Supplementary material of the article *Uncovering Causality from Multivariate Hawkes Integrated Cumulants*

Massil Achab*[1], Emmanuel Bacry[1], Stéphane Gaiffas[1], Iacopo Mastromatteo[2], and Jean-François Muzy[1,3]

[1]*Centre de Mathématiques Appliquées, CNRS, Ecole Polytechnique, UMR 7641, 91128 Palaiseau, France*
[2]*Capital Fund Management, 23 rue de l'Université, 75007 Paris, France*
[3]*Laboratoire Sciences Pour l'Environnement, Université de Corse, 7 Avenue Jean Nicoli, 20250 Corte, France*

1 Introduction

1.1 In a nutshell

We prove here the consistency of NPHC estimator using the framework of Generalized Method of Moments Hansen [1982]. The main difference with the usual Generalized Method of Moments relies in the relaxation of the moment conditions, since we have $\mathbb{E}[\widehat{g}_T(\theta_0)] = m_T \neq 0$. We adapt the proof of consistency given in Newey and McFadden [1994].

1.2 Sketch of the proof

We can relate the integral of the Hawkes process's kernels to the integrals of the cumulant densities, from Jovanović et al. [2015]. Our cumulant matching method would fall into the usual GMM framework if we could estimate - without bias - the integral of the covariance on $\mathbb{R}$, and the integral of the skewness on $\mathbb{R}^2$. Unfortunately, we can't do that easily. We can however estimate without bias $\int f_t^T C_t^{ij} dt$ and $\int f_t^T K_t^{ijk} dt$ with f^T a compact supported function on $[-H_T, H_T]$ that weakly converges to 1, with $H_T \underset{T\to\infty}{\longrightarrow} \infty$. In most cases we will take $f_t^T = \mathbb{1}_{[-H_T, H_T]}(t)$.

Denoting $\widehat{C}^{ij,(T)}$ the estimator of $\int f_t^T C_t^{ij} dt$, the term $|\mathbb{E}[\widehat{C}^{ij,(T)}] - C^{ij}| = |\int f_t^T C_t^{ij} dt - C^{ij}|$ can be considered a proxy to the *distance to the classical GMM*. This distance has to go to zero to make the rest of GMM's proof work: the estimator $\widehat{C}^{ij,(T)}$ is then asymptotically unbiased towards C^{ij} when T goes to infinity.

1.3 Notations

We observe the multivariate point process $(\boldsymbol{N}_t)$ on $\mathbb{R}^+$, with Z^i the events of the i^{th} component. We will often write covariance / skewness instead of integrated covariance / skewness. In the rest of the document, we use the following notations.

Hawkes kernels' integrals $\quad \boldsymbol{G}^{\text{true}} = \int \boldsymbol{\Phi}_t dt = (\int \phi_t^{ij} dt)_{ij} = \boldsymbol{I_d} - (\boldsymbol{R}^{\text{true}})^{-1}$

Theoretical mean matrix $\quad \boldsymbol{L} = \text{diag}(\Lambda^1, \ldots, \Lambda^d)$

Theoretical covariance $\quad \boldsymbol{C} = \boldsymbol{R}^{\text{true}} \boldsymbol{L} (\boldsymbol{R}^{\text{true}})^\top$

*massil.achab@m4x.org

Theoretical skewness $\boldsymbol{K^c} = (K^{iij})_{ij} = (\boldsymbol{R}^{\text{true}})^{\odot^2}\boldsymbol{C}^\top + 2[\boldsymbol{R}^{\text{true}} \odot (\boldsymbol{C} - \boldsymbol{R}^{\text{true}}\boldsymbol{L})](\boldsymbol{R}^{\text{true}})^\top$

Filtering function $f^T \geq 0 \qquad \text{supp}(f^T) \subset [-H_T, H_T] \qquad F^T = \int f_s^T ds \qquad \widetilde{f}_t^T = f_{-t}^T$

Events sets $Z^{i,T,1} = Z^i \cap [H_T, T + H_T] \qquad Z^{j,T,2} = Z^j \cap [0, T + 2H_T]$

Estimators of the mean $\widehat{\Lambda}^i = \frac{N^i_{T+H_T} - N^i_{H_T}}{T} \qquad \widetilde{\Lambda}^j = \frac{N^j_{T+2H_T}}{T+2H_T}$

Estimator of the covariance $\widehat{C}^{ij,(T)} = \frac{1}{T}\sum_{\tau \in Z^{i,T,1}} \left(\sum_{\tau' \in Z^{j,T,2}} f_{\tau'-\tau} - \widetilde{\Lambda}^j F^T\right)$

