assume that $\mathbb{P}\big(\lambda_l(\xi^*) = \lambda_l(\xi)\big) = 1$. In particular, for any $k = 1, \cdots, 5d$ we get

$$\mathbb{E}[\lambda_l^k(\xi)] = \mathbb{E}[\lambda_l^k(\xi^*)]. \tag{74}$$

By properties of the Hawkes processes, we can deduce that

$$\mathbb{E}[\lambda_l(\xi)] = (I - BR(\theta))^{-1}\nu. \tag{75}$$

Here, $I$ denotes the unity matrix of dimension $d \times d$. Moreover, the branching ratio matrix $BR$ of dimension $d \times d$ is defined for any line index $i = 1, \cdots, d$ and any column index $j = 1, \cdots, d$ as

$$\mathrm{BR}^{(i,j)}(\theta) = \|h^{(i,j)}(., \theta)\|_1 = \int_0^\infty h^{(i,j)}(t, \theta)\, dt.$$

Then, we have by Definition (17) that

$$h^{(i,j)}(t, \theta^{(i,j)}) = \gamma^{(i,j)}(1 + A^{(i,j)} \cos(\pi^{(i,j)}t)) \frac{\beta^{(i,j)} t^{\beta^{(i,j)}-1}}{(\alpha^{(i,j)})^{\beta^{(i,j)}}(1 + (t/\alpha^{(i,j)})^{\beta^{(i,j)}})^2}.$$

When evaluating Equation (75) at the point $\xi^*$, we get

$$\mathbb{E}[\lambda_l(\xi^*)] = (I - BR(\theta^*))^{-1}\nu^*. \tag{76}$$

Moreover, we can obtain explicit formulae of $\mathbb{E}[\lambda_l(\xi)^k]$ and $\mathbb{E}[\lambda_l(\xi)^k]$ for any $k = 2, \cdots, 5d$. These explicit formulae are different from Equations (75) and (76). This leaves us with a system of $5d^2$ equations with $5d^2$ parameters. Finally, this yields $\xi^* = \xi$ by Equation (74). □

We turn our attention to the proof of Theorem 1. This is an application of Theorem 2 in [Potiron, 2025].

*Proof of Theorem 1.* The central limit theorem (22) is obtained by an application of Theorem 2 in [Potiron, 2025] with Lemmas 4, 5, 6, 9, 10, 11, 12 and 16. With the same arguments as in the proof of Theorem 1 in Supplement C of [Potiron and Volkov, 2025], we can show the consistency of the asymptotic covariance estimator, i.e.

$$\widehat{\Gamma}^{-1} \xrightarrow{\mathbb{P}} \Gamma^{-1}. \tag{77}$$

The feasible normalized central limit theorem (23) is deduced by an application of Slusky's theorem with the central limit theorem (22) and Expression (77). □

We now deliver the proof of Corollary 1. This is a consequence to Theorem 1.

*Proof of Corollary 1.* We can get the corollary by an application of Theorem 1, a Taylor expansion and Condition 2. More specifically, we can use the arguments from the proof of Proposition 2 and Corollary 3 in Supplement C of [Potiron and Volkov, 2025]. □

Moreover, we give the proof of Proposition 1.

*Proof of Proposition 1.* We can get the proposition by an application of Theorem 1, Corollary 1, and another Taylor expansion. □

Finally, we give the proof of Corollary 2.

*Proof of Corollary 2.* Under the null hypothesis $H_0$, we can show that the Wald test statistic $S$ defined in (16) converges in distribution to a chi-squared distribution with $q$ degrees of freedom using Proposition 1. Under the alternative hypothesis $H_1$, we can show that the Wald test statistic $S$ is consistent using Proposition 1. □