Detecting Structural Shifts in Multivariate Hawkes Processes with Fréchet Statistic

Rui Luo and Vikram Krishnamurthy

Abstract— This paper proposes a new approach for change point detection in multivariate Hawkes processes using Fréchet statistic of a network. The method splits the point process into overlapping windows, estimates kernel matrices in each window, and reconstructs the signed Laplacians by treating the kernel matrices as the adjacency matrices of the causal network. We demonstrate the effectiveness of our method through experiments on both simulated and cryptocurrency datasets. Our results show that our method is capable of accurately detecting and characterizing changes in the causal structure of multivariate Hawkes processes, and may have potential applications in fields such as finance and neuroscience. The proposed method is an extension of previous work on Fréchet statistics in point process settings and represents an important contribution to the field of change point detection in multivariate point processes.

I. INTRODUCTION

Networked systems, such as social [1], [2], biological [3], and financial networks [4], often experience events that trigger cascading effects. Social interactions can escalate into conflicts, create polarized groups, or spread misinformation. Neuronal activities in the brain may incite further stimulations or inhibitions. Similarly, initial shocks in stock markets can propagate volatility. These event chains are crucial to understanding the dynamics of practical networks.

Multivariate Hawkes processes are crucial for modeling event occurrences across fields like neuroscience, finance, and social science, effectively capturing self- and cross-excitation among multiple channels that lead to cascading events. Detecting change points in these processes is vital, with current methods primarily focusing on kernel function estimation using generalized likelihood ratio (GLR) or CUSUM procedures. For instance, Yamin et al. [5] proposed a modified CUSUM and LRT-based method for supply chain disruptions, while Wang et al. [6] developed a CUSUM procedure for sequential change-point detection in Hawkes networks. Bayesian approaches, such as those by Li et al. [7] and Detommaso [8], leverage prior information and tackle complex models. Additionally, Linderman et al. [9] introduced a random graph

model to reveal latent networks in multivariate Hawkes processes. Collectively, these methods advance change point detection research and enhance our understanding of underlying dynamics and future predictions.

In contrast to previous work, we introduce a novel method that simplifies the task of detecting shifts in magnitude and frequency of events by reconstructing the kernel's integral, eliminating the need for full kernel recovery. This method utilizes Fréchet statistics of signed Laplacian matrices to construct test statistics that detect changes, offering a unique network-based perspective on the causal structures driving event dynamics.

Our main contribution is formulating change point detection in multivariate Hawkes processes as a network hypothesis testing problem, enhancing our understanding of underlying causality and providing a robust tool for analyzing multivariate time series data. The efficacy of this approach is validated through its application to practical datasets, effectively identifying significant changes. The paper is organized as follows: Section II outlines the multivariate Hawkes process and defines a kernel matrix that captures the causal network, estimated using the NPHC algorithm in Section II-A. Section III explores a novel signed Laplacian metric space, enabling the derivation of a computationally efficient Fréchet mean and variance for these causal structures. Our methodology's robustness is demonstrated in Section IV through experiments on simulated data and real cryptocurrency markets, showcasing its effectiveness in pinpointing change points.

II. Multivariate Hawkes Process

In this section, we outline the formulation of a multivariate Hawkes process. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and consider a multivariate Hawkes process $N(t) = (N_1, \cdots, N_m)$, where $N_i(t), 1 \le i \le m$, is the counting process representing the cumulative number of events up to time t for subject i. We define a set of increasing σ-algebras $\{\mathcal{F}^t\}_{t \ge 0}$, where $\mathcal{F}^t = \sigma\{N^t\}$, and the non-negative, $\mathcal{F}^t$-measurable process $\lambda(t)$ as the intensity of $N(t)$, given by:

$$
\begin{aligned}
P(dN_i(t) = 1|\mathcal{F}^t) &= \lambda_i(t)dt + o(dt), \\
P(dN_i(t) > 1|\mathcal{F}^t) &= o(dt),
\end{aligned}
\tag{1}
$$

R. Luo is with the Department of Systems Engineering, City University of Hong Kong, Hong Kong SAR, China ruiluo@cityu.edu.hk

V. Krishnamurthy is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, 14850, USA vikramk@cornell.edu

where $o(dt)$ is little-o of dt and $\lim_{dt\to 0}\frac{o(dt)}{dt} = 0$. The intensity $\lambda_i(t)$ for each component is given by:

$$\lambda_i(t) = \mu_i + \sum_{j=1}^{m}\int_0^t \phi_{ij}(t-s)dN_j(s), \qquad (2)$$

where μ_i is the background rate and $\phi_{ij}(\cdot)$ is the kernel function representing the effects of process N_j on process N_i. A popular choice for $\phi_{ij}(\cdot)$ is an exponential function $\phi_{ij}(t) = \alpha_{ij}e^{-\beta_{ij}t}$, and (2) can be rewritten as

$$\lambda_i(t) = \mu_i + \sum_{j=1}^{m}\int_0^t \alpha_{ij}e^{-\beta_{ij}(t-s)}dN_j(s). \qquad (3)$$