Estimator of the skewness[1]

$$\widehat{K}^{ijk,(T)} = \frac{1}{T}\sum_{\tau \in Z^{i,T,1}} \left(\sum_{\tau' \in Z^{j,T,2}} f_{\tau'-\tau} - \widetilde{\Lambda}^j F^T\right)\left(\sum_{\tau'' \in Z^{k,T,2}} f_{\tau'-\tau} - \widetilde{\Lambda}^k F^T\right)$$
$$- \frac{\widehat{\Lambda}^i}{T + 2H_T}\sum_{\tau' \in Z^{j,T,2}} \left(\sum_{\tau'' \in Z^{k,T,2}} (f^T \star \widetilde{f}^T)_{\tau'-\tau''} - \widetilde{\Lambda}^k (F^T)^2\right)$$

GMM related notations

$$\begin{aligned}
\theta = \boldsymbol{R} \quad &\text{and} \quad \theta_0 = \boldsymbol{R}^{\text{true}} \\
g_0(\theta) &= \textbf{vec}\begin{bmatrix} \boldsymbol{C} - \boldsymbol{R}\boldsymbol{L}\boldsymbol{R}^\top \\ \boldsymbol{K^c} - \boldsymbol{R}^{\odot^2}\boldsymbol{C}^\top - 2[\boldsymbol{R} \odot (\boldsymbol{C} - \boldsymbol{R}\boldsymbol{L})]\boldsymbol{R}^\top \end{bmatrix} \in \mathbb{R}^{2d^2} \\
\widehat{g}_T(\theta) &= \textbf{vec}\begin{bmatrix} \widehat{\boldsymbol{C}}^{(T)} - \boldsymbol{R}\widehat{\boldsymbol{L}}\boldsymbol{R}^\top \\ \widehat{\boldsymbol{K^c}}^{(T)} - \boldsymbol{R}^{\odot^2}(\widehat{\boldsymbol{C}}^{(T)})^\top - 2[\boldsymbol{R} \odot (\widehat{\boldsymbol{C}}^{(T)} - \boldsymbol{R}\widehat{\boldsymbol{L}})]\boldsymbol{R}^\top \end{bmatrix} \in \mathbb{R}^{2d^2} \\
Q_0(\theta) &= g_0(\theta)^\top W g_0(\theta) \\
\widehat{Q}_T(\theta) &= \widehat{g}_T(\theta)^\top \widehat{W}_T \widehat{g}_T(\theta)
\end{aligned}$$

2 Consistency

First, let's remind a useful theorem for consistency in GMM from Newey and McFadden [1994].

Theorem 2.1. *If there is a function $Q_0(\theta)$ such that (i) $Q_0(\theta)$ is uniquely maximized at θ_0; (ii) Θ is compact; (iii) $Q_0(\theta)$ is continuous; (iv) $\widehat{Q}_T(\theta)$ converges uniformly in probability to $Q_0(\theta)$, then $\widehat{\theta}_T = \arg\max \widehat{Q}_T(\theta) \xrightarrow{\mathbb{P}} \theta_0$.*

We can now prove the consistency of our estimator.

Theorem 2.2. *Suppose that $(\boldsymbol{N_t})$ is observed on $\mathbb{R}^+$, $\widehat{W}_T \xrightarrow{\mathbb{P}} W$, and*

1. *W is positive semi-definite and $W g_0(\theta) = 0$ if and only if $\theta = \theta_0$,*
2. *$\theta \in \Theta$, which is compact,*
3. *the spectral radius of the kernel norm matrix satisfies $||\boldsymbol{\Phi}||_* < 1$,*
4. *$\forall i, j, k \in [d]$, $\int f_u^T C_u^{ij} du \to \int C_u^{ij} du$ and $\int f_u^T f_v^T K_{u,v}^{ijk} du dv \to \int K_{u,v}^{ijk} du dv$,*

[1] When $f_t^T = \mathbb{1}_{[-H_T, H_T]}(t)$, we remind that $(f^T \star \widetilde{f}^T)_t = (2H_T - |t|)^+$. This leads to the estimator we showed in the article.

5. $(F^T)^2/T \xrightarrow{\mathbb{P}} 0$ *and* $||f||_\infty = O(1)$.

Then

$$\widehat{\theta}_T \xrightarrow{\mathbb{P}} \theta_0.$$

Remark 1. *In practice, we use a constant sequence of weighting matrices:* $\widehat{W_T} = \boldsymbol{I_d}$.

