

PCA FOR POINT PROCESSES

BY FRANCK PICARD^{1,a}, VINCENT RIVOIRARD^{2,b}, ANGELINA ROCHE^{2,c} AND
VICTOR M. PANARETOS^{3,d}

¹Laboratoire de Biologie et Modélisation de la Cellule, CNRS ENS de Lyon, franck.picard@ens-lyon.fr

²CEREMADE, Université Paris Dauphine, Vincent.Rivoirard@dauphine.fr, roche@ceremade.dauphine.fr

³Institut de Mathématiques, Ecole Polytechnique Fédérale de Lausanne, victor.panaretos@epfl.ch

We introduce a novel statistical framework for the analysis of replicated point processes that allows for the study of point pattern variability at a population level. By treating point process realizations as random measures, we adopt a functional analysis perspective and propose a form of functional Principal Component Analysis (fPCA) for point processes. The originality of our method is to base our analysis on the cumulative mass functions of the random measures, which gives us a direct and interpretable analysis. Key theoretical contributions include establishing a Karhunen–Loève expansion for the random measures and a Mercer theorem for covariance measures. We establish convergence in a strong sense, and introduce the concept of principal measures, which can be seen as latent processes governing the dynamics of the observed point patterns. We propose an easy-to-implement estimation strategy of eigenelements for which parametric rates are achieved. We fully characterize the solutions of our approach to Poisson and Hawkes processes and validate our methodology via simulations and diverse applications in seismology, single-cell biology and neurosciences, demonstrating its versatility and effectiveness. Our method is implemented in the `pppca` R-package.

1. Introduction. Point processes constitute a ubiquitous framework that is essential in probability and statistics as well as in numerous application fields. In the most general sense, point processes are discrete random sets in an arbitrary space, serving as the natural mathematical formalism for discrete random patterns. Depending on the ambient space, point processes can model spatial/temporal events, tessellations of space or random geometrical configurations. Poisson processes are arguably fundamental [41], due to their simplicity, and because they serve as building blocks for more elaborate point process models. Key models include renewal, marked, cluster or doubly stochastic point processes [22, 23], and geometrical aspects have also long been reviewed [16]. When the ambient space is the real line, point processes represent sequences of discrete events in time and are referred to as temporal point processes, among which Hawkes processes [33, 34] are crucial for enhancing the flexibility of modeling temporal events, introducing a tractable and interpretable model of temporal dependency [8, 9]. The Poisson and Hawkes processes are among the most fundamental, popular and analytically/statistically tractable, serving as basic examples, including in the present work. Given their formidable flexibility in handling particular characteristics of the data, the scope of application of point process models is vast, ranging from neurosciences [5, 17, 43] to genomics [12, 30], ecology [60], epidemiology [15], seismology [51], social sciences [20, 28, 49] and stock prices moves [26] to name but a few.

Statistical inference for point processes has a long history [1, 40], with nonparametric approaches being favored for analyzing Poisson [4, 42, 57, 61] and Hawkes processes [3, 13, 31, 58]. Maximum likelihood [6, 17, 51] and the Bayesian framework have also shown

Received May 2024; revised November 2025.

MSC2020 subject classifications. Primary 62M, 62P; secondary 60G55, 62H25.

Key words and phrases. Point processes, Karhunen–Loève expansion, PCA.

good performance [25, 55, 59]. However, current frameworks are primarily based on the observation of a single point process, while increasingly, various fields are providing data in the form of replicated point processes. For example, in the field of earth sciences, point processes are essential to assess the seismic risk, which can have significant implications for civil engineering and insurance, for example. In the seminal work of [51], dependent point processes were proposed to study preseismic quiescence and foreshocks that are expected to precede major earthquakes. This accurate modeling of earthquake hazards appears crucial to anticipate future earthquakes before they occur. Analyzing earthquake occurrences across various sites could help in identifying cities that show unusual earthquake accumulation patterns, or to more clearly identify the characteristic patterns of seismic activity before and after major earthquakes. In this work, we will consider Turkish cities subject to recorded earthquakes over time, investigating the variability of these temporal patterns among cities [39]. In single-cell genomics, it is now possible to individually characterize genomes based on epigenetic marks organized along chromosomes in a one-dimensional array [64]. Furthermore, these repeated point patterns are instrumental in characterizing the diversity of cancer cells and their emergence of drug resistance. In neuroscience, observing spiking neurons has become routine for investigating whether their coordinated activity relates to the function of specific brain regions [21]. What unifies these examples is that replicated point processes offer an opportunity to study the variability of point patterns from a new perspective, characterizing data at a population level. This poses new mathematical, statistical and methodological challenges, as most standard statistical frameworks are dedicated to continuous valued data and not to sets of points.

In this work, we propose to put forward a different statistical perspective on point processes by developing a new framework for dimension reduction and visualization of replicated point processes. Our framework is based on a functional analysis perspective—to be contrasted to the perspective based on random sets. Viewing the realization of a point process as a random measure, and thus a bounded linear functional over an appropriately regular space of functions, we obtain uniform series representations and associated notions of principal modes of variation. This allows us to develop a Principal Component Analysis framework paralleling that of functional PCA (fPCA). fPCA is the workhorse of functional data analysis [45]. It enables the visualization of functional data, seen as random curves over compact sets (as described in [54]), enables dimensionality reduction and regularization in regression/testing/classification [24, 35, 56] and provides a bridge to extend classical multivariate procedures (like model based clustering) to functional data [27, 38]. The theoretical underpinnings of fPCA rely on two fundamental results: the Karhunen–Loève theorem and Mercer’s theorem [36]. In particular, the Karhunen–Loève theorem asserts that the approximation of functional data by their principal components is uniformly convergent (in the mean-square sense). However, the theory of fPCA does not apply to point processes, which are more intricate mathematical objects.

The link between point processes and functional data is a natural approach to define equivalent of PCA for samples of point processes, and has been attempted before, albeit via the intensity function. Wu et al. [62] investigated PCA for doubly stochastic point processes in the temporal setting, while [44] extended the analysis to spatiotemporal processes. Bouzas et al. [7] also applied functional PCA to estimate the mean of a doubly stochastic Poisson process. These methods rely on kernel or projection estimators of the process intensities, leading to a Karhunen–Loève decomposition of the corresponding intensity functions. An alternative line of work, exemplified by [37] in the case of spatial processes, focuses on performing PCA on estimators of L -functions. However these approaches have notable drawbacks. First, instead of relying on the observed point process, they relate to an intensity (heuristically, a derivative) that is not observable. Furthermore, and most importantly, they do not allow

for a Karhunen–Loève expansion for the measures associated to the initial process; this is a theoretical but also practical drawback, in terms of interpretation. Other approaches include [46] who proposed a convenient Hilbertian setting for PCA of measures for the analysis of distribution functions appearing in the study of grain-size curve. Although the principle of the method is similar to ours, the setting is very different since the data is a sample of general random densities whose cumulative distribution function is observed on a fixed grid, rather than a sample of point processes. Closer in spirit to our approach, [11] recently obtained expansions of general random measures. More precisely, Theorem 4.1 of [11] proves that under mild assumptions, a finite regular random measure over $\mathbb{R}^d$ can be written as a series expansion of deterministic real finite measures weighted by uncorrelated real random variables with summable variances. The convergence of the series is in a weak sense, though, and hence does not constitute a Karhunen–Loève expansion. Consequently, the result is not used (or directly usable, for that matter) for the purposes of statistical analysis.

To describe our contributions, we consider n i.i.d. point processes $(N_1, \dots, N_n)$, and their associated random measures, $(\Pi_1, \dots, \Pi_n)$. We start with Theorem 3.2, which establishes the Karhunen–Loève expansion for the random measures Π_i , as well as a Mercer theorem for the covariance measure of the process (see Theorem 3.4). We establish convergence in a strong sense. Utilizing the series expansion provided by the Karhunen–Loève decomposition, we introduce the concept of principal measures, which can be seen as latent processes governing the dynamics of the observed measures. These results are based on the embedding of the measures Π_i in a functional space using their associated *cumulative mass functions* defined in equation (2). This embedding allows us to integrate our model within the functional PCA framework. We then construct a natural associated covariance operator whose eigenfunctions are the cumulative mass functions of the principal measures. Eigenelements are characterized for Poisson processes and Hawkes processes with exponential self-exciting function by solving second-order differential systems provided in equation (14). When the Poisson process is homogeneous, the eigenfunctions of the covariance operator consist of a Fourier-like basis, and we also show that the eigenvalues exhibit a polynomial decay. When the Poisson process is inhomogeneous, the eigenfunctions are implicit; however, a detailed qualitative analysis allows us to show that the eigenfunctions are oscillatory. Surprisingly, we demonstrate that similar phenomena occur in the case of the Hawkes process. We propose a full estimation framework based on the observation of the N_i 's that does not require any smoothing, and provide convergence of estimators at the parametric rate. Our theoretical results are complemented by a simulation study that exemplifies the convergence of eigenvalues and eigenfunctions. We then apply our framework in seismology, genomics and neuroscience, illustrating the versatility and power of our methodology. We also provide a R-package, `pppca`,¹ which implements our method.

The article is organized as follows. Section 2 introduces the setting of our work. Section 3 establishes Karhunen–Loève and Mercer theorems for point processes. In Section 4, we study the Poisson process and a specific class of Hawkes processes in detail by using the analytical point of view. Section 5 introduces our estimators of the eigenelements and establishes rates of convergence. Our numerical study is carried out in Section 6. Section 7 is devoted to applications. The proofs of our results are collected in the Supplementary Material [53].

Notation. We introduce some notation that will be useful throughout the article. The set $\mathbb{N}$ denotes the set of nonnegative integers. For a set $I \subset \mathbb{R}^d$, for $d \geq 1$, we denote

$$\mathbb{L}^2(I) = \{f : I \mapsto \mathbb{R} : \|f\| < +\infty\},$$

¹<https://github.com/franckpicard/pppca>.

where $\|f\|$ stands for the $\mathbb{L}_2$ -norm of f : $\|f\| = (\int_I |f(t)|^2 dt)^{1/2}$. The associated scalar product is denoted $\langle f, g \rangle = \int_I f(t)g(t) dt$. When $I = (0, 1)$ and $\alpha \in \mathbb{N}$, we set $\partial_\alpha f = f^{(\alpha)}$. When $I = (0, 1)^2$ and $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}^2$, we set $|\alpha| = \alpha_1 + \alpha_2$ and $\partial_\alpha f = \frac{\partial^{|\alpha|} f}{\partial s^{\alpha_1} \partial t^{\alpha_2}}$. The derivatives and partial derivatives are taken in the distribution sense meaning that $\partial_\alpha f$ verifies the equality

$$\langle \partial_\alpha f, \varphi \rangle = (-1)^{|\alpha|} \langle f, \partial_\alpha \varphi \rangle,$$

for all infinitely differentiable functions φ with support in I .

2. A signed-measure model for dimension reduction. Consider n independent and identically distributed (i.i.d.) temporal point processes $(N_1, \dots, N_n)$, observed on the time interval $[0, 1]$. These processes, defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, are associated with some random measures $(\Pi_1, \dots, \Pi_n)$ so that

$$\Pi_i(B) = \sum_{T \in N_i} \mathbf{1}_{\{T \in B\}} = \text{Card}(N_i \cap B), \quad B \in \mathcal{B},$$

where $\mathcal{B}$ is the Borel σ -field of subsets of $[0, 1]$. Hence, $\Pi_1, \dots, \Pi_n$ is a sample of random measures on $([0, 1], \mathcal{B})$. We denote by Π a random measure following the same distribution as the Π_i 's and by N the point process associated with Π . Assuming that $\mathbb{E}[\Pi([0, 1])] < +\infty$, we denote the first moment of Π by W :

$$W(B) = \mathbb{E}[\Pi(B)], \quad B \in \mathcal{B}.$$

In the initial step of our dimension reduction framework, we center the observed data by defining a signed random measure

$$\Delta_i(B) = \Pi_i(B) - W(B), \quad B \in \mathcal{B}.$$

The Δ_i 's have the same distribution as $\Delta = \Pi - W$. Then we define a second-order moment for Δ , inspired from the formalism proposed by [52] and [11]. For this purpose, we define the signed measure of covariance: for all $B \times B' \in \mathcal{B} \otimes \mathcal{B}$,

$$\begin{aligned} C_\Delta(B \times B') &:= \text{Cov}(\Pi(B), \Pi(B')) = \mathbb{E}[\Delta(B)\Delta(B')] \\ &= \mathbb{E}\left[\sum_{T \in N, T' \in N} \mathbf{1}_{\{(T, T') \in B \times B'\}}\right] - (W \otimes W)(B \times B') \end{aligned}$$

with $\mathcal{B} \otimes \mathcal{B}$ the product σ -algebra of $[0, 1] \times [0, 1]$, and where, for two measures μ and ν on $([0, 1], \mathcal{B})$ the product measure $\mu \otimes \nu$ on $([0, 1]^2, \mathcal{B} \otimes \mathcal{B})$ is defined such that

$$(\mu \otimes \nu)(B \times B') = \mu(B)\nu(B'), \quad B, B' \in \mathcal{B}.$$

We assume throughout that

$$(1) \quad \mathbb{E}[\Pi^2([0, 1])] < \infty,$$

which implies that C_Δ is a finite measure.