Expressing further in matrix form:

$$\lambda(t) = \mu + \int_0^t \phi(t-s)dN(s), \qquad (4)$$

where $\phi(\cdot)$ is an $m \times m$ kernel matrix with entries $\phi_{ij}(\cdot)$. [10] has shown that $\phi(\cdot)$ (which they referred to as the excitation matrix) with exponential kernel functions (3) can reveal the casual structure underlying the multivariate components. Specifically, $N_i(t)$ is Granger non-causal for $N_j(t)$ if and only if the corresponding kernel function $\phi_{ij}(\cdot) = 0$ [11].

To infer the causal structure from $\phi(\cdot)$, we use the integrals of its elements to construct the matrix $H = [h_{ij}]$, where

$$h_{ij} = \int_0^\infty \phi_{ij}(t)dt. \qquad (5)$$

For kernels with exponential functions (3), we have $h_{ij} = \frac{\alpha_{ij}}{\beta_{ij}}$. h_{ij} is commonly referred to as the branching ratio, which quantifies the average number of events triggered by a single event, providing a measure of the endogenous effect in finance [12]. Furthermore, the cluster representation of Hawkes processes [13] reveals that the integral in the expression for the intensity of a multivariate Hawkes process signifies the average total number of events for subject i that are directly triggered by an event for subject j.

A. Kernel Matrix Estimation based on Integrated Cumulants

The Non-Parametric Hawkes with Cumulants (NPHC) algorithm [11] is a moment-matching method that facilitates the estimation of the kernel matrix H without the need to estimate the kernel functions' shape, which is based on the idea that the integrated cumulants of a Hawkes process can be expressed explicitly as a function of H [14].

For an arbitrary m-dimensional random vector $X = (X_1, \cdots, X_m)$, the cumulant of order m, denoted by $k(X_{\bar{m}})$, where $\bar{m}$ indicates the set $\{1, \cdots, m\}$, is defined as

$$k(X_{\bar{m}}) = \sum_\pi (|\pi| - 1)!(-1)^{|\pi|-1}\prod_{B\in\pi}\langle X_B\rangle, \qquad (6)$$

where the sum goes over all partitions π of the set $\{1, \cdots, m\}$, $|\cdot|$ denotes the number of blocks of a given partition, and $\langle X_B\rangle = \langle\prod_{k\in B}X_k\rangle$. The integrated cumulants provide a measure of the average correlated activity between events of different subjects, which is a natural generalization of the covariance of two variables to higher dimensions.

Note that if H has a spectral radius ρ strictly smaller than 1, $N(t)$ has asymptotically stationary increments, and $\lambda(t)$ is asymptotically stationary [15]. Consequently, $R = (I_d - H)^{-1}$ can be defined.

The first three integrated cumulants of a multivariate Hawkes process can be used to estimate the mean intensity, integrated covariance density matrix, and skewness of the process by equations. Specifically, the mean intensity Λ_i and the integrated covariance density matrix C_{ij} can be expressed as linear combinations of the vector of background intensities μ and the matrix R_{ij}, while the skewness K_{ijk} can be expressed as a polynomial in R_{ij} and C_{ij} [11]:

$$\Lambda_i = \sum_{l=1}^{m} R_{il}\mu_l, \quad C_{ij} = \sum_{l=1}^{m}\Lambda_l R_{il}R_{jl}, \qquad (7)$$

$$K_{ijk} = \sum_{1=1}^{m}(R_{il}R_{jl}C_{kl} + R_{il}C_{jl}R_{kl} + C_{il}R_{jl}R_{kl}$$
$$- 2\Lambda_l R_{il}R_{jl}R_{kl}). \qquad (8)$$

The estimation procedure assumes that truncating the integration from $(-\infty, +\infty)$ to $[-W, W]$ introduces only a small error. Given a realization of a stationary Hawkes process $N(t), t \in [0, T]$, where Z_i represents the set of events corresponding to subject i, the estimators for the cumulants can be obtained under this assumption:

$$\hat{\Lambda}_i = \frac{N_i(T)}{T}, \qquad (9)$$

$$\hat{C}_{ij} = \frac{1}{T}\sum_{\tau\in Z_i}(N_j(\tau + W) - N_j(\tau - W) - 2W\hat{\Lambda}_j), \qquad (10)$$

$$\hat{K}_{ijk} = \frac{1}{T}\sum_{\tau\in Z_i}\Big[(N_j(\tau + W) - N_j(\tau - W) - 2W\hat{\Lambda}_j)$$
$$.(N_k(\tau + W) - N_k(\tau - W) - 2W\hat{\Lambda}_k)\Big]$$
$$- \frac{\hat{\Lambda}_i}{T}\sum_{\tau\in Z_j}\sum_{\tau'\in Z_k}(2W - |\tau' - \tau|)^+ + 4W^2\hat{\Lambda}_i\hat{\Lambda}_j\hat{\Lambda}_k. \qquad (11)$$