Proof. Proceed by verifying the hypotheses of Theorem 2.1 from Newey and McFadden [1994]. Condition 2.1(i) follows by (i) and by $Q_0(\theta) = [W^{1/2}g_0(\theta)]^\top[W^{1/2}g_0(\theta)] > 0 = Q_0(\theta_0)$. Indeed, there exists a neighborhood N of θ_0 such that $\theta \in N\backslash\{\theta_0\}$ and $g_0(\theta) \neq 0$ since $g_0(\theta)$ is a polynom. Condition 2.1(ii) follows by (ii). Condition 2.1(iii) is satisfied since $Q_0(\theta)$ is a polynom. Condition 2.1(iv) is harder to prove. First, since $\widehat{g}_T(\theta)$ is a polynom of θ, we prove easily that $\mathbb{E}[\sup_{\theta\in\Theta}|\widehat{g}_T(\theta)|] < \infty$. Then, by Θ compact, $g_0(\theta)$ is bounded on Θ, and by the triangle and Cauchy-Schwarz inequalities,

$$\begin{aligned}
&\left|\widehat{Q}_T(\theta) - Q_0(\theta)\right| \\
&\quad\leq \left|(\widehat{g}_T(\theta) - g_0(\theta))^\top \widehat{W}_T(\widehat{g}_T(\theta) - g_0(\theta))\right| + \left|g_0(\theta)^\top(\widehat{W}_T + \widehat{W}_T^\top)(\widehat{g}_T(\theta) - g_0(\theta))\right| + \left|g_0(\theta)^\top(\widehat{W}_T - W)g_0(\theta)\right| \\
&\quad\leq \|\widehat{g}_T(\theta) - g_0(\theta)\|^2\|\widehat{W}_T\| + 2\|g_0(\theta)\|\|\widehat{g}_T(\theta) - g_0(\theta)\|\|\widehat{W}_T\| + \|g_0(\theta)\|^2\|\widehat{W}_T - W\|.
\end{aligned}$$

To prove $\sup_{\theta\in\Theta}\left|\widehat{Q}_T(\theta) - Q_0(\theta)\right| \xrightarrow{\mathbb{P}} 0$, we should now prove that $\sup_{\theta\in\Theta}\|\widehat{g}_T(\theta) - g_0(\theta)\| \xrightarrow{\mathbb{P}} 0$. By Θ compact, it is sufficient to prove that $\|\widehat{\boldsymbol{L}} - \boldsymbol{L}\| \xrightarrow{\mathbb{P}} 0$, $\|\widehat{\boldsymbol{C}}^{(T)} - \boldsymbol{C}\| \xrightarrow{\mathbb{P}} 0$, and $\|\widehat{\boldsymbol{K^c}}^{(T)} - \boldsymbol{K^c}\| \xrightarrow{\mathbb{P}} 0$.

Proof that $\|\widehat{\boldsymbol{L}} - \boldsymbol{L}\| \xrightarrow{\mathbb{P}} 0$

The estimator of $\boldsymbol{L}$ is unbiased so let's focus on the variance of $\widehat{\boldsymbol{L}}$.

$$\begin{aligned}
\mathbb{E}[(\widehat{\Lambda}^i - \Lambda^i)^2] &= \mathbb{E}\left[\left(\frac{1}{T}\int_{H_T}^{T+H_T}(dN_t^i - \Lambda^i dt)\right)^2\right] \\
&= \frac{1}{T^2}\int_{H_T}^{T+H_T}\int_{H_T}^{T+H_T}\mathbb{E}[(dN_t^i - \Lambda^i dt)(dN_{t'}^i - \Lambda^i dt')] \\
&= \frac{1}{T^2}\int_{H_T}^{T+H_T}\int_{H_T}^{T+H_T} C_{t'-t}^{ii} dt dt' \\
&\leq \frac{1}{T^2}\int_{H_T}^{T+H_T} C^{ii} dt = \frac{C^{ii}}{T} \longrightarrow 0
\end{aligned}$$

By Markov inequality, we have just proved that $\|\widehat{\boldsymbol{L}} - \boldsymbol{L}\| \xrightarrow{\mathbb{P}} 0$.

