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

Massil Achab\(^1\), Emmanuel Bacry\(^1\), Stéphane Gaïffas\(^1\), 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 \(E[\hat{g}_T(\theta_0)] = m_T \neq 0\). 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 processes’ 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 \(\int f^T_t C_{ij}^t \, dt\) and \(\int f^T_t K_{ijk}^t \, dt\) with \(f^T\) a compact supported function on \([-H_T, H_T]\) that weakly converges to 1, with \(H_T \to \infty\). In most cases we will take \(f^T_t = 1_{[-H_T, H_T]}(t)\).

Denoting \(\hat{C}_{ij}(T)\) the estimator of \(\int f^T_t C_{ij}^t \, dt\), the term \(|E[\hat{C}_{ij}(T)] - C_{ij}| = |\int f^T_t C_{ij}^t \, dt - C_{ij}|\) can be considered a proxy to the difference to the classical GMM. This distance has to go to zero to make the rest of GMM’s proof work: the estimator \(\hat{C}_{ij}(T)\) is then asymptotically unbiased towards \(C_{ij}\) when \(T\) goes to infinity.

1.3 Notations

We observe the multivariate point process \((N_t)\) 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  \(G^{\text{true}} = \int \Phi_t \, dt = (\int \phi_t^{ij} \, dt)_{ij} = I_d - (R^{\text{true}})^{-1}\)

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

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

massil.achab@m4x.org
Theoretical skewness \( K^e = (K^{ij})_{ij} = (R^{\text{true}})_{ij}^2 C^T + 2 [R^{\text{true}} \odot (C - R^{\text{true}} L)] (R^{\text{true}})^T \)

Filtering function \( f^T \geq 0 \; \text{supp}(f^T) \subset [-H_T, H_T] \quad FT = \int f^T_s ds \quad \hat{f}^T_t = f^T_{-t} \)

Events sets \( Z^{i, T, 1} = Z^i \cap [H_T, T + H_T] \quad Z^{i, T, 2} = Z^i \cap [0, T + 2H_T] \)

Estimators of the mean \( \hat{\Lambda}^i = \frac{N^i_{T+H_T} - N^i_{H_T}}{T} = \frac{N^i_{T+2H_T}}{T+2H_T} \)

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

Estimator of the skewness \( \hat{K}^{ijk, (T)} = \frac{1}{T} \sum_{\tau \in Z^{i, T, 1}} \left( \sum_{\tau' \in Z^{i, T, 2}} f^{\tau - \tau'} - \hat{\Lambda}^j F^T \right) \left( \sum_{\tau'' \in Z^{k, T, 2}} f^{\tau - \tau''} - \hat{\Lambda}^k F^T \right) - \frac{\hat{\Lambda}^i}{T + 2H_T} \sum_{\tau' \in Z^{i, T, 2}} \left( \sum_{\tau'' \in Z^{k, T, 2}} (f^{T} \hat{f}^{T})^{\tau - \tau''} - \hat{\Lambda}^k (F^T)^2 \right) \)

GMM related notations

\[
\begin{align*}
\theta &= R \quad \text{and} \quad \theta_0 = R^{\text{true}} \\
 g_0(\theta) &= \text{vec} \left[ K^e - R^{\odot 2} C^T + 2 [R \odot (C - R L)] R^T \right] \in \mathbb{R}^{2d^2} \\
 \hat{g}_T(\theta) &= \text{vec} \left[ \hat{C}^{(T)} - \tilde{R} \tilde{L} R^T \right] \in \mathbb{R}^{2d^2} \\
 Q_0(\theta) &= g_0(\theta)^T W g_0(\theta) \\
 \hat{Q}_T(\theta) &= \hat{g}_T(\theta)^T \hat{W}_T \hat{g}_T(\theta) 
\end{align*}
\]

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 \( \theta_0 \); (ii) \( \Theta \) is compact; (iii) \( Q_0(\theta) \) is continuous; (iv) \( \hat{Q}_T(\theta) \) converges uniformly in probability to \( Q_0(\theta) \), then \( \hat{\theta}_T = \arg \max \hat{Q}_T(\theta) \xrightarrow{p} \theta_0 \).

We can now prove the consistency of our estimator.