Since the covariance C provides fewer independent coefficients than the kernel matrix H, the NPHC algorithm focuses on a subset of m^2 third-order cumulant coefficients K_{iij}, namely $K^c = \{K_{iij}\}_{1\le i,j\le m}$. Specifically, the estimator of R, denoted $\hat{R}$, is obtained by minimizing the Frobenius norm of two differences:

$$\hat{R} = \arg\min_R \mathcal{L}(R)$$
$$= \arg\min_R (1 - \kappa)\|K^c(R) - \hat{K}^c\|_2^2 + \kappa\|C(R) - \hat{C}\|_2^2,$$

where $\|\cdot\|_2$ is the Frobenius norm, while $\hat{C}$ and $\hat{K}^c$ are the respective estimators of C and K^c as defined in (10), (11). Finally, once $\hat{R}$ is obtained, the kernel matrix H can be estimated as $\hat{H} = I_m - \hat{R}^{-1}$.

III. Fréchet Statistics Based Change Point Detection

In this section, we propose a change point detection method that utilizes the Fréchet statistics of a sequence of causal networks. We construct these networks using the estimated kernel matrix $\hat{H}$ obtained in Section II-A. By computing the Fréchet distance between the signed Laplacian matrices of these networks, we can detect significant changes in the underlying structure of the multivariate Hawkes process.

A. Dynamic Causal Network

Our method employs a dynamic signed causal network framework with edges that can have either positive or negative weights. We calculate the node degree by summing the absolute values of the weights on the incoming edges. This degree information is incorporated into the degree matrix, which is crucial in calculating the signed Laplacian for the computation of Fréchet statistics in Section III-B.

A causal network is a directed graph where a_{ij} of its weighted adjacency matrix A indicates node j's influence on node i, with negative weights for inhibitory influences. A dynamic causal network is a sequence of graph snapshots $G^{(1)}, G^{(2)}, \cdots$, where each snapshot $G^{(t)} = (V^{(t)}, E^{(t)})$ represents the causal network observed at time t. We require that $V^{(1)} = V^{(2)} = \cdots = V$, meaning that the multivariate Hawkes process has a fixed set of components.

To construct a snapshot of the causal network, we use a sliding window approach. Specifically, we estimate the kernel matrix $\hat{H}$ from a sliding window of the multivariate Hawkes process and take the symmetric part, $A = (\hat{H} + \hat{H}^T)/2$, as the adjacency matrix for the snapshot of the causal network[1]. Denote $A^{(t)}$ as the adjacency matrix of the causal network during the t-th sliding window. The signed Laplacian $\bar{L}^{(t)}$ is defined as $\bar{L}^{(t)} = \bar{D}^{(t)} - A^{(t)}$, where $\bar{D}^{(t)}$ is the diagonal degree matrix given by $\bar{d}_{ii}^{(t)} = \sum_{j=1}^{m} |a_{ij}^{(t)}|$, that is, the diagonal entries are the unsigned degree of each node.

B. Fréchet Mean of Signed Laplacian Matrices

To measure the dissimilarity between two snapshots in a dynamic causal network, we define a metric space [16] based on the Log-Euclidean metric of the nearest symmetric positive definite (SPD) matrices of their corresponding signed Laplacian matrices. The introduced metric enables us to compare the structures of different snapshots of the dynamic causal network. Moreover,

[1]In experiments presented in Section IV, we demonstrate that the antisymmetric part of $\hat{H}$, i.e., $(\hat{H} - \hat{H}^T)/2$, is negligible.

it admits a closed-form Fréchet mean, which enables efficient computation.

The signed Laplacian matrix $\bar{L}^{(t)}$ is known to be positive semi-definite [17], which can cause problems when using SPD metrics like the Log-Euclidean metric. We address this issue by finding the nearest SPD matrix to $\bar{L}^{(t)}$ in the Frobenius norm [18]. The resulting SPD matrix, denoted by $\tilde{L}^{(t)}$, is used instead of $\bar{L}^{(t)}$ to define the metric space.

We introduce a metric space $(\tilde{\mathcal{L}}, \mathrm{d})$ for dynamic causal networks. Here, $\tilde{\mathcal{L}}$ denotes the set of nearest SPD matrices to the signed Laplacians, and d is a function $\mathrm{d}, : \tilde{\mathcal{L}} \times \tilde{\mathcal{L}} \to \mathbb{R}_+$ defined using the Log-Euclidean metric, i.e., $\delta_{LE}(X, Y) = |\log(X) - \log(Y)|_{\mathrm{F}}$.