Now, to perform dimension reduction of the N_i 's (equivalently the Π_i 's) we shall take advantage of the rich framework of the functional perspective. Following [18], for instance, given a random measure μ , we consider the associated *cumulative mass function* defined by

$$(2) \quad F_\mu(t) = \mu([0, t]), \quad t \in [0, 1].$$

Then we introduce K_Δ , the bivariate cumulative mass function of the covariance measure C_Δ , such that

$$K_\Delta(s, t) = C_\Delta([0, s] \times [0, t]) = \mathbb{E}[F_\Delta(s)F_\Delta(t)], \quad s, t \in [0, 1]$$

and the associated integral operator

$$\Gamma_{\Delta} = \mathbb{E}(F_{\Delta} \otimes F_{\Delta}),$$

where $\otimes$ is the tensor product between two elements of $\mathbb{L}^2([0, 1])$ denoted by $f \otimes g(h) = \langle f, h \rangle g$. The covariance operator Γ_{Δ} is expressed as

$$\begin{aligned} \Gamma_{\Delta}(f)(\cdot) &= \mathbb{E}(\langle f, F_{\Delta} \rangle F_{\Delta}(\cdot)), \quad f \in \mathbb{L}^2, \\ &= \int_0^1 K_{\Delta}(\cdot, t) f(t) dt. \end{aligned}$$

Then dimension reduction is based on the Mercer expansion of the covariance kernel K_{Δ} . Indeed, in the case where K_{Δ} is continuous, Mercer's theorem applies and ensures the existence of a sequence of nonnegative real numbers $(\lambda_j)_{j \geq 1} \in \ell^1(\mathbb{N}^*)$ and orthonormal functional basis $(\eta_j)_{j \geq 1}$ of $\mathbb{L}^2([0, 1])$ such that

$$(3) \quad K_{\Delta}(s, t) = \sum_{j \geq 1} \lambda_j \eta_j(s) \eta_j(t), \quad s, t \in [0, 1],$$

with uniform and absolute convergence of the series (see Theorem 4.6.5 of [36]). Furthermore, η_j is an eigenfunction of the operator Γ_{Δ} associated with the eigenvalue λ_j .

Now the question is the following. How can we use the decomposition (3) to build a series expansion for Π and C_{Δ} ? This issue is tackled in the next section.

3. Karhunen–Loève expansion and Mercer theorem for point processes. This section is devoted to prove a Karhunen–Loève theorem for the random measure Π and a Mercer theorem for the covariance measure C_{Δ} . The obtained expansions will be based on specific sequences of measures $(\mu_j)_j$, called *principal measures*.

We start from the decomposition (3). Since Γ_{Δ} is a compact self-adjoint operator, the eigenvalues $(\lambda_j)_{j \geq 1}$ are isolated, with finite multiplicities and we have

$$\lambda_j \geq \lambda_{j+1} \geq 0, \quad \forall j \geq 1,$$

with $\lambda_j \rightarrow 0$ when $j \rightarrow +\infty$. We refer the reader to Sections 6.3 and 6.4 of [10] for properties of the spectrum of a compact operator. In the sequel, we assume that in (3),

$$(4) \quad \lambda_j > 0, \quad j = 1, 2, \dots$$

The following proposition is a central result that allows us to define our principal measures from the basis functions $(\eta_j)_{j \geq 1}$.

PROPOSITION 3.1. *We suppose that Assumptions (1) and (4) are satisfied and that K_{Δ} is continuous. Then, for all $j \geq 1$, the derivative in the distributional sense of η_j is a measure, denoted by μ_j , that verifies*

$$\eta_j(t) = \mu_j([0, t]) = F_{\mu_j}(t), \quad t \in [0, 1]$$

and, for all $\varphi \in \mathcal{H}_0^1 = \{f \in \mathbb{L}^2(I) : f' \in \mathbb{L}^2(I) \text{ and } f(t) = 0 \text{ for all } t \notin I\}$,

$$(5) \quad \langle \eta_j, \varphi' \rangle = \langle F_{\mu_j}, \varphi' \rangle = - \int_0^1 \varphi(s) d\mu_j(s) = - \langle \varphi, \mu_j \rangle.$$

Now, using the μ_j 's derived in Proposition 3.1 and following the proof of Theorem 4.1 of [11], we obtain that for all measurable and bounded function φ on $[0, 1]$,

$$\sum_{T \in N} \varphi(T) = \mathbb{E} \left[\sum_{T \in N} \varphi(T) \right] + \sum_{j \geq 1} \sqrt{\lambda_j} \xi_j \langle \mu_j, \varphi \rangle,$$

where $\{\xi_j\}_{j \geq 1}$ is a sequence of uncorrelated real random variables of mean zero and variance one. This result allows us to interpret the variations of $\sum_{T \in N} \varphi(T)$ around its mean as a sum of random variables with decreasing variance. It can also be expressed as follows:

$$(6) \quad \Pi(B) = W(B) + \sum_{j \geq 1} \sqrt{\lambda_j} \xi_j \mu_j(B), \quad B \in \mathcal{B},$$

which implies for $t \in [0, 1]$,

$$(7) \quad F_\Pi(t) = \Pi([0, t]) = W([0, t]) + \sum_{j \geq 1} \sqrt{\lambda_j} \xi_j \mu_j([0, t]) = F_W(t) + \sum_{j \geq 1} \sqrt{\lambda_j} \xi_j F_{\mu_j}(t).$$

The last equation means in particular that, under Assumption (4), the random sequence $(\xi_j)_{j \geq 1}$ is uniquely defined by the following equation:

$$(8) \quad \xi_j = \frac{\langle F_\Pi - F_W, F_{\mu_j} \rangle}{\sqrt{\lambda_j}}, \quad j \geq 1.$$

To deduce the last expression, we have used that the functions $(F_{\mu_j})_{j \geq 1}$ are orthonormal. The convergence of the series holds then pointwise, that is, for all $B \in \mathcal{B}$ (or for all measurable and bounded function φ). However, the strength of the Karhunen–Loève theorem for general stochastic processes is that the convergence holds uniformly on $[0, 1]$, which prevents problems on the boundaries (typically Gibbs effect type phenomena). This is not ensured by Theorem 4.1 of [11]. Therefore, we first establish a Karhunen–Loève decomposition for point processes for which the convergence holds uniformly for all sufficiently regular functions. Uniformity, however, cannot be in the total variation distance: as illustrated in Section 4 for Poisson and Hawkes processes, the measures $(\mu_j)_j$ are signed measures with very regular densities (typically $\mathcal{C}^\infty$). On the contrary, the measure Π is discrete. This implies that the convergence cannot hold uniformly for all continuous and bounded functions, that is to say, uniformly in total variation. For this reason, in the sequel, uniform convergence will be related to Sobolev spaces instead. For this purpose, we introduce, for $I = (0, 1)$ or $I = (0, 1)^2$,

$$\mathcal{H}_0^k = \{f \in \mathbb{L}^2(I) : \partial_\alpha f \in \mathbb{L}^2(I) \text{ for all } |\alpha| \leq k \text{ and } f(t) = 0 \text{ for all } t \notin I\}.$$

The derivatives and partial derivatives are taken in the distributional sense. Sobolev norms of negative order are also introduced as the norm of the dual space of $\mathcal{H}_0^k$, namely

$$\|\mu\|_{\mathcal{H}^{-k}} = \sup \left\{ |\langle f, \mu \rangle| : f \in \mathcal{H}_0^k \text{ and } \sum_{|\alpha| \leq k} \|\partial_\alpha f\|^2 \leq 1 \right\}.$$

The definition and properties of these spaces are described in Brezis [10], Chapters 8, 9. With these definitions in place, we may now state the following.

THEOREM 3.2 (Karhunen–Loève theorem for point processes). *Suppose that Assumptions (1) and (4) are satisfied and that K_Δ is continuous. Then there exists a sequence $\{\xi_j\}_{j \geq 1}$ of uncorrelated real random variables of mean zero and variance one such that*

$$\lim_{J \rightarrow +\infty} \mathbb{E} \left[\left\| \Pi - W - \sum_{j=1}^J \sqrt{\lambda_j} \xi_j \mu_j \right\|_{\mathcal{H}^{-1}}^2 \right] = 0.$$

REMARK 3.3. Using the fact that, for any Borel function φ ,

$$\langle \varphi, \Pi - W \rangle = \sum_{T \in N} \varphi(T) - \mathbb{E} \left[\sum_{T \in N} \varphi(T) \right]$$

a direct consequence of Theorem 3.2 is that

$$\lim_{J \rightarrow +\infty} \sup_{\varphi \in \mathcal{H}_0^1, \|\varphi\|^2 + \|\varphi'\|^2 \leq 1} \mathbb{E} \left[\left(\sum_{T \in N} \varphi(T) - \mathbb{E} \left[\sum_{T \in N} \varphi(T) \right] - \sum_{j=1}^J \sqrt{\lambda_j} \xi_j \langle \mu_j, \varphi \rangle \right)^2 \right] = 0.$$

By analogy with fPCA, the elements of the sequence $\{\xi_j\}_{j \geq 1}$ are called *the scores associated to the measure Π (or associated to the point process N)*. They play a central role in the study of the data we consider (see Section 7). Indeed, for each individual i , it can be deduced from the Karhunen–Loève decomposition that for J large, Π_i is close to $W + \sum_{j=1}^J \sqrt{\lambda_j} \xi_{i,j} \mu_j$, which can be written informally as

$$(9) \quad \Pi_i(B) \approx W(B) + \sum_{j=1}^J \sqrt{\lambda_j} \xi_{i,j} \mu_j(B), \quad B \in \mathcal{B}$$

or

$$(10) \quad \Pi_i([0, t]) \approx W([0, t]) + \sum_{j=1}^J \sqrt{\lambda_j} \xi_{i,j} F_{\mu_j}(t), \quad t \geq 0$$

with

$$(11) \quad \xi_{i,j} = \frac{\langle F_{\Delta_i}, F_{\mu_j} \rangle}{\sqrt{\lambda_j}},$$

by using (8) and $F_{\Delta_i} = F_{\Pi_i} - F_W$. As to what concerns the covariance measure C_Δ , we establish the following Mercer theorem.

THEOREM 3.4 (Mercer's theorem for C_Δ). *Suppose that Assumptions (1) and (4) are satisfied and that K_Δ is continuous. Then*

$$(12) \quad \left\| \sum_{j=1}^J \lambda_j \mu_j \otimes \mu_j - C_\Delta \right\|_{\mathcal{H}^{-2}} \xrightarrow{J \rightarrow +\infty} 0.$$

In the next sections, we focus on the study of the eigenelements $(\lambda_j, F_{\mu_j})_{j \geq 1}$. Recall that we have a direct correspondence between the measures μ_j and the mass functions F_{μ_j} thanks to equation (2). In Section 4, we study Poisson processes and specific classes of Hawkes processes from the analytical point of view. For these central examples of point processes, we derive the asymptotic behavior of (λ_j, F_{μ_j}) when j goes to $+\infty$. We even obtain the explicit form of the eigenelements for the homogeneous Poisson process for all j 's. In Section 5, for the general setting, we provide an estimation scheme for the eigenelements based on realizations of $(N_1, N_2, \dots, N_n)$.

4. Study of eigenelements for Poisson and Hawkes processes. The goal of this section is to complement the general results of Section 3 with some concrete and explicit cases by adopting a constructive point of view to determine how eigenelements (λ_j, F_{μ_j}) behave, at least asymptotically when j goes to $+\infty$. We consider two classical point process models, namely Poisson processes and Hawkes processes with exponential self-exciting functions.

To present our results, we introduce, for $j \geq 1$,

$$(13) \quad H_j(t) = \int_t^1 F_{\mu_j}(s) ds, \quad t \in [0, 1]$$

and consider solutions (λ, y) of the following system of second-order integro-differential equations:

$$(14) \quad \begin{cases} -\lambda y''(t) = w(t)y(t) + \int_0^1 M(s, t)y(s) ds, & t \in (0, 1), \\ y(1) = 0, & y'(0) = 0, \end{cases}$$

where $\lambda > 0$ and y is twice continuously differentiable and:

1. for the Poisson process, w is the intensity of the process and $M = 0$,
2. for the Hawkes process with exponential self-exciting function, w is a constant proportional to the baseline intensity and M is an exponential convolution kernel (see equation (21)).

We shall prove that the eigenelements $(\lambda_j, F_{\mu_j})_{j \geq 1}$ of Γ_Δ correspond to solutions $(\lambda_j, H_j)_{j \geq 1}$ of System (14), once H_j is renormalized by a constant so that $\|F_{\mu_j}\|_2 = 1$ (observe that for any constant $c > 0$, (λ, y) is a solution of (14) if and only if $(\lambda, c \times y)$ is also a solution). This renormalization is required since $(F_{\mu_j})_{j \geq 1} = (\eta_j)_{j \geq 1}$ is an orthonormal functional basis of $\mathbb{L}^2([0, 1])$.

Relationships connecting Karhunen–Loève expansions to differential equations of order 2 have already been established in the study of stochastic processes. See, for instance, Section A of [19] for the case of the Ornstein–Uhlenbeck process. The system of equations (14) appears naturally when we study the equation

$$\lambda_j F_{\mu_j}(t) = \int_0^1 K_\Delta(t, s) F_{\mu_j}(s) ds, \quad t \in [0, 1]$$

by taking derivatives of left- and right-hand sides. See, for instance, the straightforward proof of Theorem 4.1 in the Supplementary Material. In the sequel, except in Remarks 4.2 and 4.9, we assume that the λ_j 's are positive.

4.1. Eigendecomposition for Poisson point processes. Here we suppose that N is a Poisson process with intensity $t \in [0, 1] \mapsto w(t)$ assumed to be continuous and positive on $(0, 1)$. In this case,

$$W(B) = \int_B w(t) dt, \quad B \in \mathcal{B}$$

and $\Pi(B)$ is a Poisson variable with parameter $W(B)$. Furthermore, for two disjoint Borel sets B and B' , $\Pi(B)$ and $\Pi(B')$ are independent so that

$$F_W(t) = W([0, t]) = \int_0^t w(u) du, \quad 0 \leq t \leq 1$$

and

$$K_\Delta(s, t) = F_W(\min\{s, t\}) = \int_0^{\min\{s, t\}} w(u) du, \quad 0 \leq s, t \leq 1.$$

The following result, which involves the functions H_j introduced in (13), provides the PCA dimension reduction for Poisson processes.

THEOREM 4.1. *Assume that w is continuous and positive on $(0, 1)$. Then, $(\lambda_j, F_{\mu_j})_{j \geq 1}$ are the eigenelements of the operator Γ_Δ if and only if $(\lambda_j, H_j)_{j \geq 1}$ are solutions of the following Sturm–Liouville problem:*

$$(15) \quad \begin{cases} -\lambda y''(t) = w(t)y(t), & t \in (0, 1), \\ y(1) = 0, & y'(0) = 0. \end{cases}$$

Mercer and Karhunen–Loève expansions for Poisson processes are then provided by solutions of System (15) and by using the relationship

$$F_{\mu_j}(t) = -H'_j(t), \quad t \in [0, 1].$$

REMARK 4.2. The proof of Theorem 4.1 shows that if F_{μ_0} is a continuous eigenfunction of Γ associated with the zero eigenvalue, then F_{μ_0} is solution of (15) with $\lambda = 0$ and then $F_{\mu_0} = 0$.