**Theorem 2.2.** Suppose that \( N_k \) is observed on \( \mathbb{R}^+, \tilde{W}_T \xrightarrow{p} W \), 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 \( \| \Phi \|_* < 1 \),
4. \( \forall i, j, k \in [d], \int f^T_u C^{ij}_u du \rightarrow \int C^{ij}_u du \) and \( \int f^T_u f^T_v R^{ijk}_{uv} du dv \rightarrow \int R^{ijk}_{uv} du dv \),

\footnote{When \( f^T_t = 1_{[-H_T, H_T]}(t) \), we remind that \( (f^T \hat{f}^T)_t = (2H_T - |t|)^+ \). This leads to the estimator we showed in the article.}
5. \((FT)^2/T \xrightarrow{p} 0\) and \(\|f\|_\infty = O(1)\).

Then

\[ \hat{\theta}_T \xrightarrow{p} \theta_0. \]

**Remark 1.** In practice, we use a constant sequence of weighting matrices: \(\hat{W}_T = I_d\).

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

\[
|\hat{Q}_T(\theta) - Q_0(\theta)| \\
\leq |(\hat{g}_T(\theta) - g_0(\theta))^T \hat{W}_T (\hat{g}_T(\theta) - g_0(\theta))| + |g_0(\theta)^T (\hat{W}_T + \hat{W}_T^2) (\hat{g}_T(\theta) - g_0(\theta))| + |g_0(\theta)^T (\hat{W}_T - W) g_0(\theta)| \\
\leq \|\hat{g}_T(\theta) - g_0(\theta)\|^2 \|\hat{W}_T\| + 2 \|g_0(\theta)\| \|\hat{g}_T(\theta) - g_0(\theta)\| \|\hat{W}_T\| + \|g_0(\theta)\|^2 \|\hat{W}_T - W\|.
\]

To prove \(\sup_{\theta \in \Theta} |\hat{Q}_T(\theta) - Q_0(\theta)| \xrightarrow{p} 0\), we should now prove that \(\sup_{\theta \in \Theta} |\hat{Q}_T(\theta) - g_0(\theta)| \xrightarrow{p} 0\). By \(\Theta\) compact, it is sufficient to prove that \(\|\tilde{L} - L\| \xrightarrow{p} 0\), \(\|C(\cdot)^T - C\| \xrightarrow{p} 0\), and \(\|\tilde{K}(\cdot)^T - K\| \xrightarrow{p} 0\).

**Proof that** \(\|\tilde{L} - L\| \xrightarrow{p} 0\)

The estimator of \(L\) is unbiased so let’s focus on the variance of \(\tilde{L}\).

\[
E[(\tilde{\Lambda}^i - \Lambda^i)^2] = E \left[ \left( \frac{1}{T} \int_{H_T}^{T+H_T} (dN^i_t - \Lambda^i dt) \right)^2 \right] \\
= \frac{1}{T^2} \int_{H_T}^{T+H_T} \int_{H_T}^{T+H_T} E[(dN^i_t - \Lambda^i dt)(dN^i_{t'} - \Lambda^i dt')] \\
= \frac{1}{T^2} \int_{H_T}^{T+H_T} \int_{H_T}^{T+H_T} C^{ii}_{t,t'} dt dt' \\
\leq \frac{1}{T^2} \int_{H_T}^{T+H_T} C^{ii} dt = \frac{C^{ii}}{T} \xrightarrow{T \to \infty} 0
\]

By Markov inequality, we have just proved that \(\|\tilde{L} - L\| \xrightarrow{p} 0\).

**Proof that** \(\|\tilde{C}(\cdot)^T - C\| \xrightarrow{p} 0\)

First, let’s remind that \(E(C(\cdot)^T) \neq C\). Indeed,

\[
E(\hat{C}^{ij}(\cdot)^T) = E \left( \frac{1}{T} \int_{H_T}^{T+H_T} dN^i_t \int_0^{T+2H_T} dN^j_{t'} f_{t-t'} - \hat{N}^i \hat{N}^j F^T \right) \\
= E \left( \frac{1}{T} \int_{H_T}^{T+H_T} dN^i_t \int_0^{T+2H_T} dN^j_{t'} f_{t-t'} - \Lambda^i \Lambda^j F^T \right) + \epsilon^{ij, T, H_T} F^T \\
= \frac{1}{T} \int_{H_T}^{T+H_T} \int_{-T}^{T} f_s \mathbb{E} (dN^i_t dN^j_{t+s} - \Lambda^i \Lambda^j ds) + \epsilon^{ij, T, H_T} F^T \\
= \int_{s} f_s C^{ij} ds + \epsilon^{ij, T, H_T} F^T
\]
Now,
\[
\epsilon_{ij,T,H_T} = \mathbb{E} \left( \Lambda^i \Lambda^j - \hat{\Lambda}^i \hat{\Lambda}^j \right)
\]
\[
= -\frac{1}{T^2} \int_{H_T}^{T+H_T} \int_{0}^{T+2H_T} \mathbb{E} \left( dN^i_t dN^j_{t'} - \Lambda^i \Lambda^j dt dt' \right)
\]
\[
= -\frac{1}{T^2} \int_{H_T}^{T+H_T} \int_{0}^{T+2H_T} C_{t,t'}^{ij} dt dt'
\]
\[
= -\frac{1}{T} \int \left( 1 + \frac{|H_T - |t|}{T} \right) C_{t}^{ij} dt
\]