Suppose we have a set of independent and identically distributed random variables $\tilde{L}^{(1)}, \cdots, \tilde{L}^{(n)} \sim F$ in $(\tilde{\mathcal{L}}, \mathrm{d})$. According to Theorem 3.13 in [19], the metric space $(\tilde{\mathcal{L}}, \mathrm{d})$ admits a unique closed-form Fréchet mean μ_F, which is given by $\mu_F = \exp(\mathbb{E}(\log(\tilde{L})))$. We also define the sample Fréchet mean as $\hat{\mu}_F = \exp(\frac{1}{n} \sum_{i=1}^{n} \log(\tilde{L}^{(i)}))$. The existence and uniqueness of the sample Fréchet mean imply its asymptotic consistency [20].

The Fréchet variance quantifies the spread of a random variable around its Fréchet mean. For $\tilde{L}^{(1)}, \cdots, \tilde{L}^{(n)} \sim F$, we define the population Fréchet variance as $V_F = \mathbb{E}(\mathrm{d}^2(\mu_F, \tilde{L}))$ and the sample Fréchet variance as $\hat{V}_F = \frac{1}{n} \sum_{i=1}^{n} \mathrm{d}^2(\hat{\mu}_F, \tilde{L}^{(i)}) = \frac{1}{n} \sum_{i=1}^{n} |\frac{1}{n} \sum_{j=1}^{n} \log(\tilde{L}^{(j)}) - \log(\tilde{L}^{(i)})|_{\mathrm{F}}^2$.

We establish the Central Limit Theorem (CLT) for the sample Fréchet variance $\hat{V}_F$ in the metric space $(\tilde{\mathcal{L}}, \mathrm{d})$ as follows:

Proposition 3.1 (Fréchet Variance CLT): Under the following assumptions (Assumptions 1-3 in [21]):
1) μ_F and $\hat{\mu}_F$ exist and are unique;
2) For any $l \in \tilde{\mathcal{L}}$, $\delta J(\delta, l) \to 0$ as $\delta \to 0$, where the complexity of the metric space $\tilde{\mathcal{L}}$ is quantified by a bound on the entropy integral for metric δ-balls $B_\delta(l)$:

$$J(\delta, l) = \int_0^1 [1 + \log N\{\epsilon\delta/2, B_\delta(l), \mathrm{d}\}]^{1/2} d\epsilon,$$

where $B_\delta(l)$ is the δ-ball centered at l in the metric d and $N\{\epsilon\delta/2, B_\delta(l), \mathrm{d}\}$ is the covering number for $B_\delta(l)$ using open balls of radius $\epsilon\delta/2$.
3) The entropy integral of the metric space is finite, i.e., $\int_0^1 [1 + \log N(\epsilon, \tilde{\mathcal{L}}, \mathrm{d})]^{1/2} d\epsilon < \infty$.

The following result is obtained:

$$n^{1/2}(\hat{V}_F - V_F) \to N(0, \sigma_F^2) \text{ in distribution,} \qquad (12)$$

where $\sigma_F^2 = \mathrm{Var}\{\mathrm{d}^2(\mu_F, \tilde{L})\}$.

C. Estimating the Location of a Change Point

1) Test Statistics Construction: Let us consider a sequence $\{Y^{(i)}\}_{i=1}^n$ of independent, time-ordered data points in a metric space $(\tilde{\mathcal{L}}, \mathrm{d})$ defined in Section III-B.

We assume there is at most one change point, which we denote by $0 < \tau < 1$. Specifically, $Y^{(1)}, \cdots, Y^{(\lfloor n\tau \rfloor)} \sim F_1$ and $Y^{(\lfloor n\tau \rfloor + 1)}, \cdots, Y^{(n)} \sim F_2$, where F_1 and F_2 are unknown probability measures on $(\tilde{\mathcal{L}}, d)$, and $\lfloor x \rfloor$ is the greatest integer less than or equal to x. Our aim is to test the null hypothesis of distribution homogeneity, denoted by $H_0 : F_1 = F_2$, against the alternative hypothesis of a single change point, denoted by $H_1 : F_1 \neq F_2$.

We constrain the hypothesized change point location τ to a compact interval $\mathcal{I}_c = [c, 1-c] \subset [0, 1]$ with positive constant c to ensure accurate representation of each segment's Fréchet mean and variance. For statistical analysis of segments separated by $u \in \mathcal{I}_c$, we compute the sample Fréchet mean before and after $\lfloor nu \rfloor$ observations as:

$$\hat{\mu}_{[0,u]} = \arg\min_{l \in \tilde{\mathcal{L}}} \frac{1}{\lfloor n\tau \rfloor} \sum_{i=1}^{\lfloor n\tau \rfloor} d^2(Y^{(i)}, l),$$

$$\hat{\mu}_{[u,1]} = \arg\min_{l \in \tilde{\mathcal{L}}} \frac{1}{n - \lfloor n\tau \rfloor} \sum_{i=\lfloor n\tau \rfloor + 1}^{n} d^2(Y^{(i)}, l),$$

and the corresponding sample Fréchet variances are:

$$\hat{V}_{[0,u]} = \frac{1}{\lfloor n\tau \rfloor} \sum_{i=1}^{\lfloor n\tau \rfloor} d^2(Y^{(i)}, \hat{\mu}_{[0,u]}),$$
$$\hat{V}_{[u,1]} = \frac{1}{n - \lfloor n\tau \rfloor} \sum_{i=\lfloor n\tau \rfloor + 1}^{n} d^2(Y^{(i)}, \hat{\mu}_{[u,1]}), \tag{13}$$