Unfortunately, solutions of System (15) are typically not explicit. However, the specific case where w is constant, corresponding to the homogeneous Poisson process, is of interest and can be easily dealt with.

4.1.1. *Homogeneous Poisson processes.* In the case of homogeneous Poisson process, the function w is constant on $[0, 1]$: for some $w_0 \in \mathbb{R}_+^*$,

$$w(t) = w_0, \quad t \in [0, 1]$$

and the solutions of equation (15) can be obtained explicitly

$$\lambda_j = \frac{4w_0}{\pi^2(2j-1)^2}, \quad H_j(t) = A_j \cos(\pi(2j-1)t/2), \quad t \in [0, 1], j \geq 1,$$

where, in the last expression, A_j can be any constant independent of t . But, taking into account the constraint $\|H'_j\| = \|F_{\mu_j}\| = 1$, we must have $|A_j| = \sqrt{2}/(\pi j)$. Without loss of generality, we take $A_j = -\sqrt{2}/(\pi j)$ so that

$$(16) \quad F_{\mu_j}(t) = \sqrt{2} \sin(\pi(2j-1)t/2), \quad t \in [0, 1], j \geq 1.$$

In particular, the eigenfunctions F_{μ_j} are highly oscillating functions with exactly j zeros on $[0, 1]$ at points $x_k = 2k/(2j-1)$, $k = 0, \dots, j-1$.

Then we can finally write that if N is a homogeneous Poisson process with intensity w_0 ,

$$\Pi([0, t]) = w_0 t + \frac{2\sqrt{2}w_0}{\pi} \sum_{j \geq 1} \frac{\xi_j}{2j-1} \sin(\pi(2j-1)t/2), \quad t \in [0, 1],$$

with $\{\xi_j\}_{j \geq 1}$ uncorrelated centered real random variables of unit variance.

REMARK 4.3. Observe that in the homogenous Poisson setting, $K_\Delta(s, t) = w_0 \min\{s, t\}$, which corresponds to the covariance kernel of Brownian motion, so that the Karhunen–Loève decomposition of Π can be deduced from the one of Brownian motion (see, e.g., [2], pp. 41–42). The main difference with Brownian motion is that the scores $\{\xi_j\}_{j \geq 1}$ are not Gaussian. Determining the distribution of the scores in the Poisson setting remains an open and interesting problem. It is beyond the scope of this paper.

4.1.2. *Inhomogeneous Poisson processes.* The general Poisson case is much more involved, since when w is not constant, we cannot derive explicit solutions of (15), in full generality. But we prove in this section that solutions of (15) exist and some useful qualitative properties of them can be obtained. Actually, system (15) is a self-adjoint Sturm–Liouville problem with separated boundary conditions, which has been extensively considered in the literature. We refer the reader to Chapter 4 of [63] and more precisely to equations (4.1.1)–(4.1.4) with

$$p \equiv 1, \quad q \equiv 0, \quad \lambda^{-1} \text{ instead of } \lambda.$$

Therefore, we have the following result [63].

THEOREM 4.4 (Theorems 4.3.1 and 4.6.2 of [63]). *We consider the solutions $(\lambda_j, H_j)_j$ of the Sturm–Liouville problem (15) with $w \in \mathbb{L}_1[0, 1]$ and assume that w is positive on $(0, 1)$. We have:*

1. *The eigenvalues $(\lambda_j)_{j \geq 1}$ are simple and asymptotically, we have*

$$\lambda_j \underset{j \rightarrow +\infty}{\sim} \frac{(\int_0^1 \sqrt{w(u)} du)^2}{(j\pi - \pi/2)^2}.$$

2. *For $j \geq 1$, if H_j is the solution associated with λ_j , then H_j is unique up to a multiplicative constant and has exactly $j - 1$ zeros on $(0, 1)$.*

3. *The sequence $(H_j)_{j \geq 1}$ can be normalized to be an orthonormal sequence in $\mathbb{L}_2([0, 1], w)$, that is,*

$$\int_0^1 H_j(u) H_{j'}(u) w(u) du = \begin{cases} 0 & \text{if } j \neq j', \\ 1 & \text{if } j = j'. \end{cases}$$

REMARK 4.5. By using the third point of Theorem 4.4, we observe that for $j \neq j'$,

$$\begin{aligned} \int_0^1 F_{\mu_j}(u) F_{\mu_{j'}}(u) du &= \int_0^1 H_j'(u) H_{j'}'(u) du \\ &= H_j'(1) H_{j'}'(1) - H_j'(0) H_{j'}'(0) - \int_0^1 H_j''(u) H_{j'}'(u) du \\ &= \lambda_j^{-1} \int_0^1 H_j(u) H_{j'}(u) w(u) du = 0, \end{aligned}$$

so that, once conveniently normalized, $(F_{\mu_j})_{j \geq 1}$ is an $\mathbb{L}_2$ -orthonormal basis, as required by our theory.

Unlike the homogeneous case, we do not obtain the exact values of the eigenvalues λ_j for all j 's, but the first point of Theorem 4.4 provides their asymptotic behavior when $j \rightarrow +\infty$. The numerical study of Section 6 goes further since it shows that, for even small values of j , λ_j is close to $(\int_0^1 \sqrt{w(u)} du)^2 / (j\pi - \pi/2)^2$, exactly as for the homogeneous case (see Figure 2).

Combining the second point of Theorem 4.4 and the boundary condition $H_j(1) = 0$, we obtain that H_j has exactly j zeros on $(0, 1]$. Since $F_{\mu_j}(t) = -H_j'(t)$ for $t \in [0, 1]$ and $F_{\mu_j}(0) = 0$, Rolle's theorem yields that F_{μ_j} has at least j zeros on $[0, 1)$ and is therefore a highly oscillating functions, exactly as for the homogeneous case for which the functions F_{μ_j} are sine functions (see (16)). These qualitative properties are confirmed by the numerical study of Section 6 (see Figure 1). The numerical examples considered show that the F_{μ_j} 's have exactly j zeros on $[0, 1)$.

To conclude, when w is not constant, we cannot give a closed expression of eigenelements of the operator Γ_Δ , but through the study of their behavior provided by the Sturm–Liouville theory, we have established that they can be viewed as perturbations of the eigenelements of the homogeneous Poisson case.

4.2. Eigendecomposition for Hawkes processes with exponential self-exciting functions.

We extend our analytical results by providing results on dimension reduction for (linear) Hawkes processes. Hawkes processes were designed by Alan G. Hawkes [33, 34] to model the properties of self-excitation. This model seems particularly natural to account for dependencies in our point process model. Hawkes processes can be viewed as a branching process

over a homogeneous Poisson process but they are also characterized by their stochastic intensity $w(t)$ defined by

$$w(t) = \lim_{\delta \rightarrow 0} \delta^{-1} \mathbb{E}[\Pi[0, t + \delta] - \Pi[0, t] | \mathcal{F}_t],$$

where the filtration $\mathcal{F}_t$ stands for the information available up to (but not including) time t ; $w(t)$ can also be interpreted as the probability to have a new occurrence at time t given the past.

DEFINITION 4.6. The (linear) Hawkes process with baseline intensity $w_0 > 0$ and self-exciting function $h : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is a point process N with intensity

$$(17) \quad w(t) = w_0 + \int_{-\infty}^{t-} h(t-s) dN_s = w_0 + \sum_{T \in N, T < t} h(t-T).$$

Observe that the larger h , the stronger the influence of the past. If $h = 0$, the associated Hawkes process corresponds to a homogeneous Poisson process whose intensity is the baseline intensity w_0 . In the sequel, we mainly consider an exponential self-exciting function, that is,

$$(18) \quad h(t) = \alpha \exp(-\beta t), \quad t > 0,$$

and in this case, the intensity becomes

$$w(t) = w_0 + \alpha \int_{-\infty}^{t-} \exp(-\beta(t-s)) dN_s = w_0 + \alpha \sum_{T \in N, T < t} \exp(-\beta(t-T)),$$

meaning that the influence of the past decreases exponentially in time. The intensity associated with an exponential self-exciting function becomes a Markov process [50].

In the sequel, we assume that $\|h\|_1 < 1$. For the case of exponential self-exciting function h given in (18), this condition is equivalent to $\alpha < \beta$. Assuming $\|h\|_1 < 1$ ensures that there exists a unique stationary version of the Hawkes process N with intensity given by (17) (see Theorem 1 of [8]). Furthermore, under this assumption, the mean and the covariance kernel of N are well-defined and if we consider the exponential case, we obtain explicit expressions.

PROPOSITION 4.7. We assume that for $t > 0$, $h(t) = \alpha e^{-\beta t}$, with $0 < \alpha < \beta$. Then

$$(19) \quad F_W(t) = W([0, t]) = \frac{w_0 \beta t}{\beta - \alpha}, \quad t > 0,$$

and for $0 < s \leq t$,

$$(20) \quad K_\Delta(s, t) = \frac{\beta^3 w_0}{(\beta - \alpha)^3} s + \frac{w_0 \alpha \beta (2\beta - \alpha)}{2(\beta - \alpha)^4} (-1 - e^{(\alpha - \beta)(t-s)} + e^{(\alpha - \beta)t} + e^{(\alpha - \beta)s}).$$

The proof of Proposition 4.7 can easily be deduced from results in Section 3 of [29]. Subsequently, using the functions H_j introduced in (13), the following result provides the PCA decomposition for Hawkes processes with exponential self-exciting function when studied on the compact interval $[0, 1]$.

THEOREM 4.8. We assume that for $t > 0$, $h(t) = \alpha e^{-\beta t}$, with $0 < \alpha < \beta$. Then $(\lambda_j, F_{\mu_j})_{j \geq 1}$ are the eigenelements of the operator Γ_Δ if and only if $(\lambda_j, H_j)_{j \geq 1}$ are solutions of the following system:

$$(21) \quad \begin{cases} -\lambda y''(t) = \frac{\beta w_0}{\beta - \alpha} y(t) + \frac{w_0 \alpha \beta (2\beta - \alpha)}{2(\beta - \alpha)^2} \int_0^1 e^{-(\beta - \alpha)|t-s|} y(s) ds, & t \in (0, 1), \\ y(1) = 0, & y'(0) = 0. \end{cases}$$

REMARK 4.9. Similar to the Poisson case, if F_{μ_0} is a continuous eigenfunction of Γ associated with the zero eigenvalue, then $F_{\mu_0} = 0$.

The following result shows that under Assumption (22), the solutions of the ordinary differential equation (21) are oscillating functions. In this particular setting, we then obtain a behavior close to the behavior observed for the Poisson case.

THEOREM 4.10. Assume $\alpha < \beta$ and

$$(22) \quad \frac{\alpha(2\beta - \alpha)}{\beta - \alpha} \left(2 + \frac{3 - 2e^{-(\beta-\alpha)/2} - e^{-(\beta-\alpha)}}{\beta - \alpha} \right) < 1.$$

Let

$$w_1 = \frac{\beta w_0}{\beta - \alpha}.$$

Then the eigenvalues $(\lambda_j)_{j \geq 1}$ are simple and for all $j \geq 1$, there exists a solution (λ_j, H_j) to the system (21) such that $\|F_{\mu_j}\|_2 = 1$ and we have

$$(23) \quad \left| \lambda_j - \frac{4w_1}{\pi^2(2j-1)^2} \right| \leq C_1 j^{-4}$$

and

$$(24) \quad \sup_{t \in [0,1]} |F_{\mu_j}(t) - \sqrt{2} \sin(\pi(2j-1)t/2)| \leq C_2 j^{-1},$$

for C_1 and C_2 two constants not depending on j .

Let us discuss Assumption (22), which is used to solve the system (21). More precisely, this condition allows to control norms of operators introduced in Lemma S.2.2 (in the Supplementary Material) to obtain the expression of H_j provided by equation (S.2.3.21). Whether Assumption (22) is necessary to state the conclusions of Theorem 4.10 remains an open question. Since

$$2 + \frac{3 - 2e^{-(\beta-\alpha)/2} - e^{-(\beta-\alpha)}}{\beta - \alpha} \in [2, 4]$$

for any $0 < \alpha < \beta$, the main term in the left-hand side of Assumption (22) is

$$\frac{\alpha(2\beta - \alpha)}{\beta - \alpha} = \frac{\alpha(2 - \|h\|_1)}{1 - \|h\|_1}.$$

It means that Assumption (22) is satisfied as soon as α is not too large and $\|h\|_1$ is not too close to 1. Interpreting α as the size of the jumps and $\|h\|_1$ as the influence of the past, it means that the size of the jumps cannot be too large and the influence of the past cannot be too strong. Note that when $\|h\|_1$ or α go to 0, the Hawkes process becomes closer to a homogeneous Poisson process with parameter w_1 . This can be seen in the definition of the intensity w but also in the reformulation of System (21):

$$\begin{cases} -\lambda y''(t) = w_1 y(t) + \frac{w_1 \alpha (2 - \|h\|_1)}{2(1 - \|h\|_1)} \int_0^1 e^{-\alpha(\frac{1}{\|h\|_1} - 1)|t-s|} y(s) ds, & t \in (0, 1), \\ y(1) = 0, & y'(0) = 0. \end{cases}$$

If y is bounded, then the second term of the right-hand side goes to 0 when $\|h\|_1$ or α goes to 0.

Surprisingly, we observe that the eigenvalues have an asymptotic behavior similar to the homogeneous Poisson case. Since the upper bound (23) is polynomial in j , approximations are sharp. Observe that the parameter w_1 is the average number of points in $[0, 1]$. It is the analog of parameter w_0 for the homogeneous Poisson process studied in Section 4.1.1. In both cases, we have that λ_j is close to $\frac{\mathbb{E}[\Pi([0,1])]}{(j\pi - \pi/2)^2}$. The Poisson analogy for eigenfunctions F_{μ_j} is valid as well (see (16)). To complement these theoretical contributions, a numerical study is lead in Section 6. In particular, we examine in Section 6 whether Assumption (22) is a technical artefact or not.