Since \( f \) satisfies \( F^T = o(T) \), we have \( \mathbb{E}(\hat{C}(T)) \to C \). It remains now to prove that \( \| \hat{C}(T) - \mathbb{E}(\hat{C}(T)) \|_p \to 0 \).

Let’s now focus on the variance of \( \hat{C}^{ij}(T) \) : \( \forall (\hat{C}^{ij}(T)) = \mathbb{E} \left( (\hat{C}^{ij}(T))^2 \right) - \mathbb{E}(\hat{C}^{ij}(T))^2 \).

Now,
\[
\mathbb{E} \left( (\hat{C}^{ij}(T))^2 \right) = \mathbb{E} \left( \frac{1}{T^2} \sum_{(\tau,\eta',\eta) \in (Z^T)^2} (f_{\tau'-\tau} - F^T / (T + 2H_T))(f_{\eta' - \eta} - F^T / (T + 2H_T)) \right)
\]
\[
= \mathbb{E} \left( \frac{1}{T^2} \int_{t,s \in [H_T, T + H_T]} \int_{t',s' \in [0, T + 2H_T]} dN^i_t dN^j_{t'} dN^i_s dN^j_{s'} (f_{t'-t} - F^T / (T + 2H_T))(f_{s'-s} - F^T / (T + 2H_T)) \right)
\]
\[
= \frac{1}{T^2} \int_{t,s \in [H_T, T + H_T]} \int_{t',s' \in [0, T + 2H_T]} \mathbb{E} \left( dN^i_t dN^j_{t'} dN^i_s dN^j_{s'} \right) (f_{t'-t} - F^T / (T + 2H_T))(f_{s'-s} - F^T / (T + 2H_T))
\]

And,
\[
\mathbb{E} \left( \hat{C}^{ij}(T) \right)^2 = \frac{1}{T^2} \int_{t,s \in [H_T, T + H_T]} \int_{t',s' \in [0, T + 2H_T]} \mathbb{E} \left( dN^i_t dN^j_{t'} \right) \mathbb{E} \left( dN^i_s dN^j_{s'} \right) (f_{t'-t} - F^T / (T + 2H_T))(f_{s'-s} - F^T / (T + 2H_T))
\]

Then, the variance involves the integration towards the difference of moments \( \mu_r s t u - \mu_r s t u \). Let’s write it as a sum of cumulants, since cumulants density are integrable.

\[
\mu_r s t u - \mu_r s t u = \kappa_r \kappa_s \kappa_t \kappa_u [4] + \kappa_r \kappa_s \kappa_t \kappa_u [3] + \kappa_r \kappa_s \kappa_t \kappa_u [6] + \kappa_r \kappa_s \kappa_t \kappa_u - (\kappa_r \kappa_s + \kappa_r \kappa_u)(\kappa_t \kappa_s + \kappa_t \kappa_u)
\]

In the rest of the proof, we denote \( a_t = 1_{t \in [H_T, T + H_T]} \), \( b_t = 1_{t \in [0, T + 2H_T]} \), \( c_t = 1_{t \in [-H_T, H_T]} \), \( g_t = f_t - \frac{1}{T + 2H_T} F_T \).

Before starting the integration of each term, let’s remark that:

1. \( \Psi_t = \sum_{n \geq 1} \Phi_t^{(r^n)} \geq 0 \) since \( \Phi_t \geq 0 \).
2. The regular parts of \( C^{ij} \), \( S^{ijk} \) (skewness density) and \( K^{ijkl} \) (fourth cumulant density) are positive as polynomials of integrals of \( \psi^{ab} \) 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. \( \forall t \in \mathbb{R}, a_t (b \ast \bar{g})_t = 0 \), where \( \bar{g}_s = g_{-s} \).