One can obtain the contaminated version of Fréchet variances by replacing the Fréchet mean of a segment with the mean of the complementary segment. This leads to the definitions:

$$\hat{V}_{[0,u]}^C = \frac{1}{\lfloor n\tau \rfloor} \sum_{i=1}^{\lfloor n\tau \rfloor} d^2(Y^{(i)}, \hat{\mu}_{[u,1]}),$$

$$\hat{V}_{[u,1]}^C = \frac{1}{n - \lfloor n\tau \rfloor} \sum_{i=\lfloor n\tau \rfloor + 1}^{n} d^2(Y^{(i)}, \hat{\mu}_{[0,u]}),$$

which are guaranteed to be at least as large as the correct version (13). The differences $\hat{V}_1^C - \hat{V}_{[0,u]}$ and $\hat{V}_2^C - \hat{V}_{[u,1]}$ can be interpreted as measures of the between-group variance of the two segments.

For some fixed $u \in \mathcal{I}_c$, the statistic $\sqrt{u(1-u)}(\sqrt{n}/\sigma)(\hat{V}_{[0,u]} - \hat{V}_{[u,1]})$ has an asymptotic standard normal distribution under the null hypothesis H_0. Here, σ denotes the asymptotic variance of the empirical Fréchet variance. This result allows us to test hypotheses about differences in Fréchet variances between two segments of data. A sample based estimator for σ^2 is:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} d^4(\hat{\mu}, \tilde{L}_i) - \left(\frac{1}{n} \sum_{i=1}^{n} d^2(\hat{\mu}, \tilde{L}_i) \right)^2,$$

which is consistent under H_0 [21]. Also,

$$\hat{\mu} = \arg\min_{l \in \tilde{\mathcal{L}}} \frac{1}{n} \sum_{i=1}^{n} d^2(Y^{(i)}, l), \quad \hat{V} = \frac{1}{n} \sum_{i=1}^{n} d^2(Y^{(i)}, \hat{\mu}).$$

The test statistic proposed in [22] can detect differences in both Fréchet means and Fréchet variances of the distributions F_1 and F_2. The equation for the test statistic is provided as follows:

$$T_n(u) = \frac{u(1-u)}{\hat{\sigma}^2} \Big[(\hat{V}_{[0,u]} - \hat{V}_{[u,1]})^2 \\ + (\hat{V}_{[0,u]}^C - \hat{V}_{[0,u]} + \hat{V}_{[u,1]}^C - \hat{V}_{[u,1]})^2 \Big] \tag{14}$$

In this equation, $(\hat{V}_{[0,u]} - \hat{V}_{[u,1]})^2$ indicates the difference in Fréchet variances between two segments of data, while $(\hat{V}_{[0,u]}^C - \hat{V}_{[0,u]} + \hat{V}_{[u,1]}^C - \hat{V}_{[u,1]})^2$ captures the difference in Fréchet means between the two segments.

Under H_0 that no change point exists, we establish the weak convergence[2] of $nT_n(u)$:

Proposition 3.2 (Weak Convergence of $nT_n(u)$):
Under H_0 and the following assumptions (Assumptions 1-4 in [22]):

1) μ_F and $\hat{\mu}_F$ exist and are unique;
2) For any $\alpha = \{\alpha_1, \ldots, \alpha_n : 0 \leq \alpha_i \leq 1, \sum_{i=1}^{n} \alpha_i = 1\}$, $\hat{\mu}_\alpha = \arg\min_{l \in \tilde{\mathcal{L}}} \sum_{i=1}^{n} \alpha_i d^2(Y^{(i)}, l)$ exists and is unique almost surely;
3) For any $l \in \tilde{\mathcal{L}}$, $\int_0^1 \sqrt{\log N(\epsilon\delta, B_\delta(l), d)} d\epsilon = O(1)$ as $\delta \to 0$;
4) There exist $\delta > 0$ and $C > 0$ such that for all $l \in B_\delta(\mu_F)$: $\mathbb{E}(d^2(Y, l)) - \mathbb{E}(d^2(Y, \mu_F)) = Cd^2(l, \mu_F) + O(\delta^2)$ as $\delta \to 0$,

The following result is obtained:

$$\{nT_n(u) : u \in \mathcal{I}_c\} \Rightarrow \{\mathcal{G}^2(u) : u \in \mathcal{I}_c\}, \tag{15}$$

where $\mathcal{G} = \left\{ \frac{\mathcal{B}(u)}{\sqrt{u(1-u)}} : u \in \mathcal{I}_c \right\}$ and $\mathcal{B}(u)$ is a Brownian bridge on $\mathcal{I}_c$, which is a Gaussian process indexed by $\mathcal{I}_c$ with zero mean and covariance structure given by $K(s, t) = \min(s, t) - st$.