5. Estimation of eigenelements. In practice, dimension reduction of point processes is performed by estimating the principal measures μ_j , the associated eigenvalues λ_j and the scores $\xi_{i,j}$ (see (9)). The estimation of the principal measures μ_j is based on the estimation of the associated cumulative mass functions F_{μ_j} , and hence necessitates the estimation of eigenfunctions and eigenvalues of the covariance operator Γ_Δ . To proceed, by using the random measures $(\Pi_1, \dots, \Pi_n)$, we first define an empirical version of the covariance operator Γ_Δ using the empirical counterparts of the covariance measure C_Δ and kernel K_Δ . We recall that we assume that the point processes $(N_1, \dots, N_n)$ are independent and identically distributed. Let

$$\widehat{W} = \frac{1}{n} \sum_{i=1}^n \Pi_i \quad \text{and} \quad \widehat{\Delta}_i = \Pi_i - \widehat{W}.$$

Then we define

$$\widehat{C}_\Delta(B \times B') = \frac{1}{n} \sum_{i=1}^n \sum_{T, T' \in N_i} \mathbf{1}_{\{(T, T') \in B \times B'\}} - \widehat{W}(B) \times \widehat{W}(B'), \quad B, B' \in \mathcal{B},$$

which provides the definition of the empirical covariance kernel:

$$\widehat{K}_\Delta(s, t) = \widehat{C}_\Delta([0, s] \times [0, t]), \quad s, t \in [0, 1]$$

along with the empirical integral operator:

$$\widehat{\Gamma}_\Delta(f)(\cdot) = \int_0^1 \widehat{K}_\Delta(\cdot, t) f(t) dt,$$

that can be expressed using the cumulative mass functions of the empirical measures, namely $F_{\widehat{\Delta}_i}(t) = \widehat{\Delta}_i([0, t])$, which leads to

$$\widehat{\Gamma}_\Delta = \frac{1}{n} \sum_{i=1}^n F_{\widehat{\Delta}_i} \otimes F_{\widehat{\Delta}_i}.$$

Then $\text{Im}(\widehat{\Gamma}_\Delta) \in \text{span}\{F_{\widehat{\Delta}_1}, \dots, F_{\widehat{\Delta}_n}\}$, which implies that $\widehat{\Gamma}_\Delta$ is a finite-rank operator. Since it is also self-adjoint, the covariance operator is compact and the diagonalization theorem ensures the existence of a basis $(\widehat{\eta}_j)_{j \geq 1}$ of eigenfunctions of $\widehat{\Gamma}_\Delta$. We denote by $(\widehat{\lambda}_j)_{j \geq 1}$ the associated eigenvalue sequence, sorted in nonincreasing order, so that

$$\widehat{\Gamma}_\Delta = \sum_{j \geq 1} \widehat{\lambda}_j \widehat{\eta}_j \otimes \widehat{\eta}_j.$$

In particular, the eigenelements $(\lambda_j, \eta_j)_{j \geq 1}$ are estimated by $(\widehat{\lambda}_j, \widehat{\eta}_j)_{j \geq 1}$. Finally, for any process N_i with $i \in \{1, \dots, n\}$ and any $j \geq 1$, we estimate the scores $\xi_{i,j}$ defined in Section 3 by

$$\widehat{\xi}_{i,j} = \frac{\langle \widehat{\eta}_j, F_{\widehat{\Delta}_i} \rangle}{\sqrt{\widehat{\lambda}_j}} \quad \text{if } \widehat{\lambda}_j > 0.$$

The case $\hat{\lambda}_j = 0$ is not considered since it corresponds to eigenfunctions $\hat{\eta}_j$ that are in the kernel of $\hat{\Gamma}_{\hat{\Delta}}$; this is not encountered in practice.

Since the operator $\hat{\Gamma}_{\hat{\Delta}}$ is a finite-rank operator, it is possible to build a matrix, denoted by $\hat{G}_{\hat{\Delta}}$, such that the eigenelements of $\hat{\Gamma}_{\hat{\Delta}}$ can be derived explicitly from the eigenelements of $\hat{G}_{\hat{\Delta}}$. To construct $\hat{G}_{\hat{\Delta}}$, we consider all occurrences sorted in nondecreasing order:

$$\mathcal{T} = \left(\bigcup_{i=1}^n N_i \right) \cup \{0, 1\} = \{T_0, \dots, T_{|\mathcal{T}|}\},$$

so that $T_0 = 0$ and $T_{|\mathcal{T}|} = 1$. If some ties occur when we take the union of the processes N_i , we identify them in only one T_ℓ so that for any $\ell \neq \ell'$, we have $T_\ell \neq T_{\ell'}$. Then we define the histogram system $(e_1, \dots, e_{|\mathcal{T}|})$ associated to this grid

$$e_\ell(t) = \frac{1}{\sqrt{T_\ell - T_{\ell-1}}} 1_{[T_{\ell-1}; T_\ell)}(t), \quad \ell = 1, \dots, |\mathcal{T}|,$$

and we set

$$\hat{G}_{\hat{\Delta}} = (\langle \hat{\Gamma}_{\hat{\Delta}}(e_\ell), e_{\ell'} \rangle)_{1 \leq \ell, \ell' \leq |\mathcal{T}|}.$$

Since $\hat{\Gamma}_{\hat{\Delta}}$ is self-adjoint operator, $\hat{G}_{\hat{\Delta}}$ is a self-adjoint matrix. In the following lemma, we show that $\hat{G}_{\hat{\Delta}}$ can be easily computed. Furthermore, we show that eigenelements of $\hat{\Gamma}_{\hat{\Delta}}$ can be deduced from eigenelements of $\hat{G}_{\hat{\Delta}}$.

LEMMA 5.1.

1. *The elements of $\hat{G}_{\hat{\Delta}}$ are of closed form: for any $1 \leq \ell, \ell' \leq |\mathcal{T}|$,*

$$\begin{aligned} (\hat{G}_{\hat{\Delta}})_{\ell, \ell'} &= \frac{\sqrt{(T_\ell - T_{\ell-1})(T_{\ell'} - T_{\ell'-1})}}{n} \\ &\quad \times \sum_{i=1}^n (F_{\Pi_i}(T_{\ell'-1})F_{\Pi_i}(T_{\ell-1}) - F_{\widehat{W}}(T_{\ell'-1})F_{\widehat{W}}(T_{\ell-1})). \end{aligned}$$

In particular, if for all ℓ , there exists a unique i such that $T_\ell \in N_i$, we have $F_{\widehat{W}}(T_{\ell-1}) = (\ell - 1)/n$ and

$$(\hat{G}_{\hat{\Delta}})_{\ell, \ell'} = \frac{\sqrt{(T_\ell - T_{\ell-1})(T_{\ell'} - T_{\ell'-1})}}{n} \times \sum_{i=1}^n \left(F_{\Pi_i}(T_{\ell'-1})F_{\Pi_i}(T_{\ell-1}) - \frac{\ell - 1}{n} \frac{\ell' - 1}{n} \right).$$

2. *Let $\text{Sp}(\hat{\Gamma}_{\hat{\Delta}})$ (resp., $\text{Sp}(\hat{G}_{\hat{\Delta}})$) be the set of eigenvalues of the operator $\hat{\Gamma}_{\hat{\Delta}}$ (resp., of $\hat{G}_{\hat{\Delta}}$). We have*

$$\text{Sp}(\hat{\Gamma}_{\hat{\Delta}}) \setminus \{0\} \subset \text{Sp}(\hat{G}_{\hat{\Delta}}) \subset \text{Sp}(\hat{\Gamma}_{\hat{\Delta}}).$$

3. *Let $\hat{v}_j = (\hat{v}_j^1, \dots, \hat{v}_j^{|\mathcal{T}|})^t$ a unit-norm eigenvector of $\hat{G}_{\hat{\Delta}}$ associated with the eigenvalue $\hat{\lambda}_j$ and let*

$$\hat{\eta}_j = \sum_{\ell=1}^{|\mathcal{T}|} \hat{v}_j^\ell e_\ell.$$

Then $\hat{\eta}_j$ is a unit-norm eigenfunction of the operator $\hat{\Gamma}_{\hat{\Delta}}$ associated with the eigenvalue $\hat{\lambda}_j$ and there exists a unique measure $\hat{\mu}_j$ such that $F_{\hat{\mu}_j} = \hat{\eta}_j$.

REMARK 5.2. If $0 \in \text{Sp}(\hat{G}_{\hat{\Delta}})$, then the previous result shows that $\text{Sp}(\hat{G}_{\hat{\Delta}}) = \text{Sp}(\hat{\Gamma}_{\hat{\Delta}})$.

REMARK 5.3. Of course, the definition of $\hat{\eta}_j$ is arbitrary since several eigenvectors may be associated to the eigenvalue $\hat{\lambda}_j$. In particular, we could consider $-\hat{\eta}_j$. Actually, in practice, the dimension of the estimated eigenspaces associated to each $\hat{\lambda}_j$ is equal to 1. In this case, the definition of $\hat{\eta}_j$ is unique up to its sign.

REMARK 5.4. The estimated scores can be written

$$\begin{aligned}\hat{\xi}_{i,j} &= \frac{\langle \hat{\eta}_j, F_{\hat{\Delta}_i} \rangle}{\sqrt{\hat{\lambda}_j}} = \hat{\lambda}_j^{-1/2} \sum_{\ell=1}^{|\mathcal{T}|} \frac{\hat{v}_j^\ell}{\sqrt{T_{\ell-1} - T_\ell}} \int_{T_{\ell-1}}^{T_\ell} F_{\hat{\Delta}_i}(t) dt \\ &= \hat{\lambda}_j^{-1/2} \sum_{\ell=1}^{|\mathcal{T}|} \hat{v}_j^\ell \sqrt{T_\ell - T_{\ell-1}} F_{\hat{\Delta}_i}(T_{\ell-1}),\end{aligned}$$

where the last equality comes from equation (S3.1.24) of the Supplementary Material.

REMARK 5.5. Since $\hat{\Gamma}_{\hat{\Delta}}$ is a finite-rank operator, we establish in Section S.3 of the Supplementary Material that $\text{Im}(\hat{\Gamma}_{\hat{\Delta}}) \subset \text{span}\{e_1, \dots, e_{|\mathcal{T}|}\}$ (see (S.3.1.25)). In particular, eigenvectors associated with nonzero eigenvalues of $\hat{\Gamma}_{\hat{\Delta}}$ are histograms. The expression provided by the third point of Lemma 5.1 specifies the values taken by each histogram $\hat{\eta}_j$.

The following theorem shows that the estimates introduced previously achieve the parametric rate for estimating eigenelements λ_j , η_j and μ_j .

THEOREM 5.6. *We assume that*

$$(25) \quad \mathbb{E}[\|F_{\Delta}\|^4] < +\infty.$$

We have

$$(26) \quad \mathbb{E}\left[\sup_{j \geq 1} |\hat{\lambda}_j - \lambda_j|^2\right] \leq 4 \frac{\mathbb{E}[\|F_{\Delta}\|^4]}{n}.$$

If we further assume that the eigenvalues of Γ_{Δ} are simple, then by introducing the eigengaps

$$\delta_1 = \lambda_1 - \lambda_2$$

and

$$\delta_j = \min\{\lambda_j - \lambda_{j+1}; \lambda_{j-1} - \lambda_j\}, \quad j \geq 2,$$

with $\tilde{\eta}_j = \text{sign}(\langle \hat{\eta}_j, \eta_j \rangle) \eta_j$, we have

$$(27) \quad \mathbb{E}[\|\hat{\eta}_j - \tilde{\eta}_j\|^2] \leq 32\delta_j^{-2} \frac{\mathbb{E}[\|F_{\Delta}\|^4]}{n}$$

and

$$(28) \quad \mathbb{E}[\|\hat{\mu}_j - \tilde{\mu}_j\|_{\mathcal{H}^{-1}}^2] \leq 32\delta_j^{-2} \frac{\mathbb{E}[\|F_{\Delta}\|^4]}{n},$$

with $\tilde{\mu}_j = \text{sign}(\langle \hat{\eta}_j, \eta_j \rangle) \mu_j$.

The moment Assumption (25) is very mild. Furthermore, Theorem 4.4 shows that for Poisson processes, the eigenvalues of Γ_{Δ} are simple. For Hawkes processes studied in Section 4.2, this condition holds under Assumption (22) (see Theorem 4.10). Even if our framework is nonparametric in nature, let us observe that regularization is not necessary to achieve parametric rates.

6. Numerical study. We now complement our theoretical contributions with some empirical results. We probe, in particular, the following questions: how does the truncated approximation behave? In particular, to what extent are the properties of eigenfunctions preserved in a nonasymptotic setting? For Hawkes processes, some restrictions due to Assumption (22) appear. Are these limitations purely theoretical ones? We discuss these points in the sequel.

We first simulate $n = 100$ Poisson processes using the `hawkesbow` package [14]. We consider different intensity functions: constant, $w(t) = 100$, linear, $w(t) = t$, $w(t) = \exp(-0.005t)$, sinusoidal, $w(t) = (1 + \sin(0.11t))$. Then we compute the average eigenfunctions and eigenvalues over 50 replicates (Figures 1 and 2, respectively). For eigenvalues, the empirical asymptotic regime in j conforms to the theoretical regime in $(j\pi)^{-2}$ (with a slight modification $(j\pi - \pi/2)^{-2}$ that better fits for small j 's, which matches the remarks following Theorem 4.4). The form of the intensity does not impact much the convergence of eigenvalue estimates, even for large scales (j small). As expected, eigenfunctions show increasing oscillations with j . However, the form of the intensity induces different oscillatory patterns in eigenfunctions but the number of zero crossings is preserved with j zeros on the interval $[0, 1)$ for all functions.