Proof that $\|\widehat{\boldsymbol{C}}^{(T)} - \boldsymbol{C}\| \xrightarrow{\mathbb{P}} 0$

First, let's remind that $\mathbb{E}(\widehat{\boldsymbol{C}}^{(T)}) \neq \boldsymbol{C}$. Indeed,

$$\begin{aligned}
\mathbb{E}\left(\widehat{C}^{ij,(T)}\right) &= \mathbb{E}\left(\frac{1}{T}\int_{H_T}^{T+H_T} dN_t^i \int_0^{T+2H_T} dN_{t'}^j f_{t'-t} - \widehat{\Lambda}^i\widetilde{\Lambda}^j F^T\right) \\
&= \mathbb{E}\left(\frac{1}{T}\int_{H_T}^{T+H_T} dN_t^i \int_{-t}^{T+2H_T-t} dN_{t+s}^j f_s - \Lambda^i\Lambda^j F^T\right) + \epsilon^{ij,T,H_T} F^T \\
&= \frac{1}{T}\int_{H_T}^{T+H_T}\int_{-H_T}^{H_T} f_s\mathbb{E}\left(dN_t^i dN_{t+s}^j - \Lambda^i\Lambda^j ds\right) + \epsilon^{ij,T,H_T} F^T \\
&= \int f_s C_s^{ij} ds + \epsilon^{ij,T,H_T} F^T
\end{aligned}$$

Now,

$$\begin{aligned}
\epsilon^{ij,T,H_T} &= \mathbb{E}\left(\Lambda^i\Lambda^j - \widehat{\Lambda}^i\widetilde{\Lambda}^j\right) \\
&= -\frac{1}{T^2}\int_{H_T}^{T+H_T}\int_0^{T+2H_T}\mathbb{E}\left(dN_t^i dN_{t'}^j - \Lambda^i\Lambda^j dt dt'\right) \\
&= -\frac{1}{T^2}\int_{H_T}^{T+H_T}\int_0^{T+2H_T} C_{t-t'}^{ij} dt dt' \\
&= -\frac{1}{T}\int\left(1+\left(\frac{H_T-|t|}{T}\right)^-\right)^+ C_t^{ij} dt
\end{aligned}$$

Since f satisfies $F^T = o(T)$, we have $\mathbb{E}(\widehat{\boldsymbol{C}}^{(T)}) \longrightarrow \boldsymbol{C}$. It remains now to prove that $\|\widehat{\boldsymbol{C}}^{(T)} - \mathbb{E}(\widehat{\boldsymbol{C}}^{(T)})\| \xrightarrow{\mathbb{P}} 0$.
Let's now focus on the variance of $\widehat{C}^{ij,(T)}$: $\quad \mathbb{V}(\widehat{C}^{ij,(T)}) = \mathbb{E}\left((\widehat{C}^{ij,(T)})^2\right) - \mathbb{E}(\widehat{C}^{ij,(T)})^2$.
Now,

$$\begin{aligned}
\mathbb{E}\left((\widehat{C}^{ij,(T)})^2\right) &= \mathbb{E}\left(\frac{1}{T^2}\sum_{(\tau,\eta,\tau',\eta')\in(Z^{i,T,1})^2\times(Z^{j,T,2})^2}(f_{\tau'-\tau} - F^T/(T+2H_T))(f_{\eta'-\eta} - F^T/(T+2H_T))\right) \\
&= \mathbb{E}\left(\frac{1}{T^2}\int_{t,s\in[H_T,T+H_T]}\int_{t',s'} dN_t^i dN_{t'}^j dN_s^i dN_{s'}^j (f_{t'-t} - F^T/(T+2H_T))(f_{s'-s} - F^T/(T+2H_T))\right) \\
&= \frac{1}{T^2}\int_{t,s\in[H_T,T+H_T]}\int_{t',s'\in[0,T+2H_T]}\mathbb{E}\left(dN_t^i dN_{t'}^j dN_s^i dN_{s'}^j\right)(f_{t'-t} - F^T/(T+2H_T))(f_{s'-s} - F^T/(T+2H_T))
\end{aligned}$$

And,

$$\mathbb{E}(\widehat{C}^{ij,(T)})^2 = \frac{1}{T^2}\int_{t,s\in[H_T,T+H_T]}\int_{t',s'\in[0,T+2H_T]}\mathbb{E}\left(dN_t^i dN_{t'}^j\right)\mathbb{E}\left(dN_s^i dN_{s'}^j\right)(f_{t'-t} - F^T/(T+2H_T))(f_{s'-s} - F^T/(T+2H_T))$$

Then, the variance involves the integration towards the difference of moments $\mu^{r,s,t,u} - \mu^{r,s}\mu^{t,u}$. Let's write it as a sum of cumulants, since cumulants density are integrable.