To perform a hypothesis test between H_0 and H_1, the statistic $\sup_{u \in \mathcal{I}_c} nT_n(u)$ is used. Here, $T_n(u)$ is a test statistic for hypothesis testing, calculated for each potential change point $u \in \mathcal{I}_c$. The $(1-\alpha)$th quantile of $\sup_{u \in \mathcal{I}_c} \mathcal{G}^2(u)$ is denoted as $q_{1-\alpha}$, which is obtained by a bootstrap approach, as described in Section 3.3 of [22].

Under the null hypothesis H_0:

$$\sup_{u \in \mathcal{I}_c} nT_n(u) \Rightarrow \sup_{u \in \mathcal{I}_c} \mathcal{G}^2(u), \tag{16}$$

where $T_n(u)$ is the test statistic and $\mathcal{G}^2(u)$ is the limiting distribution. We use this result to define the rejection region for a level α significance test as:

$$R_{n,\alpha} = \left\{ \sup_{u \in \mathcal{I}_c} nT_n(u) > q_{1-\alpha} \right\}, \tag{17}$$

where $q_{1-\alpha}$ is the $(1-\alpha)$ quantile of $\mathcal{G}^2(u)$.

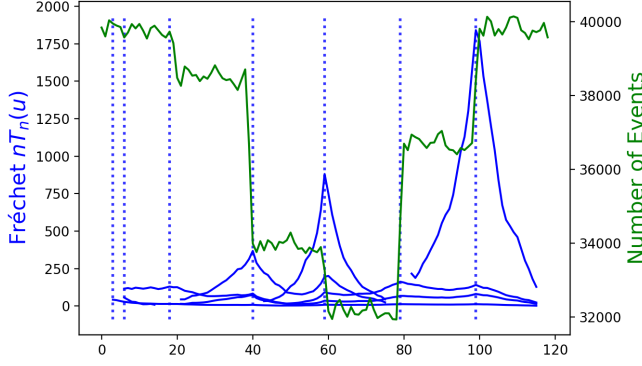[2]Weak convergence is a function space generalization of convergence in distribution [23].

Fig. 1. Our algorithm effectively detected all change points in a synthetic 10-dimensional point process of length 600000 using Fréchet statistics. The data was split into 119 overlapping windows of length 10000, with the number of events in each window shown in green. The Fréchet test statistic $nT_n(u) : u \in \mathcal{I}_c$ was also plotted for each analyzed segment due to the use of binary segmentation. The accuracy of our algorithm was validated by the ground truth, as the peak of the test statistics occurs around the significant changes in the number of events.
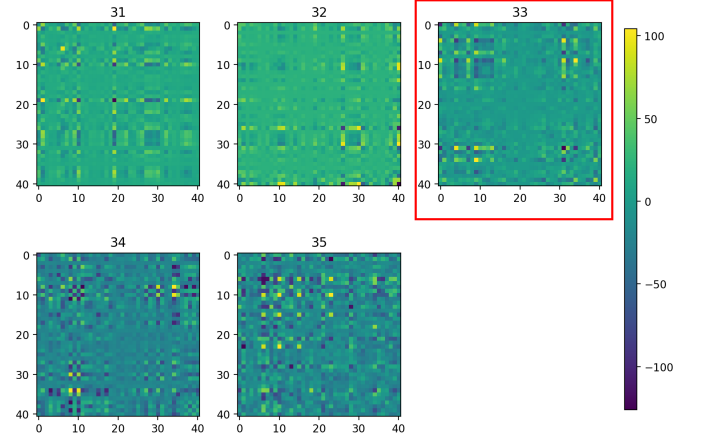

Fig. 2. The sequence of symmetrized kernel matrices $A^{(t)} = (H^{(t)} + H^{(t)T})/2$ estimated by the NPHC algorithm (see Section II-A) is displayed before and after the detected change point at time 33. The change in the underlying causal network dynamics is reflected in the shift of the kernel matrix patterns from the earlier (left) to the later (right) time periods.