Then, to investigate the impact of dependencies in the data, we simulated $n = 1000$ Hawkes processes on $[0, 100]$, with $w_0 = 100$, and self-exciting function $h(t) = \alpha e^{-\beta t}$. To explain briefly the simulation method, the package considers a reproduction function h that is decomposed such as $h = mh^*$, with $m = \int_0^\infty h(t) dt$ and h^* a true density function such that $\int_0^\infty h^*(t) dt = 1$. In our setting, this means that $h^*(t) = \beta \exp(-\beta t)$, and that

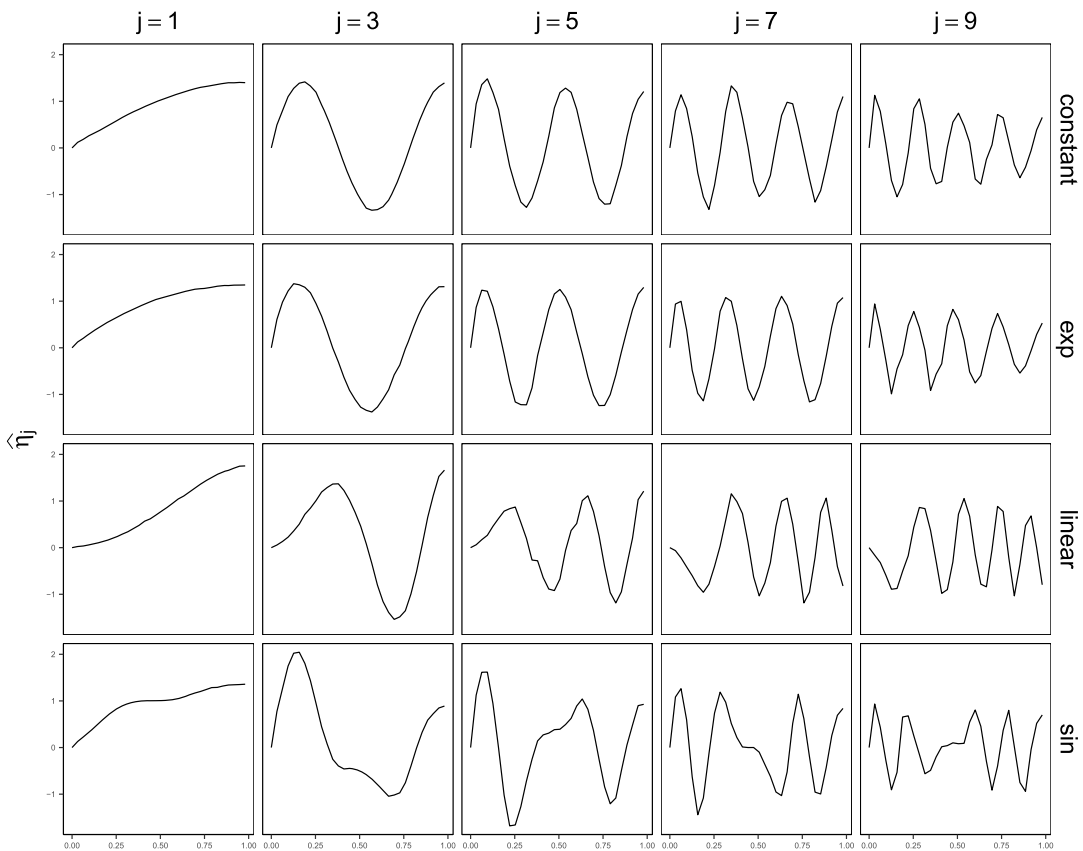


FIG. 1. Average eigenfunctions for Poisson processes with different intensity functions over 50 replicates.

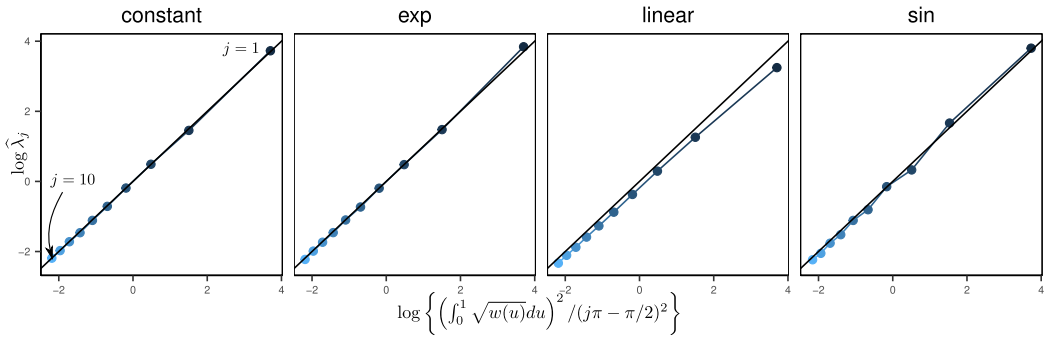


FIG. 2. Average eigenvalues (log-scale) for Poisson processes over 50 replicated. Each dot corresponds to a value of $j \in \{1, \dots, 10\}$. The empirical average is plotted vs. the expected theoretical asymptotic regime of eigenvalues in $(\int_0^1 \sqrt{w(u)} du)^2 / (j\pi - \pi/2)^2$, as expected from Theorem 4.4. Note that Theorem 4.4 provides a $(j\pi)^{-2}$ regime. The black line corresponds to the first bisector, so that the points align if the empirical convergence matches the theoretical regime.

$h(t) = \|h\|_1 \times h^*(t)$. Since Theorem 4.10 exhibits the central role of $\|h\|_1$ and α , in our simulation we consider $\alpha \in \{0.01, 0.1, 1, 10\}$ and $\|h\|_1 \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$. Data were finally rescaled in $[0, 1]$. This setting allows us to explore different strengths and ranges of dependency.

The eigenfunctions show a periodic pattern with zero crossings depending on j (Figure 3). Interestingly, the number of zero crossings does not depend on the parameters $(\alpha, \|h\|_1)$ and agrees with the values predicted by the theory for all studied values of j (note that for visualization purposes, eigenfunctions are studied for a small range of values of j). Figure 4 shows that convergence of eigenvalues holds for all parameters $(\alpha, \|h\|_1)$ but the convergence is slower when α or $\|h\|_1$ is large. Observe that in our setting, Assumption (22) is satisfied

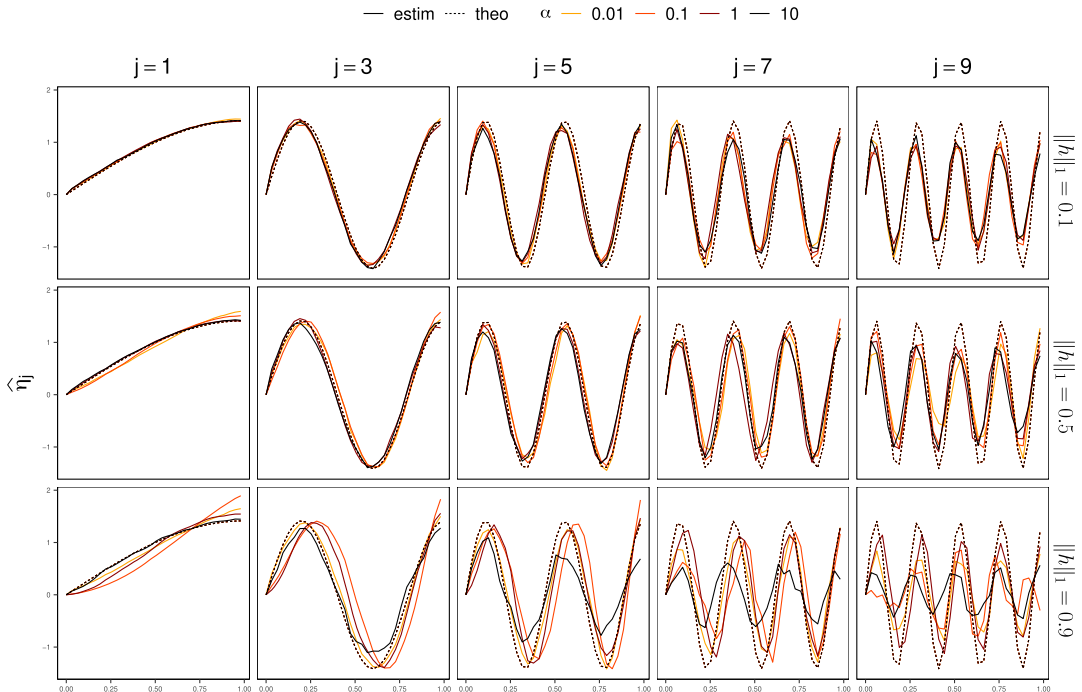


FIG. 3. Average eigenfunctions for Hawkes' processes with different transfert functions over 50 replicates. Dotted lines: asymptotic eigenfunctions $t \mapsto \sqrt{2} \sin(\pi(2j-1)t/2)$ (see Theorem 4.10).

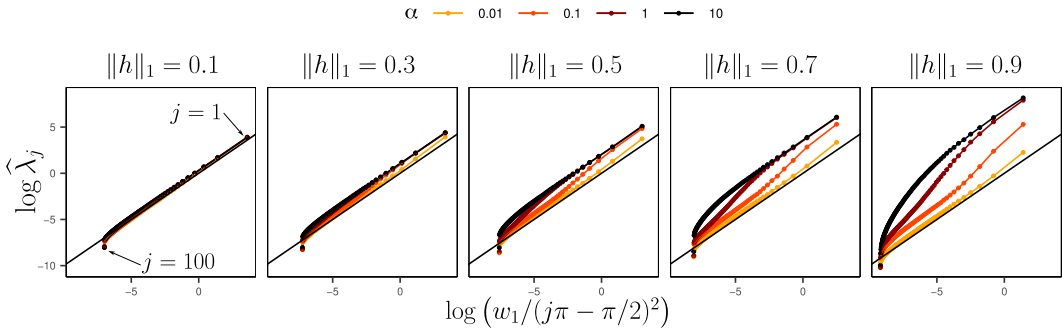


FIG. 4. Average eigenvalues (log-scale) for Hawkes' processes over 50 replicated. Each dot corresponds to a value of $j \in \{1, \dots, 50\}$. The empirical average is plotted vs. the expected theoretical asymptotic regime of eigenvalues in $w_1/(j\pi - \pi/2)^2$, as expected from Theorem 4.10. The black line corresponds to the first bisector, so that the points align if the empirical convergence matches the theoretical regime.

only in few cases: if $\alpha = 0.01$, or if $\alpha = 0.1$ and $\|h\|_1 \in \{0.1, 0.3\}$, so we conjecture that Assumption (22) may be a technical artefact and Theorem 4.10 may hold in more general situations.

From this study, we can conclude that many of the theoretical conclusions of Section 4, which are valid when j is large, seem to extend for small values of j for Poisson and Hawkes' processes (for the latter at least when $\|h\|_1$ and α are not too large). Note that the eigenfunctions for the Poisson and Hawkes' processes have a very similar oscillating behavior, and this at each scale.

7. Applications.

7.1. PCA on point processes reveals multiscale variability in earthquake occurrence data.

We begin by illustrating our method through the analysis of data obtained from the Kandilli Observatory and Earthquakes Research Institute at Boğaziçi University.² The dataset comprises earthquake occurrences in Turkey and neighboring regions of Greece, recorded between January 2013 and January 2023, spanning 1181 cities (Figure 5A). In recent years, the Gulf of Gökova in Southwest Turkey has witnessed two significant seismic events: the Bodrum earthquake on July 16, 2017 (magnitude 6.6, also felt on the Greek island of Kos) [39], and the Aegean Sea earthquake on November 1, 2020, with a moment magnitude of 7.0 (the highest magnitude observed during the period). In the following, we illustrate how our framework can be employed to investigate the fine-scale dynamics of earthquake occurrences in the region, focusing on the 195 cities that experienced more than two earthquakes over the specified period. Finally, note that our framework assumes the observed processes (earthquake occurrences) are independent from one city to another, even though spatial dependencies are likely to exist. The spatiotemporal modeling of such data lies beyond the scope of the present work.

We start by inspecting the significance of the first eigenfunction $\hat{\eta}_1$ (Figure 5C) and the scores on this axis (Figure 5D). We see that $\hat{\eta}_1$ corresponds to the temporal accumulation of the number of earthquakes across cities and over time, and that the scores $(\hat{\xi}_{i1})_i$ directly correspond to the total number of earthquakes in each city over the specified period. Since this axis carries $\hat{\lambda}_1/\sum_j \hat{\lambda}_j = 89\%$ of the variability, we conclude that the first source of variance in the data is associated with the deviation of cities in the accumulation of earthquakes over time compared to the average temporal pattern of earthquakes in the region. We identify Akdeniz as an outlier with an unusually high number of earthquakes compared to the regional

²<http://www.koeri.boun.edu.tr/sismo/2/earthquake-catalog/>.

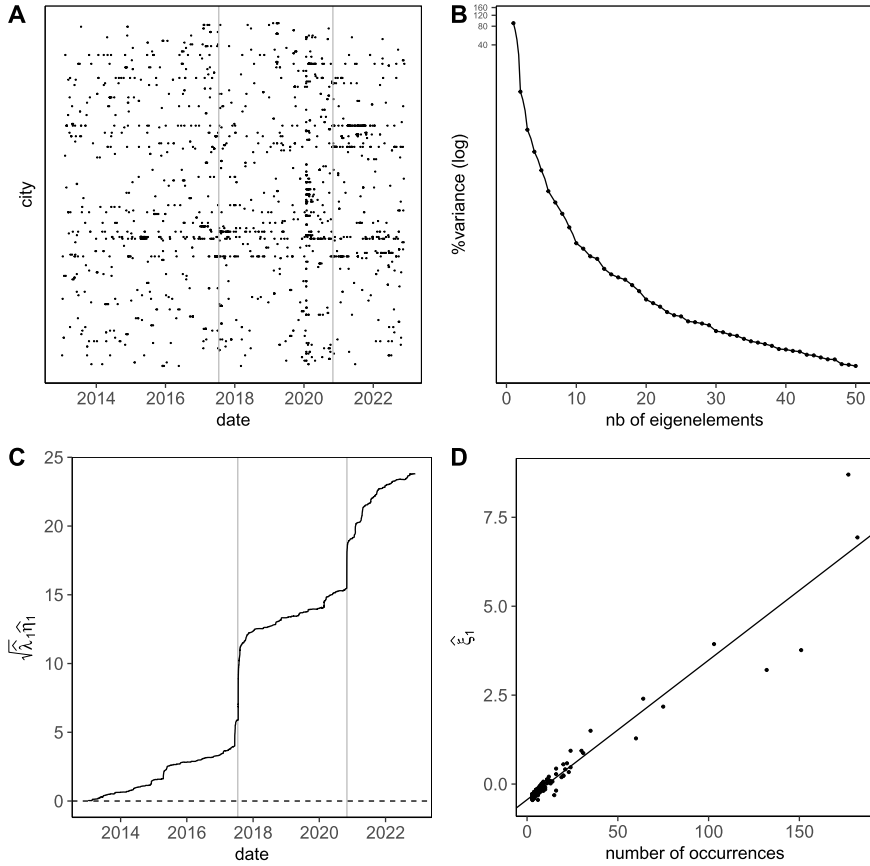


FIG. 5. A: Raster plot of earthquakes during the period 2013–2023. Each line corresponds to a city and each dot to an earthquake occurrence. Breakpoint dates (grey vertical lines) correspond to 2017.07.16 and 2020.11.01 B: Percentage of variance according to the number of eigenlements (log scale). C: First rescaled eigenfunction $\sqrt{\hat{\lambda}_1} \hat{\eta}_1$ according to the date of earthquakes (left). Breakpoint dates (grey vertical lines) correspond to 2017.07.16 and 2020.11.01. D: Plot of PCA scores for the first axis $\hat{\xi}_1$ according to the number of occurrences $W([0, t])$.

average. This city will be excluded from subsequent analyses that investigate the finer-scale dynamics of earthquake activity in the area.