$$\begin{aligned}
\mu^{r,s,t,u} - \mu^{r,s}\mu^{t,u} &= \kappa^{r,s,t,u} + \kappa^{r,s,t}\kappa^u[4] + \kappa^{r,s}\kappa^{t,u}[3] + \kappa^{r,s}\kappa^t\kappa^u[6] + \kappa^r\kappa^s\kappa^t\kappa^u - (\kappa^{r,s} + \kappa^r\kappa^s)(\kappa^{t,u} + \kappa^t\kappa^u) \\
&= \kappa^{r,s,t,u} \\
&+ \kappa^{r,s,t}\kappa^u + \kappa^{u,r,s}\kappa^t + \kappa^{t,u,r}\kappa^s + \kappa^{s,t,u}\kappa^r \\
&+ \kappa^{r,t}\kappa^{s,u} + \kappa^{r,u}\kappa^{s,t} \\
&+ \kappa^{r,t}\kappa^s\kappa^u + \kappa^{r,u}\kappa^s\kappa^t + \kappa^{s,t}\kappa^r\kappa^u + \kappa^{s,t}\kappa^r\kappa^u
\end{aligned}$$

In the rest of the proof, we denote $a_t = \mathbb{1}_{t\in[H_T,T+H_T]}$, $b_t = \mathbb{1}_{t\in[0,T+2H_T]}$, $c_t = \mathbb{1}_{t\in[-H_T,H_T]}$, $g_t = f_t - \frac{1}{T+2H_T}F^T$
Before starting the integration of each term, let's remark that:

1. $\boldsymbol{\Psi}_t = \sum_{n\geq 1}\boldsymbol{\Phi}_t^{(\star n)} \geq 0$ since $\boldsymbol{\Phi}_t \geq 0$.
2. The regular parts of C_u^{ij}, $S_{u,v}^{ijk}$ (skewness density) and $K_{u,v,w}^{ijkl}$ (fourth cumulant density) are positive as polynoms of integrals of $\psi_\cdot^{ab}$ with positive coefficients. The integrals of the singular parts are positive as well.
3. (a) $\int a_t b_{t'} f_{t'-t} dt dt' = TF^T$
   (b) $\int a_t b_{t'} g_{t'-t} dt dt' = 0$
   (c) $\int a_t b_{t'} |g_{t'-t}| dt dt' \leq 2TF^T$

4. $\forall t \in \mathbb{R}, a_t(b \star \widetilde{g})_t = 0$, where $\widetilde{g}_s = g_{-s}$.

Fourth cumulant We want here to compute $\int \kappa^{i,j,i,j}_{t,t',s,s'} a_t b_{t'} a_s b_{s'} g_{t'-t} g_{s'-s} dt dt' ds ds'$.
We remark that $|g_{t'-t} g_{s'-s}| \leq (||f||_\infty(1 + 2H_T/T))^2 \leq 4||f||_\infty^2$.

$$\begin{aligned}
\left|\frac{1}{T^2}\int \kappa^{i,j,i,j}_{t,t',s,s'} a_t b_{t'} a_s b_{s'} g_{t'-t} g_{s'-s} dt dt' ds ds'\right| &\leq \left(\frac{2||f||_\infty}{T}\right)^2 \int dt a_t \int dt' b_{t'} \int ds a_s \int ds' b_{s'} K^{ijij}_{t'-t,s-t,s'-t} \\
&\leq \left(\frac{2||f||_\infty}{T}\right)^2 \int dt a_t \int dt' b_{t'} \int ds a_s \int dw K^{ijij}_{t'-t,s-t,w} \\
&\leq \left(\frac{2||f||_\infty}{T}\right)^2 \int dt a_t \int K^{ijij}_{u,v,w} du dv dw \\
&\leq \frac{4||f||_\infty^2}{T} K^{ijij} \underset{T\to\infty}{\longrightarrow} 0
\end{aligned}$$

Third $\times$ First We have four terms, but only two different forms since the roles of (s, s') and (t, t') are symmetric.
First form

$$\begin{aligned}
\int \kappa^{i,j,i}_{t,t',s} \Lambda^j G_{\boldsymbol{t}} d\boldsymbol{t} &= \frac{\Lambda^j}{T^2}\int \kappa^{i,j,i}_{t,t',s} a_t b_{t'} a_s b_{s'} g_{t'-t} g_{s'-s} dt dt' ds ds' \\
&= \frac{\Lambda^j}{T^2}\int \kappa^{i,j,i}_{t,t',s} a_t b_{t'} a_s (b \star \widetilde{g})_s g_{t'-t} dt dt' ds \\
&= 0 \qquad \text{since } a_s(b \star \widetilde{g})_s = 0
\end{aligned}$$