Under the alternative hypothesis H_1, which assumes a change point is present at $\tau \in \mathcal{I}_c$, we can locate it by finding the maximizer of the process $T_n(u)$:

$$\hat{\tau} = \arg\max_{u \in \mathcal{I}_c} T_n(u) = \arg\max_{\lfloor nc \rfloor \leq k \leq n - \lfloor nc \rfloor} T_n\left(\frac{k}{n}\right), \quad (18)$$

where $\hat{\tau}$ is the estimated change point that maximizes the test statistic across all potential change points. Furthermore, we propose a binary segmentation procedure that extends the proposed statistic $T_n(u)$ to multiple change point scenario [16].

IV. Empirical Analysis on Simulated and Practical Networks

We evaluate the performance of the proposed method on change point detection on both a simulated dataset and a cryptocurrency market dataset. Our results are completely reproducible; the code and datasets used in the experiments are publicly available at https://tinyurl.com/Hawkes-CPD.

A. Simulation Study

We generated a synthetic dataset using the multivariate Hawkes process model, implemented through the Python library tick[3]. We first set the base intensities λ_i for all 10 subjects to be 0.3. We then defined the kernel function using an exponential decay model, as shown in (3), with decay parameters $\beta_{ij} = 0.5 + u_{ij}$, where u_{ij} is sampled from a uniform distribution between 0 and 1. To set the amplitude of the kernel function, we used the following values: for $i = j$, $\hat{\alpha}_{ij} = 1/8$, and for $i \neq j$, $\hat{\alpha}_{ij}$ was sampled from a normal distribution with mean 0 and variance 1/64. We symmetrized these values to obtain $\alpha_{ij} = (\hat{\alpha}_{ij} + \hat{\alpha}_{ji})/2$. We ensured that the spectral

radius of the resulting kernel matrix H was less than 1, which guarantees the stationarity of the process. The integral of each kernel function is computed as $h_{ij} = \frac{\alpha_{ij}}{\beta_{ij}}$.

We induce change points by updating the α_{ij} every 100000 units, resulting in a modification of the causal network H of the process.

To evaluate the performance of our algorithm in detecting multiple change points, we analyze a simulated multivariate process of length 600000. The data is divided into overlapping windows of length 10000, with adjacent windows having an overlapping window of length 5000. Figure 1 shows that the Fréchet statistics based algorithms accurately detect all the change points, as confirmed by the ground truth generated during dataset construction. We also plot the test statistics, that is, the Fréchet test statistic $nT_n(u) : u \in \mathcal{I}_c$ (14) for our algorithm. Since we use binary segmentation, we display the Fréchet test statistic for each analyzed segment.

B. Cryptocurrency Price Data

We use Binance's cryptocurrency price data[4] for 41 DeFi-related coins. The data was sampled every 5 minutes for 396 days between 1st March 2022 and 31st March 2023. Events were identified whenever a coin price changed by $\pm 0.5\%$ of its current price, resulting in a 41-dimensional multivariate point process. We divided the process into 56 overlapping 2-week windows with 1-week overlap for adjacent windows.

We present the results of our analysis of the cryptocurrency price dataset using a 41-dimensional multivariate Hawkes process model. Figure 3 displays the multiple change points detected in the model, which highlight periods of increased activity and volatility in the cryptocurrency market. To better understand the underlying

---

[3]https://github.com/X-DataInitiative/tick

[4]https://github.com/binance-us/binance-official-api-docs

Start Date	No.	Event
06/06/2022	14	Terra-LUNA contagion.
07/11/2022	19	Voyager and Celsius filed for bankruptcy.
10/17/2022	33	Over US\$718 million stolen from decentralized finance[5].
12/05/2022	40	BlockFi, which had taken a US\$250 million loan from FTX, declared bankruptcy.
01/16/2023	46	Bitcoin price is up 39% since the start of January.

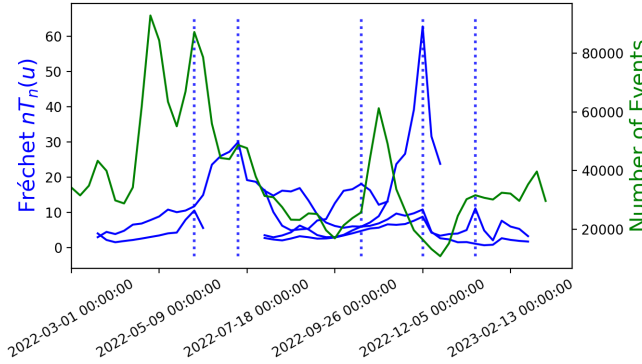TABLE I

Timeline of cryptocurrency market events.

Fig. 3. Multiple change point detection results of the cryptocurrency price dataset. The number of events in each window is displayed in green.

dynamics, we show the adjacency matrices of the causal network of the multivariate Hawkes process before and after one of the detected change points in Figure 2.

To validate the detected change points, we cross-referenced them with significant events in the cryptocurrency market, as summarized in Table 1. Our findings offer valuable insights into the communication patterns and organizational behavior during market transformations or even crashes, and demonstrate the potential of multivariate Hawkes process models in capturing the complex dynamics of financial markets.