In order to interpret the subsequent axes, we recall that our method summarizes the dynamics of earthquake occurrences through simple functions, as expressed in equation (9). Consequently, we focus on the variations of the estimated eigenfunctions $(\hat{\eta}_j)_j$, and represent the positions of cities according to their scores $(\hat{\xi}_j)_j$ on successive axes as in any PCA analysis (Figure 6). These representations allow us to identify cities (like Gokova Korfezi, Onika Adalar) that have atypical accumulations of earthquakes over the period (these cities correspond to observations with atypical scores on the axes).

Interestingly, the second axis, $\hat{\eta}_2$ (Figure 6C) reveals a distinct change in seismic activity between the two main earthquakes (2017 and 2020). Cities exhibiting positive scores on this axis (Figure 6A) indicate a global increase in earthquake rates between July 2017 and November 2020, followed by a decrease below the average regional rate after November 2020. Conversely, cities with negative scores on $\hat{\eta}_3$ (Figure 6A) show an accumulation of earthquakes before July 2017. Overall, our method provides a highly accurate description of the variability in earthquake occurrences among different cities. It offers a means to represent and position cities relative to each other based on their earthquake dynamics variability. Moreover, our estimation framework, relying on occurrence data without smoothing, cap-

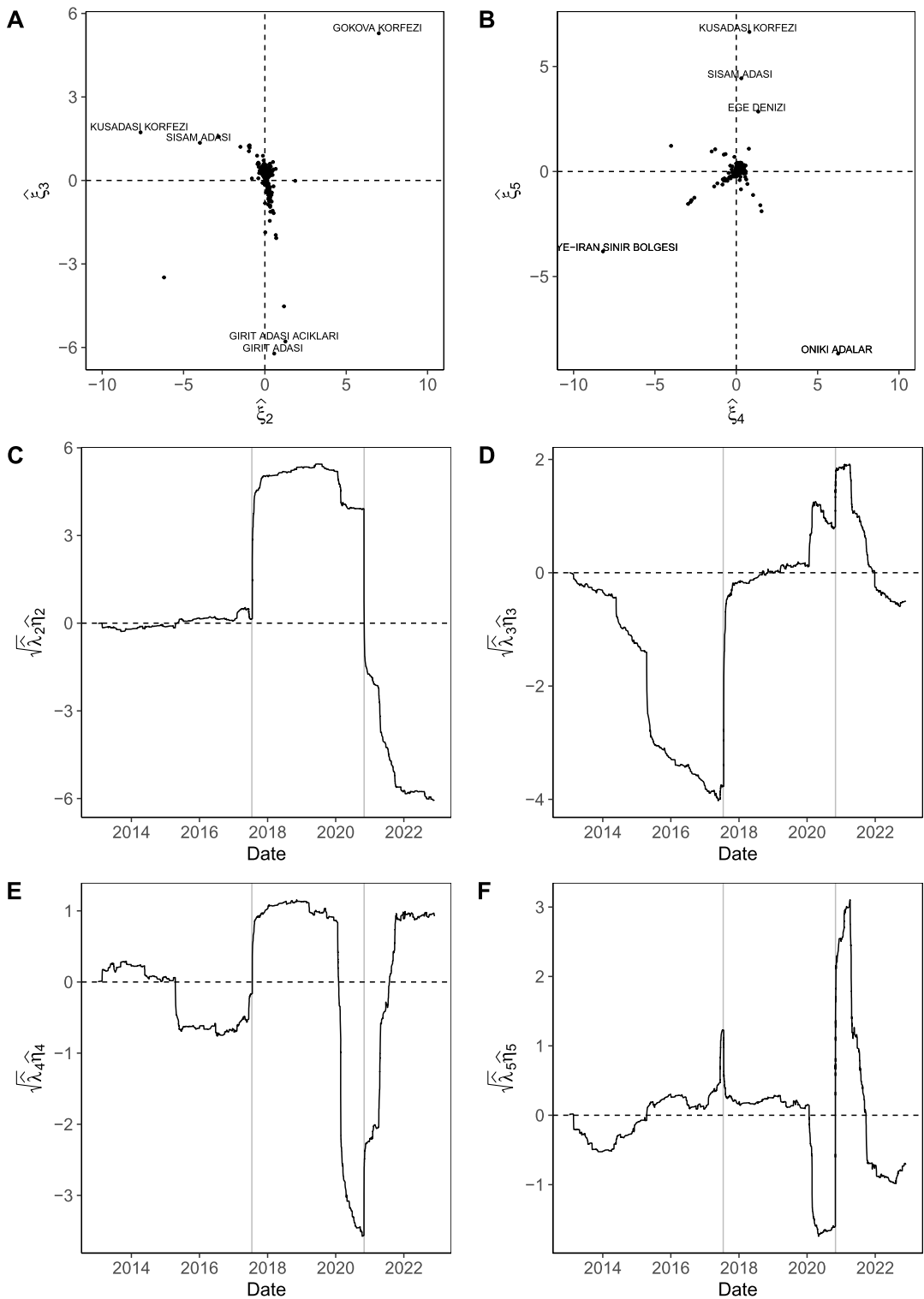


FIG. 6. A: Estimated scores of cities $\hat{\xi}_{i2}$ vs. $\hat{\xi}_{i3}$. B: Estimated scores of cities $\hat{\xi}_{i4}$ vs. $\hat{\xi}_{i5}$. C–D–E–F: Estimated rescaled eigenfunctions $\sqrt{\hat{\lambda}_j} \hat{\eta}_j$ over time for $j = 2, 3, 4, 5$.

tures sharp and fine-scale variations in this dynamics, as shown by the distinct peaks in $\hat{\eta}_4$ and $\hat{\eta}_5$ (Figure 6E–F).

7.2. PCA on point processes to investigate single-cell epigenomic variations. The field of single-cell biology has emerged as a significant framework for generating detailed molecular portraits of cell populations, providing valuable insights into the complexity of living tissues [64]. While gene expression analysis has traditionally been instrumental in understanding fundamental biological processes, epigenomic-based gene regulation is receiving increasing attention, especially when mediated by chromatin modifications measured by sequencing-based chromatin immuno-precipitation assays (chIP-Seq). Briefly, chromatin comprises the fiber that constitutes chromosomes. ChIP-Seq assays enable the mapping of chromatin modifications, generating data in the form of 1D-spatial coordinates. However, there is currently no dedicated method for analyzing single-cell epigenomic profiles that adequately accounts for both the spatial nature of the signal and the single-cell resolution.

A recent study [47] has shown that a particular histone modification (H3K27me3) is involved in the emergence of drug persistence in breast cancer cells. Drug persistence occurs when only a subset of cells, known as persister cells, survives the initial drug treatment, thereby creating a reservoir of cells from which resistant cells will emerge. Persister cells exhibit dynamic changes in H3K27me3 modifications at the single-cell level, but the variability of this pattern makes it difficult to relate to the emergence of resistance within tumor cells. Previous analyses have suggested that a pool of untreated cells could contribute to the persister cell population later upon exposure to chemotherapy. However, the lack of an appropriate methodological framework to handle the spatial characteristics of the data has made the clear identification of this pool of precursor cells difficult.

We propose using our PCA framework to explore the variability of H3K27me3 binding along the chromosomes 2 and 3 genomes among untreated cancer cells. To proceed, we performed PCA on $n = 4722$ cells, with each cell described by the 1D spatial positions of H3K27me3 modifications along chromosomes. Eigenfunctions estimated by this approach allow us to visualize the basic components of spatial variability in H3K27me3 occurrences among untreated cancer cells (Figure 7A–B). Our PCA framework then provides principal scores $\hat{\xi}_{ij}$, on which we conducted k -means clustering. Using $J = 50$ principal components and selecting the number of clusters based on a silhouette criterion, we identified clusters of cells with similar profiles of histone modification accumulation across the genome. Since k -means clustering assigns label variables to each observed process ($\hat{Z}_{ik} = 1$ if cell i is in cluster k , and $\hat{n}_k = \sum_i \hat{Z}_{ik}$ the size of cluster k), we represent the estimated clusterwise cumulative mass function ($\hat{\Pi}_k([0, t])$, Figure 7C–D) and associated histograms ($\hat{w}_k(t)$, Figure 7E–F):

$$\hat{\Pi}_k([0, t]) = \frac{1}{\hat{n}_k} \sum_{i=1}^n \hat{Z}_{ik} \Pi_i([0, t]), \quad \hat{w}_k(t) = \frac{1}{\hat{n}_k} \sum_{i=1}^n \hat{Z}_{ik} \sum_{T \in N_i} \mathbf{1}_{\{T=t\}}.$$

This analysis reveals the spatial variability of histone modifications among cancer cells, offering a new perspective on epigenomic heterogeneity by revealing three subpopulations within untreated cancer cells, characterized by zones along chromosomes that concentrate higher or lower numbers of occurrences. Further research is needed to fully understand the biological implications of these variabilities and their role in forming a pool of persister cells. Nonetheless, our PCA for point processes framework marks an important step towards investigating spatial epigenomic variations among cell populations.

7.3. PCA on point process for multielectrodes spike data. Neuroscience has witnessed high-paced developments in recent years, thanks to the convergence of experimental and computational approaches. Signal processing and machine learning, in general, are deeply connected to the analysis of neuroscience data, often in the form of spike trains that correspond to neuronal activities. These spike trains are typically modeled as point processes, and

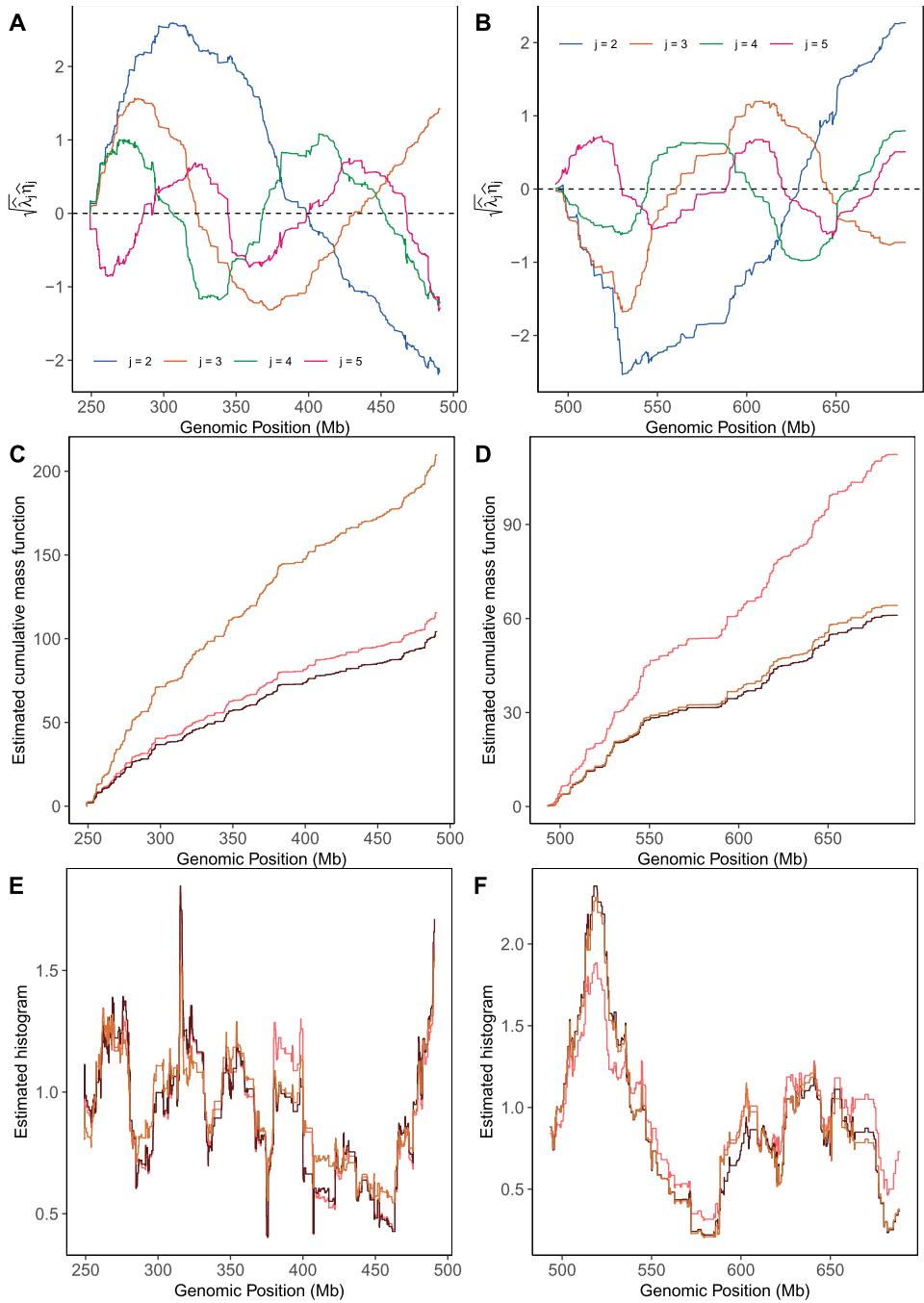


FIG. 7. A–B: Eigenfunctions weighted by their respective eigenvalues $\sqrt{\hat{\lambda}_j} \hat{\eta}_j$ according to the position along the genome. Chromosome 2 (left) and Chromosome 3 (right). C–D: Estimated clusterwise cumulative mass functions $\hat{\Pi}_k$. E–F: Estimated clusterwise histograms $\hat{w}_k$.

technological advancements now allow for the recording of such spikes in populations of neurons, paving the way for the analysis framework of multiple neural spike train data [21]. Among many methodological challenges, dimension reduction emerges as a key challenge for visualizing, exploring and understanding the structure and variability of population activity. The goal of dimensionality reduction is to characterize how the firing rates of different neurons covary, while discarding the spiking variability as noise. We present how PCA on

point processes can be utilized in this context. In particular, the principal measure we infer can be interpreted as a latent common point process from which individual action potentials are generated in a stochastic manner. Hence, our representation with principal measures enables us to define a latent space that represents shared activity patterns prominent in the population response.