Second form

$$\begin{aligned}
\left|\int \kappa^{i,j,j}_{t,t',s'} \Lambda^i G_{\boldsymbol{t}} d\boldsymbol{t}\right| &= \left|\frac{\Lambda^i}{T^2}\int \kappa^{i,j,j}_{t,t',s'} a_t b_{t'} a_s b_{s'} g_{t'-t} g_{s'-s} dt dt' ds ds'\right| \\
&= \left|\frac{\Lambda^i}{T^2}\int \kappa^{i,j,j}_{t,t',s'} a_t b_{t'} g_{t'-t} b_{s'} (a \star g)_{s'} dt dt' ds'\right| \\
&\leq \frac{\Lambda^i}{T^2} 2||f||_\infty \int ds' b_{s'} (a \star |g|)_{s'} \int dt a_t \int dt' b_{t'} S^{ijj}_{t'-s',t-s'} \\
&\leq 4||f||_\infty S^{ijj} \Lambda^i \frac{F^T}{T} \underset{T\to\infty}{\longrightarrow} 0
\end{aligned}$$

Second $\times$ Second
First form

$$\begin{aligned}
\left|\int \kappa^{i,i}_{t,s} \kappa^{j,j}_{t',s'} G_{\boldsymbol{t}} d\boldsymbol{t}\right| &\leq \frac{2||f||_\infty}{T^2}\int C^{ii}_{t-s} C^{jj}_{t'-s'} a_t b_{t'} |g_{t'-t}| a_s b_{s'} dt dt' ds ds' \\
&\leq \frac{2||f||_\infty}{T^2} C^{ii} C^{jj} \int a_t b_{t'} |g_{t'-t}| dt dt' \\
&\leq 4||f||_\infty C^{ii} C^{jj} \frac{F^T}{T} \underset{T\to\infty}{\longrightarrow} 0
\end{aligned}$$

Second form

$$\left|\int \kappa^{i,j}_{t,s'} \kappa^{i,j}_{t',s} G_{\boldsymbol{t}} d\boldsymbol{t}\right| \leq 4||f||_\infty (C^{ij})^2 \frac{F^T}{T} \underset{T\to\infty}{\longrightarrow} 0$$

Second $\times$ First $\times$ First
First form

$$\int \kappa^{i,j}_{t,t'} \Lambda^i \Lambda^j G_{\boldsymbol{t}} d\boldsymbol{t} = \frac{\Lambda^i \Lambda^j}{T^2}\int \kappa^{i,j}_{t,t'} a_t b_{t'} g_{t'-t} dt dt' \int a_s b_{s'} g_{s'-s} ds ds' = 0$$

Second form

$$\int \kappa^{i,i}_{t,s} \Lambda^j \Lambda^j G_{\boldsymbol{t}} d\boldsymbol{t} = \left(\frac{\Lambda^j}{T}\right)^2 \int \kappa^{i,i}_{t,s} a_t b_{t'} g_{t'-t} a_s (b \star \widetilde{g})_s dt dt' ds = 0$$

We have just proved that $\mathbb{V}(\widehat{\boldsymbol{C}}^{(T)}) \xrightarrow{\mathbb{P}} 0$. By Markov inequality, it ensures us that $\|\widehat{\boldsymbol{C}}^{(T)} - \mathbb{E}(\widehat{\boldsymbol{C}}^{(T)})\| \xrightarrow{\mathbb{P}} 0$, and finally that $\|\widehat{\boldsymbol{C}}^{(T)} - \boldsymbol{C}\| \xrightarrow{\mathbb{P}} 0$. □

Proof that $\|\widehat{\boldsymbol{K^c}}^{(T)} - \boldsymbol{K^c}\| \xrightarrow{\mathbb{P}} 0$

The scheme of the proof is similar to the previous one. The upper bounds of the integrals involve the same kind of terms, plus the new term $(F^T)^2/T$ that goes to zero thanks to the assumption 5 of the theorem.

References

Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054, 1982.

Stojan Jovanović, John Hertz, and Stefan Rotter. Cumulants of hawkes point processes. *Physical Review E*, 91(4): 042802, 2015.

W. K Newey and D. McFadden. Large sample estimation and hypothesis testing. *Handbook of econometrics*, 4: 2111–2245, 1994.