V. Conclusions

In this study, we propose a method for detecting change points in causal networks of multivariate Hawkes processes using Fréchet statistics. Specifically, we divide the point process into a sequence of overlapping windows, estimate the kernel matrices in each window, and reconstruct the signed Laplacians by treating the kernel matrices as the graph adjacency matrices of the causal network. We evaluate the method on both simulated and practical datasets, which demonstrate that the proposed method is a powerful tool for detecting and characterizing changes in the causal structure of multivariate Hawkes processes.

References

[1] R. Luo, B. Nettasinghe, and V. Krishnamurthy, "Controlling segregation in social network dynamics as an edge formation game," IEEE transactions on network science and engineering, vol. 9, no. 4, pp. 2317–2329, 2022.

[2] R. Luo and N. Colombo, "Conformal load prediction with transductive graph autoencoders," arXiv preprint arXiv:2406.08281, 2024.

[3] S. W. Linderman, Bayesian Methods for Discovering Structure in Neural Spike Trains, Ph.D. thesis, Harvard University, 2016.

[4] E. Bacry, I. Mastromatteo, and J.-F. Muzy, "Hawkes processes in finance," Market Microstructure and Liquidity, vol. 1, no. 01, pp. 1550005, 2015.

[5] K. Yamin, H. Wang, B. Montreuil, and Y. Xie, "Online detection of supply chain network disruptions using sequential change-point detection for hawkes processes," arXiv preprint arXiv:2211.12091, 2022.

[6] H. Wang, L. Xie, Y. Xie, A. Cuozzo, and S. Mak, "Sequential change-point detection for mutually exciting point processes," Technometrics, pp. 1–13, 2022.

[7] Q. Li, F. Guo, and I. Kim, "A non-parametric bayesian change-point method for recurrent events," Journal of Statistical Computation and Simulation, vol. 90, no. 16, pp. 2929–2948, 2020.

[8] G. Detommaso, H. Hoitzing, T. Cui, and A. Alamir, "Stein variational online changepoint detection with applications to hawkes processes and neural networks," arXiv preprint arXiv:1901.07987, 2019.

[9] S. Linderman and R. Adams, "Discovering latent network structure in point process data," in International conference on machine learning. PMLR, 2014, pp. 1413–1421.

[10] J. Etesami, N. Kiyavash, K. Zhang, and K. Singhal, "Learning network of multivariate hawkes processes: A time series approach," arXiv preprint arXiv:1603.04319, 2016.

[11] M. Achab, E. Bacry, S. Gaïffas, I. Mastromatteo, and J.-F. Muzy, "Uncovering causality from multivariate hawkes integrated cumulants," in International Conference on Machine Learning. PMLR, 2017, pp. 1–10.

[12] S. J. Hardiman and J.-P. Bouchaud, "Branching-ratio approximation for the self-exciting hawkes process," Physical Review E, vol. 90, no. 6, pp. 062807, 2014.

[13] A. G. Hawkes and D. Oakes, "A cluster process representation of a self-exciting process," Journal of applied probability, vol. 11, no. 3, pp. 493–503, 1974.

[14] S. Jovanović, J. Hertz, and S. Rotter, "Cumulants of hawkes point processes," Physical Review E, vol. 91, no. 4, pp. 042802, 2015.

[15] A. G. Hawkes, "Spectra of some self-exciting and mutually exciting point processes," Biometrika, vol. 58, no. 1, pp. 83–90, 1971.

[16] R. Luo and V. Krishnamurthy, "Fréchet-statistics-based change point detection in dynamic social networks," IEEE Transactions on Computational Social Systems, pp. 1–9, 2023.

[17] J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. W. De Luca, and S. Albayrak, "Spectral analysis of signed graphs for clustering, prediction and visualization," in Proceedings of the 2010 SIAM international conference on data mining. SIAM, 2010, pp. 559–570.

[18] N. J. Higham, "Computing the nearest correlation matrix—a problem from finance," IMA journal of Numerical Analysis, vol. 22, no. 3, pp. 329–343, 2002.

[19] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Geometric means in a novel vector space structure on symmetric positive-definite matrices," SIAM journal on matrix analysis and applications, vol. 29, no. 1, pp. 328–347, 2007.

[20] A. Petersen and H.-G. Müller, "Fréchet regression for random objects with Euclidean predictors," The Annals of Statistics, vol. 47, no. 2, pp. 691 – 719, 2019.

[21] P. Dubey and H.-G. Müller, "Fréchet analysis of variance for random objects," Biometrika, vol. 106, no. 4, pp. 803–821, 2019.

[22] P. Dubey and H.-G. Müller, "Fréchet change-point detection," The Annals of Statistics, vol. 48, no. 6, pp. 3312 – 3335, 2020.

[23] P. Billingsley, Convergence of probability measures, John Wiley & Sons, 2013.