We consider data from an optogenetic therapy experiment performed on nonhuman primates retina. The aim of this experiment is to restore light sensitivity in the residual retinal tissue after photoreceptor generation like macular degeneration. These diseases affect photoreceptor cells, but the retinal ganglion cells (RGC) can still communicate with the brain via the optic nerve. Then the data consists in the recording the neuronal activity of retinal fragments placed on Multi-Electrodes-Array (MEA, $n = 256$ electrodes) grid. These 256 electrodes form a 16×16 grid onto a device that is positioned on the retina of macaques. The activity of each electrode can be related to a particular retina region, so that the data are typically spatiotemporal, but we only consider its temporal components. Different visual stimuli are applied to retinal fragments and the activity of electrodes is recorded.

Thanks to our framework, we can easily visualize the variability in the response of each electrode (Figure 8A), as well as the temporal dynamics that explain this variability in spike train occurrences (Figure 8B). These principal components can be interpreted as latent spike trains that compose the dynamics of the population. These latent processes could be of interest for understanding the neuronal coordination underlying the population response. Additionally, the score of each electrode can be represented with respect to the position of the electrodes on the device (Figure 8C–D) to visualize variations in each principal component across the device. Identifying regions on the device that show high or low scores could be related to brain regions whose responses are governed by the corresponding latent spiking process. Finally, this dynamical system can also be summarized by representing trajectories in the principal measure space (Figure 8E–F). In this representation, each time point t corresponds to a single point in the latent firing rate space $(\hat{\eta}_2(t), \hat{\eta}_3(t))$. This representation allows us to follow the temporal dynamics of the observed neuron population. Hence, we expect that PCA on point processes can become a natural framework to decompose these complex phenomena of population-based spike train analysis by providing simple decompositions that help unravel the diversity and variability of neuronal population spiking activities. In this application, the assumption of independence between neurons may seem as a strong one. Exploring the spatial dependency between neurons could be an interesting research direction to enhance the method.

8. Perspectives. In this work, we propose a novel mathematical and statistical framework for dimension reduction in replicated one-dimensional point processes. Our approach is based on functional PCA for random measures, for which we derive a Karhunen–Loève expansion that enables us to define the concept of principal measure. The key tool is the associated covariance operator whose eigenfunctions are the cumulative mass functions of the principal measures. We provide a detailed characterization of these eigenelements in simple point process models such as the Poisson and Hawkes’ models. Assuming that the point processes are independent and identically distributed, we further develop an easily implementable, fully data-driven estimation procedure for these eigenelements that requires no smoothing. The resulting estimators achieve the parametric rate of convergence. These rates may, however, fail to hold when dependence is present across replicates (e.g., across spatial locations or neurons). Extending the methodology to dependent settings would be an interesting direction for future research. Our contribution also includes various applications across different fields, as well as a software package that implements our methodology.

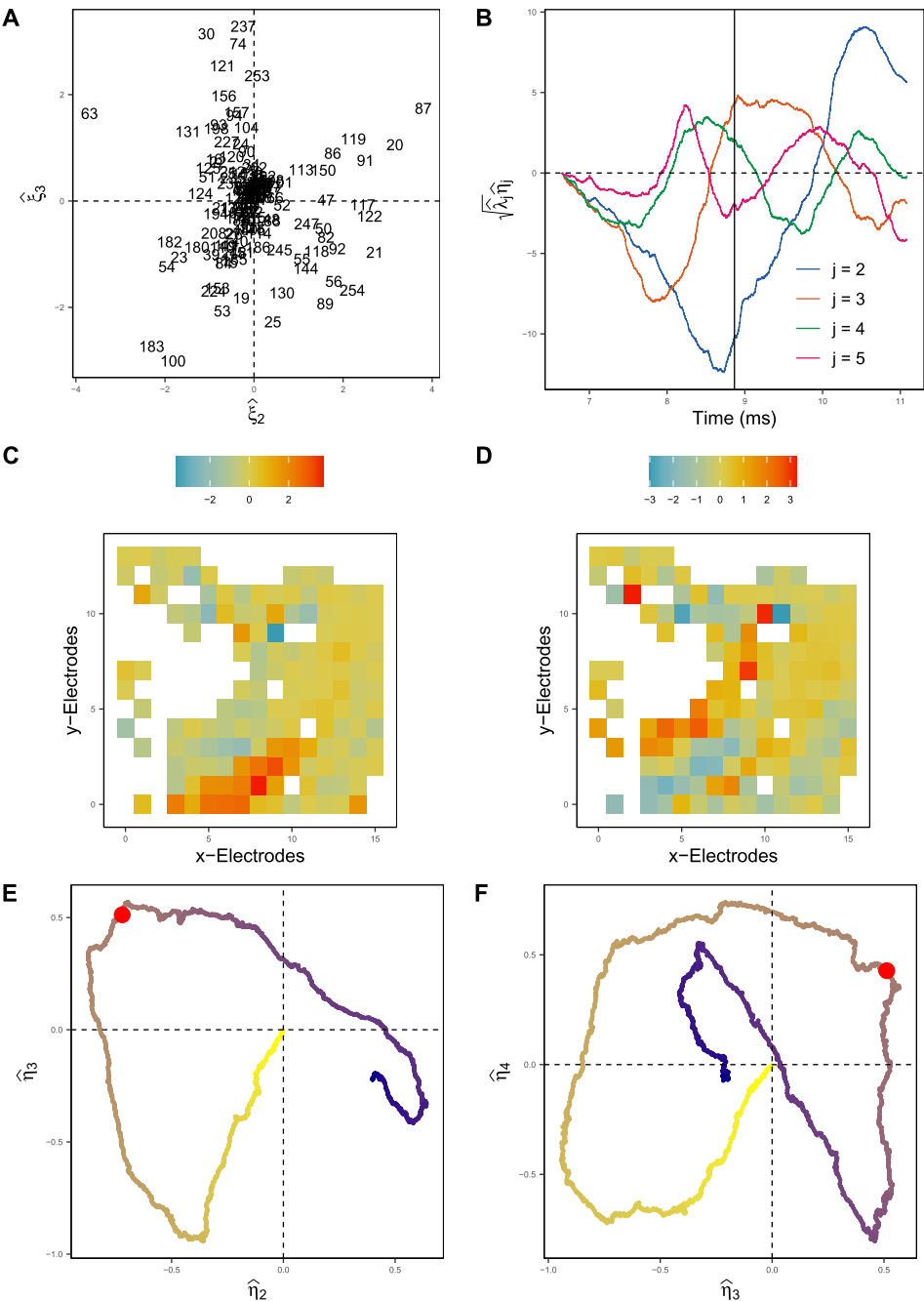


FIG. 8. *A*: scores $\hat{\xi}_2$ vs. $\hat{\xi}_3$ for electrodes $n = 1, \dots, 256$. *B*: Eigenfunctions weighted by their respective eigenvalues $\sqrt{\hat{\lambda}_j} \hat{\eta}_j$ according to time (in ms). The vertical line corresponds to the stimulus at $t = 8.86$ ms. *C-D*: visualization of scores $\hat{\xi}_2$ (*C*) and $\hat{\xi}_3$ (*D*) according to the position of each electrode on the retinal device. Red: high value, blue: low value. *E-F*: representation of the data in the latent space (*E*, $\hat{\eta}_3$ vs. $\hat{\eta}_2$, and *F*, $\hat{\eta}_4$ vs. $\hat{\eta}_3$). Time is represented by colors (yellow: beginning of the assay, purple: end of the assay). Red dot: stimulus time.

This work opens the door to many potential extensions, with a particularly interesting direction being the adaptation of our framework to multivariate processes. For example, the observed processes could be spatial point patterns, where $(N_1, \dots, N_n)$ are observed on $[0, 1]^d$. In this case, our framework could be applied to analyze image data, for instance, with $d = 2$ or $d = 3$. Another scenario involves each process N_i being observed on $[0, 1]$, where the

observed data consist of multiple dependent processes, such that $N_i = (N_i^1, \dots, N_i^d)$. Our methodology could then be extended to multivariate functional PCA, which seeks a joint representation of the d dependent processes [32]. In neuroscience, for example, repeated assays over time often involve multiple dependent neurons. Multivariate functional PCA for multivariate point processes could help uncover latent components that explain the variability in spiking events across time and neurons. Similarly, in single-cell genomics, new technologies allow for the simultaneous recording of multiple dependent chromatin modifications within a single assay at the single-cell level [48]. Extending our framework to multivariate processes requires a separate and dedicated study, in particular because our current methodology relies heavily on the concept of the cumulative mass function F_μ , which is defined on the one-dimensional interval $[0, 1]$ and takes values in $\mathbb{R}$. Extending our approach will thus necessitate the development of specific mathematical tools tailored to higher-dimensional settings.

Acknowledgments. The authors would like to thank Emmanuel Trélat for valuable suggestions concerning solutions of Sturm–Liouville theory, Celine Vallot for discussions on the single-cell epigenomic data and Gregory Gauvain, Fabrice Arcizet and Pierre Pouget for sharing the neuroscience data. The authors would like to warmly thank the Associate Editor and the anonymous referees for very valuable comments and suggestions.

Funding. The research was supported by a grant from the Agence Nationale de la Recherche ANR-18-CE45-0023 SingleStatOmics, by the projects AI4scMed, France 2030 ANR-22-PESN-0002.

SUPPLEMENTARY MATERIAL

Proofs and code (DOI: [10.1214/25-AOS2596SUPP](https://doi.org/10.1214/25-AOS2596SUPP); .zip). This Supplementary Material contains the complete proofs of our work. The method is implemented and available in the `pppca` R-package. Here is a brief example, with details in the vignette of the package: <https://github.com/franckpicard/pppca/>.

```
library(pppca)
data(earthquake)
earthquake_num <- lapply(earthquake, as.numeric)
pppca_res <- pppca(earthquake_num, Jmax = 50, mc.cores = 1)
```

REFERENCES

- [1] ANDERSEN, P. K., BORGAN, Ø., GILL, R. D. and KEIDING, N. (1993). *Statistical Models Based on Counting Processes*. Springer Series in Statistics. Springer, New York. [MR1198884 https://doi.org/10.1007/978-1-4612-4348-9](https://doi.org/10.1007/978-1-4612-4348-9)
- [2] ASH, R. B. and GARDNER, M. F. (1975). *Topics in Stochastic Processes. Probability and Mathematical Statistics, Vol. 27*. Academic Press, New York–London. [MR0448463](https://doi.org/10.1016/j.jneumeth.2022.109550)
- [3] BACRY, E., BOMPAIRE, M., GAIFFAS, S. and MUZY, J.-F. (2020). Sparse and low-rank multivariate Hawkes processes. *J. Mach. Learn. Res.* **21** Paper No. 50, 32. [MR4095329](https://doi.org/10.1016/j.jspi.2014.03.009)
- [4] BELITSER, E., SERRA, P. and VAN ZANTEN, H. (2015). Rate-optimal Bayesian intensity smoothing for inhomogeneous Poisson processes. *J. Statist. Plann. Inference* **166** 24–35. [MR3390131 https://doi.org/10.1016/j.jspi.2014.03.009](https://doi.org/10.1016/j.jspi.2014.03.009)
- [5] BONNET, A., DION-BLANC, C., GINDRAUD, F. and LEMLER, S. (2022). Neuronal network inference and membrane potential model using multivariate Hawkes processes. *J. Neurosci. Methods* **372** 109550. <https://doi.org/10.1016/j.jneumeth.2022.109550>
- [6] BONNET, A., MARTINEZ HERRERA, M. and SANGNIER, M. (2023). Inference of multivariate exponential Hawkes processes with inhibition and application to neuronal activity. *Stat. Comput.* **33** Paper No. 91, 26. [MR4606276 https://doi.org/10.1007/s11222-023-10264-w](https://doi.org/10.1007/s11222-023-10264-w)