**Fourth cumulant** We want here to compute \( \int \kappa_{t,t',s,s'}^{i,i,j,j} a_t b_{t'} a_s b_{s'} g_{t'-t} g_{s'-s} dt dt' ds ds' \).

We remark that \( |g_{t'-t} g_{s'-s}| \leq (||f||_\infty (1 + 2H_T/T))^2 \leq 4||f||^2_\infty \).

\[
\left| \frac{1}{T^2} \int \kappa_{t,t',s,s'}^{i,i,j,j} a_t b_{t'} a_s b_{s'} g_{t'-t} g_{s'-s} dt dt' ds ds' \right| \leq \left( \frac{2||f||_\infty}{T} \right)^2 \int dt dt' \int ds a_s \int ds' b_{s'} K_{t'-t,s-t,s'-t}^{ijjj} \\
\leq \left( \frac{2||f||_\infty}{T} \right)^2 \int dt dt' \int ds a_s \int dw K_{t'-t,s-t,w}^{ijij} \\
\leq \left( \frac{2||f||_\infty}{T} \right)^2 \int dt dt' \int K_{t'-t,s-t,w}^{jjjj} \to 0
\]

**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

\[
\int \kappa_{t,t',s,s'}^{i,i,j,j} N^j G_t dt = \frac{\Lambda^j}{T^2} \int \kappa_{t,t',s',s'}^{i,i,j,j} a_t b_{t'} a_s b_{s'} g_{t'-t} g_{s'-s} dt dt' ds ds' \\
= \frac{\Lambda^j}{T^2} \int \kappa_{t,t',s',s'}^{i,i,j,j} a_t b_{t'} a_s (b \ast \bar{g})_s g_{t'-t} dt dt' ds \\
= 0 \text{ since } a_s (b \ast \bar{g})_s = 0
\]

Second form

\[
\left| \int \kappa_{t,t',s,s'}^{i,i,j,j} N^j G_t dt \right| = \left| \frac{\Lambda^j}{T^2} \int \kappa_{t,t',s',s'}^{i,i,j,j} a_t b_{t'} a_s b_{s'} g_{t'-t} g_{s'-s} dt dt' ds ds' \right| \\
= \left| \frac{\Lambda^j}{T^2} \int \kappa_{t,t',s',s'}^{i,i,j,j} a_t b_{t'} g_{t'-t} b_{s'} g_{s'-s} dt dt' ds ds' \right| \\
\leq \frac{\Lambda^j}{T^2} 2||f||_\infty \int ds b_{s'} (a \ast |g|)_{s'} \int dt dt' b_{t'} S_{t'-s',t-s'}^{jj} \\
\leq 4||f||_\infty S_{t'-t}^{jj} \Lambda^j F_T T \to 0
\]

**Second \times Second**

First form

\[
\left| \int \kappa_{t,t',s,s'}^{i,i,j,j} G_t dt \right| \leq \frac{2||f||_\infty}{T^2} \int C_{t'-t}^{ii} C_{t'-s}^{jj} a_t b_{t'} |g_{t'-t}| a_s b_{s'} dt dt' ds ds' \\
\leq \frac{2||f||_\infty}{T^2} C_{t'-t}^{ii} C_{t'-s}^{jj} \int a_t b_{t'} |g_{t'-t}| dt dt' \\
\leq 4||f||_\infty C_{t'-t}^{ii} C_{t'-s}^{jj} F_T T \to 0
\]

Second form

\[
\left| \int \kappa_{t,t',s,s'}^{i,j,i,j} G_t dt \right| \leq 4||f||_\infty (C_{t'-t}^{ij})^2 \frac{F_T}{T} T \to 0
\]

**Second \times First \times First**

First form

\[
\int \kappa_{t,t',s,s'}^{i,j,i,j} N^j G_t dt = \frac{\Lambda^j}{T^2} \int \kappa_{t,t',s,s'}^{i,j,i,j} a_t b_{t'} g_{t'-t} dt dt' \int a_s b_{s'} g_{s'-s} ds ds' = 0
\]
Second form

\[
\int \kappa_{t,s}^i \Lambda^j G_t \, dt = \left( \frac{\Lambda^j}{T} \right)^2 \int \kappa_{t,s}^i a_t b_t g_{t-t'} a_s (b \ast \bar{g})_s \, dt \, ds = 0
\]

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

**Proof that** \( \| \hat{K}^{(T)} - K^c \| \xrightarrow{p} 0 \)

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**