- [7] BOUZAS, P. R., VALDERRAMA, M. J., AGUILERA, A. M. and RUIZ-FUENTES, N. (2006). Modelling the mean of a doubly stochastic Poisson process by functional data analysis. *Comput. Statist. Data Anal.* **50** 2655–2667. [MR2227341](#) <https://doi.org/10.1016/j.csda.2005.04.015>
- [8] BRÉMAUD, P. and MASSOULIÉ, L. (1996). Stability of nonlinear Hawkes processes. *Ann. Probab.* **24** 1563–1588. [MR1411506](#) <https://doi.org/10.1214/aop/1065725193>
- [9] BRÉMAUD, P. and MASSOULIÉ, L. (2001). Hawkes branching point processes without ancestors. *J. Appl. Probab.* **38** 122–135. [MR1816118](#) <https://doi.org/10.1017/s0021900200018556>
- [10] BREZIS, H. (2011). *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer, New York. [MR2759829](#)
- [11] CARRIZO VERGARA, R. (2022). Karhunen-Loève expansion of Random Measures. Available at [arXiv:2203.14202](#).
- [12] CARSTENSEN, L., SANDELIN, A., WINTHER, O. and HANSEN, N. R. (2010). Multivariate Hawkes process models of the occurrence of regulatory elements. *BMC Bioinform.* **11** 456.
- [13] CHEN, S., WITTEN, D. and SHOJAIE, A. (2017). Nearly assumptionless screening for the mutually-exciting multivariate Hawkes process. *Electron. J. Stat.* **11** 1207–1234. [MR3634334](#) <https://doi.org/10.1214/17-EJS1251>
- [14] CHEYSSON, F. (2023). *hawkesbow: Estimation of Hawkes Processes from Binned Observations* R package version 1.0.2.
- [15] CHIANG, W.-H., LIU, X. and MOHLER, G. (2022). Hawkes process modeling of COVID-19 with mobility leading indicators and spatial covariates. *Int. J. Forecast.* **38** 505–520. <https://doi.org/10.1016/j.ijforecast.2021.07.001>
- [16] CHIU, S. N., STOYAN, D., KENDALL, W. S. and MECKE, J. (2013). *Stochastic Geometry and Its Applications*, 3rd ed. *Wiley Series in Probability and Statistics*. Wiley, Chichester. [MR3236788](#) <https://doi.org/10.1002/9781118658222>
- [17] CHORNOBOY, E. S., SCHRAMM, L. P. and KARR, A. F. (1988). Maximum likelihood identification of neural point process systems. *Biol. Cybernet.* **59** 265–275. [MR0961117](#) <https://doi.org/10.1007/BF00332915>
- [18] COHN, D. L. (1993). *Measure Theory*. Birkhäuser, Boston, MA. [MR1454121](#)
- [19] CORLAY, S. and PAGÈS, G. (2015). Functional quantization-based stratified sampling methods. *Monte Carlo Methods Appl.* **21** 1–32. [MR3318550](#) <https://doi.org/10.1515/mcma-2014-0010>
- [20] CRANE, R. and SORNETTE, D. (2008). Robust dynamic classes revealed by measuring the response function of a social system. *Proc. Natl. Acad. Sci. USA* **105** 15649–15653.
- [21] CUNNINGHAM, J. P. and YU, B. M. (2014). Dimensionality reduction for large-scale neural recordings. *Nat. Neurosci.* **17** 1500–1509.
- [22] DALEY, D. J. and VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes. Vol. I. Elementary Theory and Methods*, 2nd ed. *Probability and Its Applications (New York)*. Springer, New York. [MR1950431](#)
- [23] DALEY, D. J. and VERE-JONES, D. (2008). *An Introduction to the Theory of Point Processes. Vol. II. General Theory and Structure*, 2nd ed. *Probability and Its Applications (New York)*. Springer, New York. [MR2371524](#) <https://doi.org/10.1007/978-0-387-49835-5>
- [24] DELAIGLE, A., HALL, P. and BATHIA, N. (2012). Componentwise classification and clustering of functional data. *Biometrika* **99** 299–313. [MR2931255](#) <https://doi.org/10.1093/biomet/ass003>
- [25] DONNET, S., RIVOIRARD, V. and ROUSSEAU, J. (2020). Nonparametric Bayesian estimation for multivariate Hawkes processes. *Ann. Statist.* **48** 2698–2727. [MR4152118](#) <https://doi.org/10.1214/19-AOS1903>
- [26] EMBRECHTS, P., LINIGER, T. and LIN, L. (2011). Multivariate Hawkes processes: An application to financial data. *J. Appl. Probab.* **48A** 367–378. [MR2865638](#) <https://doi.org/10.1239/jap/1318940477>
- [27] ESCABIAS, M., AGUILERA, A. M. and VALDERRAMA, M. J. (2004). Principal component estimation of functional logistic regression: Discussion of two different approaches. *J. Nonparametr. Stat.* **16** 365–384. [MR2073031](#) <https://doi.org/10.1080/10485250310001624738>
- [28] FARAJTABAR, M., WANG, Y., GOMEZ-RODRIGUEZ, M., LI, S., ZHA, H. and SONG, L. (2017). COE-VOLVE: A joint point process model for information diffusion and network evolution. *J. Mach. Learn. Res.* **18** Paper No. 41, 49. [MR3655306](#)
- [29] GAO, X. and ZHU, L. (2018). Functional central limit theorems for stationary Hawkes processes and application to infinite-server queues. *Queueing Syst.* **90** 161–206. [MR3850052](#) <https://doi.org/10.1007/s11134-018-9570-5>
- [30] GUSTO, G. and SCHBATH, S. (2005). FADO: A statistical method to detect favored or avoided distances between occurrences of motifs using the Hawkes’ model. *Stat. Appl. Genet. Mol. Biol.* **4** Art. 24, 28. [MR2170440](#) <https://doi.org/10.2202/1544-6115.1119>
- [31] HANSEN, N. R., REYNAUD-BOURET, P. and RIVOIRARD, V. (2015). Lasso and probabilistic inequalities for multivariate point processes. *Bernoulli* **21** 83–143. [MR3322314](#) <https://doi.org/10.3150/13-BEJ562>

- [32] HAPP, C. and GREVEN, S. (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. *J. Amer. Statist. Assoc.* **113** 649–659. [MR3832216 https://doi.org/10.1080/01621459.2016.1273115](https://doi.org/10.1080/01621459.2016.1273115)
- [33] HAWKES, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58** 83–90. [MR0278410 https://doi.org/10.1093/biomet/58.1.83](https://doi.org/10.1093/biomet/58.1.83)
- [34] HAWKES, A. G. (1971). Point spectra of some mutually exciting point processes. *J. Roy. Statist. Soc. Ser. B, Methodol.* **33** 438–443. [MR0358976 https://doi.org/10.1093/biomet/33.3.438](https://doi.org/10.1093/biomet/33.3.438)
- [35] HILGERT, N., MAS, A. and VERZELEN, N. (2013). Minimax adaptive tests for the functional linear model. *Ann. Statist.* **41** 838–869. [MR3099123 https://doi.org/10.1214/13-AOS1093](https://doi.org/10.1214/13-AOS1093)
- [36] HSING, T. and EUBANK, R. (2015). *Theoretical Foundations of Functional Data Analysis, with an Introduction to Linear Operators*. Wiley Series in Probability and Statistics. Wiley, Chichester. [MR3379106 https://doi.org/10.1002/9781118762547](https://doi.org/10.1002/9781118762547)
- [37] ILLIAN, J., BENSON, E., CRAWFORD, J. and STAINES, H. (2006). Principal component analysis for spatial point processes—assessing the appropriateness of the approach in an ecological context. In *Case Studies in Spatial Point Process Modeling. Lect. Notes Stat.* **185** 135–150. Springer, New York. [MR2232127 https://doi.org/10.1007/0-387-31144-0_7](https://doi.org/10.1007/0-387-31144-0_7)
- [38] JACQUES, J. and PRED, C. (2014). Model-based clustering for multivariate functional data. *Comput. Statist. Data Anal.* **71** 92–106. [MR3131956 https://doi.org/10.1016/j.csda.2012.12.004](https://doi.org/10.1016/j.csda.2012.12.004)
- [39] KARASÖZEN, E., NISSEN, E., BÜYÜKAKPINAR, P., CAMBAZ, M. D., KAHRAMAN, M., KALKAN ER-TAN, E., ABGARMİ, B., BERGMAN, E., GHODS, A. et al. (2018). The 2017 July 20 Mw 6.6 Bodrum–Kos earthquake illuminates active faulting in the Gulf of Gökova, SW Turkey. *Geophys. J. Int.* **214** 185–199.
- [40] KARR, A. F. (1991). *Point Processes and Their Statistical Inference*, 2nd ed. *Probability: Pure and Applied* **7**. Dekker, New York. [MR1113698 https://doi.org/10.1007/978-1-4612-1136-9](https://doi.org/10.1007/978-1-4612-1136-9)
- [41] KINGMAN, J. F. C. (1993). *Poisson Processes*. *Oxford Studies in Probability* **3**. The Clarendon Press, New York. [MR1207584 https://doi.org/10.1017/CBO9780511526756](https://doi.org/10.1017/CBO9780511526756)
- [42] KOLACZYK, E. D. (1999). Wavelet shrinkage estimation of certain Poisson intensity signals using corrected thresholds. *Statist. Sinica* **9** 119–135. [MR1678884 https://doi.org/10.1007/BF02429584](https://doi.org/10.1007/BF02429584)
- [43] LAMBERT, R., TULEAU-MALOT, C., BESSAIH, T., RIVOIRARD, V., BOURET, Y., LERESCHE, N. and REYNAUD-BOURET, P. (2017). Reconstructing the functional connectivity of multiple spike trains using Hawkes models. *J. Neurosci. Methods* **297**. <https://doi.org/10.1016/j.jneumeth.2017.12.026>
- [44] LI, Y. and GUAN, Y. (2014). Functional principal component analysis of spatiotemporal point processes with applications in disease surveillance. *J. Amer. Statist. Assoc.* **109** 1205–1215. [MR3265691 https://doi.org/10.1080/01621459.2014.885434](https://doi.org/10.1080/01621459.2014.885434)
- [45] LI, Y., WANG, N. and CARROLL, R. J. (2013). Selecting the number of principal components in functional data. *J. Amer. Statist. Assoc.* **108** 1284–1294. [MR3174708 https://doi.org/10.1080/01621459.2013.788980](https://doi.org/10.1080/01621459.2013.788980)
- [46] MANTÉ, C., YAO, A.-F. and DEGIOVANNI, C. (2007). Principal component analysis of measures, with special emphasis on grain-size curves. *Comput. Statist. Data Anal.* **51** 4969–4983. [MR2364553 https://doi.org/10.1016/j.csda.2006.08.003](https://doi.org/10.1016/j.csda.2006.08.003)
- [47] MARSOLIER, J., PROMPSY, P. and VALLOT, C. (2022). H3K27me3 conditions chemotolerance in triple-negative breast cancer. *Nat. Genet.* **54** 459–468.
- [48] MEERS, M. P., LLAGAS, G., JANSSENS, D. H., CODOMO, C. A. and HENIKOFF, S. (2023). Multifactorial profiling of epigenetic landscapes at single-cell resolution using MulTI-Tag. *Nat. Biotechnol.* **41** 708–716.
- [49] MOHLER, G. O., SHORT, M. B., BRANTINGHAM, P. J., SCHOENBERG, F. P. and TITA, G. E. (2011). Self-exciting point process modeling of crime. *J. Amer. Statist. Assoc.* **106** 100–108. [MR2816705 https://doi.org/10.1198/jasa.2011.ap09546](https://doi.org/10.1198/jasa.2011.ap09546)
- [50] OAKES, D. (1975). The Markovian self-exciting process. *J. Appl. Probab.* **12** 69–77. [MR0362522 https://doi.org/10.1017/s0021900200033106](https://doi.org/10.1017/s0021900200033106)
- [51] OGATA, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *J. Amer. Statist. Assoc.* **83** 9–27.
- [52] PANARETOS, V. M. and ZEMEL, Y. (2016). Amplitude and phase variation of point processes. *Ann. Statist.* **44** 771–812. [MR3476617 https://doi.org/10.1214/15-AOS1387](https://doi.org/10.1214/15-AOS1387)
- [53] PICARD, F., RIVOIRARD, V., ROCHE, A. and PANARETOS, V. (2026). Supplement to “PCA for point processes.” <https://doi.org/10.1214/25-AOS2596SUPP>
- [54] RAMSAY, J. O. and SILVERMAN, B. W. (2002). *Applied Functional Data Analysis: Methods and Case Studies*. Springer Series in Statistics. Springer, New York. [MR1910407 https://doi.org/10.1007/b98886](https://doi.org/10.1007/b98886)
- [55] RASMUSSEN, J. G. (2013). Bayesian inference for Hawkes processes. *Methodol. Comput. Appl. Probab.* **15** 623–642. [MR3085883 https://doi.org/10.1007/s11009-011-9272-5](https://doi.org/10.1007/s11009-011-9272-5)

- [56] REISS, P. T. and OGDEN, R. T. (2007). Functional principal component regression and functional partial least squares. *J. Amer. Statist. Assoc.* **102** 984–996. [MR2411660](#) <https://doi.org/10.1198/016214507000000527>
- [57] REYNAUD-BOURET, P. (2003). Adaptive estimation of the intensity of inhomogeneous Poisson processes via concentration inequalities. *Probab. Theory Related Fields* **126** 103–153. [MR1981635](#) <https://doi.org/10.1007/s00440-003-0259-1>
- [58] REYNAUD-BOURET, P. and SCHBATH, S. (2010). Adaptive estimation for Hawkes processes; application to genome analysis. *Ann. Statist.* **38** 2781–2822. [MR2722456](#) <https://doi.org/10.1214/10-AOS806>
- [59] SULEM, D., RIVOIRARD, V. and ROUSSEAU, J. (2024). Bayesian estimation of nonlinear Hawkes processes. *Bernoulli* **30** 1257–1286. [MR4699552](#) <https://doi.org/10.3150/23-bej1631>
- [60] WARD, O. G., WU, J., ZHENG, T., SMITH, A. L. and CURLEY, J. P. (2022). Network Hawkes process models for exploring latent hierarchy in social animal interactions. *J. R. Stat. Soc., Ser. C, Appl. Stat.* **71** 1402–1426. [MR4511116](#) <https://doi.org/10.1111/rssc.12581>
- [61] WILLETT, R. M. and NOWAK, R. D. (2007). Multiscale Poisson intensity and density estimation. *IEEE Trans. Inf. Theory* **53** 3171–3187. [MR2417680](#) <https://doi.org/10.1109/TIT.2007.903139>
- [62] WU, S., MÜLLER, H.-G. and ZHANG, Z. (2013). Functional data analysis for point processes with rare events. *Statist. Sinica* **23** 1–23. [MR3076156](#)
- [63] ZETTL, A. (2005). *Sturm-Liouville Theory. Mathematical Surveys and Monographs* **121**. Amer. Math. Soc., Providence, RI. [MR2170950](#) <https://doi.org/10.1090/surv/121>
- [64] ZHENG, G. X. Y., TERRY, J. M. and BIELAS, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8** 14049.